Dan Zhang
Bin Wei *Editors*

Mechatronics and Robotics Engineering for Advanced and Intelligent Manufacturing

Springer

Lecture Notes in Mechanical Engineering

About this Series

Lecture Notes in Mechanical Engineering (LNME) publishes the latest developments in Mechanical Engineering—quickly, informally and with high quality. Original research reported in proceedings and post-proceedings represents the core of LNME. Also considered for publication are monographs, contributed volumes and lecture notes of exceptionally high quality and interest. Volumes published in LNME embrace all aspects, subfields and new challenges of mechanical engineering. Topics in the series include:

- Engineering Design
- Machinery and Machine Elements
- Mechanical Structures and Stress Analysis
- Automotive Engineering
- Engine Technology
- Aerospace Technology and Astronautics
- Nanotechnology and Microengineering
- Control, Robotics, Mechatronics
- MEMS
- Theoretical and Applied Mechanics
- Dynamical Systems, Control
- Fluid Mechanics
- Engineering Thermodynamics, Heat and Mass Transfer
- Manufacturing
- Precision Engineering, Instrumentation, Measurement
- Materials Engineering
- Tribology and Surface Technology

More information about this series at http://www.springer.com/series/11236

Dan Zhang · Bin Wei
Editors

Mechatronics and Robotics Engineering for Advanced and Intelligent Manufacturing

Springer

Editors
Dan Zhang
Department of Mechanical Engineering,
  Lassonde School of Engineering
York University
Toronto, ON
Canada

Bin Wei
Faculty of Engineering and Applied Science
University of Ontario Institute
  of Technology
Oshawa, ON
Canada

Preface

The 2nd International Conference on Mechatronics and Robotics Engineering, ICMRE 2016, was held in Nice, France, during February 18–22, 2016. The aim of ICMRE 2016 is to provide a platform for researchers, engineers, academics as well as industry professionals from all over the world to present their research results and development activities in the area of mechatronics and robotics engineering. This book introduces recent advances and state-of-the-art technologies in the field of robotics engineering and mechatronics for the advanced and intelligent manufacturing. This systematic and carefully detailed collection provides a valuable reference source for mechanical engineering researchers who want to learn about the latest developments in advanced manufacturing and automation, readers from industry seeking potential solutions for their own applications, and those involved in the robotics and mechatronics industry.

This proceedings volume contains 36 papers that have been selected after review for oral presentation. These papers cover several aspects of the wide field of advanced mechatronics and robotics concerning theory and practice for advanced and intelligent manufacturing. The book contains three parts, the first part focuses on the *Design and Manufacturing of the Robot*, the second part deals with the *Mechanical Engineering and Power System*, and the third part investigates the *Automation and Control Engineering*.

We would like to express grateful thanks to our Program Committee members and Organization Committee members of the 2nd International Conference on Mechatronics and Robotics Engineering, special thanks to the keynote speakers: Prof. Alexander Balinsky, Cardiff University, UK, Prof. Farouk Yalaoui, Université de Technologie de Troyes, France, Prof. Dan Zhang, York University, Canada, and Prof. Elmar Bollin, Offenburg University of Applied Sciences, Germany. We would like to express our deep appreciation to all the authors for their significant contributions to the book. Their commitment, enthusiasm, and technical expertise are what made this book possible. We are also grateful to the publisher for supporting this project and would especially like to thank Arumugam Deivasigamani, Anthony Doyle, and Janet Sterritt for their constructive assistance and cooperation,

both with the publishing venture in general and the editorial details. We hope that the readers find this book informative and useful.

Finally, the editors would like to sincerely acknowledge all the friends and colleagues who have contributed to this book.

Toronto, Canada   Dan Zhang
Oshawa, Canada   Bin Wei
February 2016

Contents

Part I
Design and Manufacturing
of the Robot

Critical Review and Progress of Adaptive Controller Design for Robot Arms

Dan Zhang and Bin Wei

Abstract Recent progress of adaptive control, particularly the model reference adaptive control (MRAC) for robotic arm is illustrated. The model reference adaptive controller design issues that researchers face nowadays are discussed, and its recent methodologies are summarized. This paper provides a guideline for future research in the direction of model reference adaptive control for robotic arms.

Keywords Adaptive control · Robot arm · Model reference approach

1 Introduction

In general terms, the robot control problem is formulated as follows, given a desired trajectory, a mathematical model of the manipulator and its interactions with the environment, find the control algorithm which sends torque commands to the actuators so that the robot can achieve expected motion. Control the robot to perform in a certain way is one of the most challenging problems because the robot mechanism is highly nonlinear, i.e. the robot dynamic equation is expressed by nonlinear dynamics that include couplings between the variables, and also the dynamic parameters of the robot vary with position of the joint variables (when the joint moves). Conventional control methods model the manipulator as uncoupled linear subsystems, these methods can produce satisfactory performances at low speeds, but it is not efficient anymore when used for high speed and high accuracy operations. In order to address the above problem, adaptive control can be relied on. Model reference adaptive approach is most popular and established technique.

Adaptive control is the control method used by a controller which must adapt to a controlled system with parameters which vary, or are initially uncertain. For non-adaptive controller, the controller is designed based on the priori information of the system, i.e. one knows the system and designs the controller (e.g. PID controller)

D. Zhang · B. Wei (✉)
University of Ontario Institute of Technology, Oshawa, ON, Canada
e-mail: Bin.Wei@uoit.ca

gears to that system and assume there is no change in the system. Whereas for the adaptive controller, the controller does not necessary need to depend on previous information of the system, and if there is sudden change in environment, the controller can cope with it to adapt to the changed conditions. If we consider a system that we know its transfer function, we design a fixed classical controller, that controller will remain fixed parameters as long as it applies to the system, so we say that this controller depends on its structure and designed on a priori information, that is non-adaptive controller. However, if the controller is depending on posteriori information, for example, if one is changing the parameters of the controller, because of the changes of the parameters of the system or because of the disturbances coming from the environment, that controller is called adaptive. If the system is subject to unknown disturbances, or the system is expected to undergo changes in its parameters in a way which is not pre-determined from the beginning, in that case we use adaptive control. However, in some cases we know how the system operating condition will change, for example, for an aircraft, we know that the aircraft controller is determined by its altitude and speed, and we expect that aircraft to fly at specific value for altitude and speed, in that case one can design a controller for each expected operating point and we switch between the different controllers, this is called gain-scheduling. In other cases we know that the parameters of the system change, but we know also a range for the change of every parameter, in that case it is possible to design a fixed controller that can cope with different changes of the parameters, and guarantee the stability and performance, this kind of controller is robust controller.

From Fig. 1, one can see that for non-adaptive control, firstly when one needs to improve the performance error, the modelling accuracy will also be increased, secondly it cannot improve itself, and thirdly it is assumed that future will be much like present, ignoring environment changes and change in dynamics. So adaptive controller is needed to address the above problem. Now for the adaptive control, it improves itself under unforeseen and adverse conditions, and it achieves a given system performance asymptotically, it does not trade performance for modelling accuracy, as shown in Fig. 1.

Fig. 1 Adaptive control
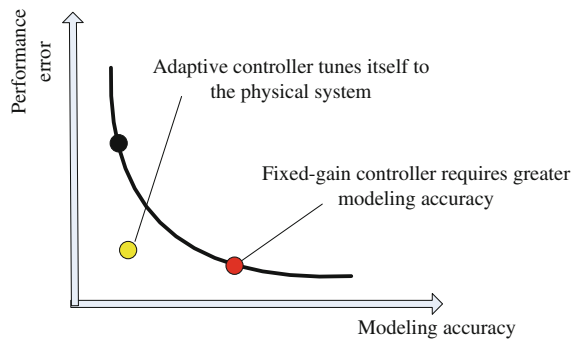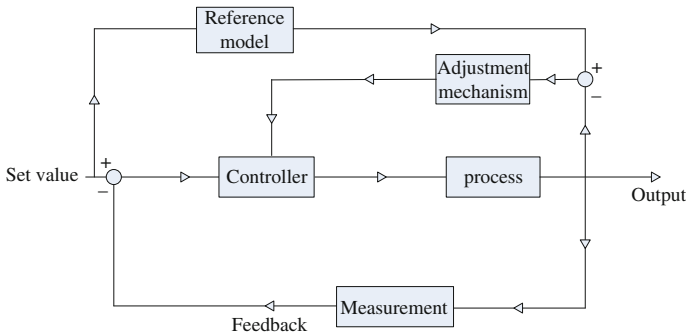
Fig. 2 Diagram of MRAC system

The adaptive control can be categorized into the following, model reference adaptive control, self-tuning adaptive control and gain-scheduled control. With the model-reference adaptive control, an accurate model of the system is developed. The set value is used as an input to both the actual and the model systems, and difference between the actual output and the output from the model is compared. The difference in these signals is then used to adjust the parameters of the controller to minimize the difference, as shown in Fig. 2.

Compared to other control methods, adaptive control is possible to achieve good performance over a wide range of motions and payloads. The advantage of the model reference adaptive control is that the plant parameters need not be fully known, instead, estimates of the plant parameters are used and the adaptive controller utilizes past input/output information to improve these estimates. However there are two drawbacks to MRAC. Stability analysis of the system is critical as it is not easy to design a stable adaptive law. The other problem is that MRAC relies on cancellation of the non-linear terms by the reference model (Sutherland 1987). In reality, exact cancellation cannot be expected, but the non-linear terms may be made so small so as to be negligible. Model reference adaptive control method was initially introduced in Whitaker et al. (1958), when they considered adaptive aircraft flight control systems, using a reference model to obtain error signals between the actual and desired behavior. These error signals were used to modify the controller parameters to attain ideal behavior in spite of uncertainties and varying system dynamics. The goal of an adaptive control system is to achieve and maintain an acceptable level in the performance of the control system in the presence of plant parameter variations. Whereas a conventional feedback control system is mainly dedicated to the elimination of the effect of disturbances upon the controlled variables. An adaptive control system is mainly dedicated to the elimination of the effect of parameter disturbances/variations upon the performance of the control system.

2 General Adaptive Control

In traditional control system, feedback is used to reject the disturbance effect that are acting on the controlled variables in order to bring the controlled variables back to their desired value. To do that, the variables are measured and compared to the desired values and the difference is fed into the controller. In these feedback systems, the designer adjusts the parameters of the controller so that a desired control performance is achieved. This is done by having a priori knowledge of the plant dynamics. When the parameters of the plant dynamic models change with time due to disturbances, the conventional control cannot deal with it anymore as the control performance will be degraded. At this time, one needs to resort to the adaptive control. A structured approach for the design of distributed and reconfigurable control system is presented in Valente and Carpanzano (2011). Distributed architectures are conceived as interconnected independent modules with standard interfaces which can be modified and reused without affecting the overall control structure. Whereas for the centralized control architectures, any change of the machine structure requires an extensive replacement of the control system. In RMS, modular and distributed architecture is essential to guarantee the capability of each single module or portions of the control to be adapted when a hardware reconfiguration occurs. But the paper did not explain in details on how the distributed and adaptive controller have been designed.

In Valentea and Mazzolinib (2015), a control approach is developed which consists of control conceptual design, application development and evaluation of solution robustness. In order to enable the control system reconfiguration, an essential feature of the control architecture is the modularity and distribution of the control decisions across various entities. The control system should be conceived as a set of independent and distributed control modules, capable of nesting one to each other.

The basic concept of adaptive control and several kinds of categories are introduced in Landau (2011), i.e. open-loop adaptive control, direct adaptive control, indirect adaptive control, robust control, and conventional control, etc. The design of a conventional feedback control is oriented to the elimination of the effect of disturbances on the controlled variables, controlled variables are, for examples, temperature if one controls the temperature, position if one controls the position of the end-effector, etc.; whereas the design of adaptive control is oriented to the elimination of effect of parameter disturbances on the performance of the control system. Simply put, the adaptive control can be seen as a conventional feedback control system but where the controlled variable is the performance index. So there are two loops for the adaptive control, one is the conventional feedback loop and the other is the adaptation loop.

The neural networks is used in Wilson and Rock (1995) for the control reconfiguration design for a space robot. The traditional controller was presented, and by using the neural networks, the traditional controller is updated to a reconfigurable controller. Two neural-network-control were developed to achieve quick adaptation

controller. Firstly, a fully-connected architecture was used that has the ability to incorporate an a priori approximate linear solution instantly, this permits quick stabilization by an approximate linear controller. Secondly, a back-propagation learning method was used that allows back-propagation with discrete-valued functions. This paper presents a new reconfigurable neural-network-based adaptive control system for the space robot, but it did not explain in details.

3 Adaptive Control for Robotic Manipulators

Non-adaptive controller designs often ignores the nonlinearities and dynamic couplings between joint motions, when robot motions require high speed and accelerations, it greatly deteriorate its control performance. Furthermore, non-adaptive controller designs requires the exact knowledge and explicit use of the complex system dynamics and system parameters. Uncertainties will cause dynamic performance degradation and system instability. There are many uncertainties in all robot dynamic models, model parameters such as link length, mass and inertia, variable payloads, elasticities and backlashes of gear trains are either impossible to know precisely or varying unpredictably. That is why adaptive control is needed to address the above problem.

Model reference adaptive control and its usage to robotic arms were introduced in Neuman and Stone (1983) and Amerongen (1981). Some design problems in adaptive robot control are briefly stated. Dubowsky and Desforges (1979) is the first one that applies the model reference adaptive control in the robotic manipulator. The approach follows the method in Donalson and Leondes (1963). A linear, second-order, time-invariant differential equation was used as the reference model for each degree of freedom of the manipulator arm. The manipulator was controlled by adjusting the position and velocity feedback gains to follow the model. A steepest-descent method was used for updating the feedback gains. Firstly the reference model dynamics was written, but the paper did not explain how the author had the reference model dynamic equation, subsequently the nonlinear manipulator (plant) dynamic equation was written, but how this equation is related to the Lagrange equation is not clear, thirdly an error function was written and the paper follows the method of steepest descent and derived the a set of equations for the parameter adjustment mechanism, which will minimize the difference between the actual closed-loop system response and the reference model response.

An adaptive algorithm was developed in Horowitz and Tomizuka (1986) for serial robotic arm for the purpose of compensating nonlinear term in dynamic equations and decoupling the dynamic interaction among the joints. The adaptive method proposed in this paper is different from Dubowsky's approach (Dubowsky and Desforges 1979). Three main differences are concluded as follows: firstly, in Horowitz's paper, the overall control system has an inner loop model reference adaptive system controller and an outer loop PID controller, whereas the control system in Dubowsky's method is entirely based on the model reference adaptive

controller; secondly, in Dubowsky's paper, the coupling among joints and non-linear terms in the manipulator equations are ignored whereas this is considered in Horowitz's method; thirdly, in Horowitz's paper, the design method is based on the hyper-stability method whereas the adaptive algorithm design in Dubowsky and Desforges (1979) is based on the steepest descent method.

Model reference adaptive control, self-tuning adaptive control and linear per-turbation adaptive control are briefly reviewed in Hsia (1986). For the model ref-erence adaptive control, the main idea is to synthesize/design a control signal u to the robot dynamic equation, which will force the robot to behave in a certain manner specified by the reference model, and the adaptive algorithm is designed based on the Lyapunove stability criterion.

The MRAC methods presented in Srinivasan (1987) is based on the theory of partitioning control, which makes them capable of compensating for non-linear terms in the dynamic equations and also to decouple the dynamic interactions between the links. It followed and used Horowitz's method (Horowitz 1983) and Sutherland's method (Sutherland 1987). Future research would focus on further simplification of MRAC schemes since the implementation of MRAC methods for the real time control of manipulators has proven to be a challenging task. There is no contribution in this thesis as it just followed and summarized the Horowitz's method and Asare and Wilson's method (Asare and Wilson 1987), and it did not propose its own method or theory.

A MRAC system of 3-DOF serial robotic manipulator was presented in Horowitz (1983), but derivation for the adaptive algorithm is not explained. It was concerned with the application of MRAC to mechanical manipulators. Due to the dynamic equations of mechanical manipulators are highly nonlinear and complex, and also the payload sometimes varies or unknown, the author applied the MRAC to the mechanical manipulators. An adaptive algorithm was developed for com-pensating nonlinear terms in the dynamic equations and for decoupling the dynamic interactions. Finally a 3-DOF serial manipulator was used as computer simulation and the results show that the adaptive control scheme is effective in reducing the sensitivity of the manipulator performance to configuration and payload variations. The core content of Horowitz's method can be concluded as four steps: first step, deterministic nonlinearity compensation and decoupling control. Because one needs to calculate the inertia matrix Mp and nonlinear term V, the second step is pro-posed, i.e. adaptive nonlinearity compensation and decoupling control, which is to adaptively adjust the inertia matrix Mp and nonlinear term V instead of calculating them, and the adaptive algorithm was developed; final step, complete the overall control system by adding the feedback gain Kp, Kv and KI. In Horowitz (1983), it did not entirely use the Landau's hyperstability design (Landau 1979), he used some part of it, and he himself developed the adaptive algorithm. Because according to Hsia (1986), Horowitz's method was separated from the Landau's hyperstability design. And also from Sutherland (1987), it is stated that "While Landau's method replied on a pre-specified parameter matrix for a model and continuous adaptation of the plant parameters, it will be seen later that it is possible to estimate the model parameters and adapt them continuously", from this

statement, it is obvious that Horowitz has his own theory to derive the adaptive algorithm, he did not use Landau's method to derive the adaptive algorithm, but how the adaptive algorithm was derived was not explicitly addressed. In Sutherland (1987), it used the same approach with Horowitz's to a 2-DOF serial robotic manipulator and a flexible manipulator.

In Tomizuka et al. (1986) and Tomizuka and Horowitz (1988), the experiment on the continuous time and discrete time adaptive control on 1-DOF test stand robot arm and Toshiba TSR-500 V robot were briefly conducted. Horowitz et al. (1987) is the continuation of Tomizuka et al. (1986) on a single axis direct drive robotic arm. It applies to a two axis direct drive robotic arm.

In Tomizuka et al. (1985), it presented the experiment evaluation of model reference adaptive controller and robust controller for positioning of a robotic arm under variation of payload. The results show that both method can be insensitive of the payload variation. Four adaptive control methods for the robotic arm were summarized in Jarnali (1989), i.e. computed torque technique, variable structure systems, adaptive linear model following control, and adaptive perturbation control, and the adaptive nonlinear model following control was proposed subsequently, which combines the self-tuning regulator and the model reference adaptive control.

Paper (Sadegh and Horowitz 1987) proposed a modified version of Horowitz's method and the assumption that matrix M and N is constant during adaptation can be removed by modifying the control law and parameter adaptation law. It is demonstrated that by modifying the control law (i.e. making the Coriolis and centripetal acceleration compensation controller a bilinear function of the joint and model reference velocities instead of a quadratic function of the joint velocities) and by modifying the parameter adaptation law (i.e. decomposing the nonlinear parameters in the manipulator dynamic equations into the product of two quantities: one constant unknown quantity, which includes the numerical values of the masses and moments of inertia of the links and the payload and the link dimensions, and the other a known nonlinear function of the manipulator structural dynamics. The nonlinear functions are then assumed to be known and calculable. The parameter adaptation law is only used to estimate the unknown constant quantities), the assumption that matrix M and N is constant during adaptation can be removed. Finally the stability of the above adaptive control law is proved. The above called "exact compensation adaptive control law (ECAL)". In the conclusion, the author found that in order to implement the adaptive controller, one needs to calculate the elements of W(xp, xv, xv) (Sadegh and Horowitz 1987), this procedure is excessively time consuming since it involves computations of highly nonlinear functions of joint position and velocities, to overcome this difficulty, later in Sadegh and Horowitz (1990) and Sadegh (1987), he proposed further modified version. The modification consists in utilizing the desired joint positions and velocities in the computation of the nonlinearity compensation controller and the parameter adaptation law instead of the actual quantities, this is known as "desired compensation adaptive control law (DCAL)" The above whole modification process is shown in Fig. 3.

Fig. 3 Modification process
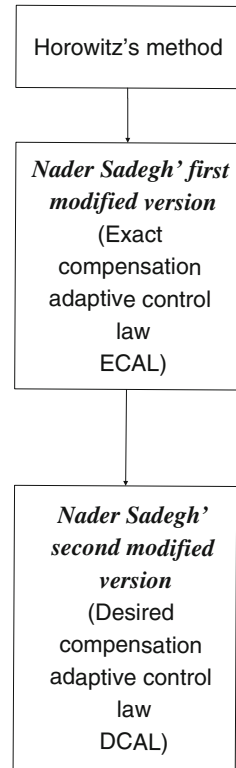
Nader Sadegh applied Craig's method (Craig et al. 1986) to the Horowitz's method, so the condition M and N assumed constant during adaptation can be removed.

Craig's method is re-parametrization, i.e. decompose the manipulator dynamic equation's nonlinear parameters into the product of two quantities: one constant unknown quantity, which includes the numerical values of the masses and moments of inertia of the links and the payload and link dimensions, and a known nonlinear function of the manipulator structural dynamics. The nonlinear functions are assumed to be known and calculable. The parameter adaptation law is only used to estimate the unknown constant quantities.

One method of reparametrizing the manipulator's dynamic equations consists in decomposing each element of the matrices M(x), N(x)'s and the vector g(x) into products of unknown constant terms and known functions of the joint displacement vector. Or a second method consists in the re-parametrization of dynamic equation into the product of unknown constant vector, and a matrix formed by known functions of joint position.

4 Conclusion

Recent progress of model reference adaptive control for robotic arm is presented. The model reference adaptive controller design issues are discussed, and its recent methodologies are summarized. This paper provides a guideline for future research in the direction of model reference adaptive control for robotic arms.

References

Amerongen, J. (1981). MRAS: Model reference adaptive systems. *Journal A, 22*(4), 192–198.

Asare, H., & Wilson, D. (1987). Evaluation of Three Model reference adaptive control algorithms for robotic manipulators. Proceedings of IEEE International Conference on Robotics and Automation. pp. 1531–1542.

Craig. J. J., Hsu. P., & Sastry. S. S. (1986). Adaptive control of mechanical manipulators. In *Proceedings of the 1986 IEEE International Conference on Robotics and Automation*, San Francisco, April.

Donalson, D., & Leondes, T. (1963). A model referenced parameter tracking technique for adaptive control systems. *IEEE Transactions on Applications and Industry, 82*(68), 241–252.

Dubowsky, S., & Desforges, D. (1979). The application of model-referenced adaptive control to robotic manipulators. *Journal of Dynamic Systems, Measurement, and Control, 101*, 193–200.

Horowitz, R. (1983). Model reference adaptive control of mechanical manipulators. PhD thesis, University of California.

Horowitz, R., & Tomizuka, M. (1986). An adaptive control scheme for mechanical manipulators—Compensation of nonlinearity and decoupling control. *Journal of Dynamic Systems, Measurement, and Control, 108*(2), 1–9.

Horowitz, R., Tsai, M. C., Anwar, G., & Tomizuka, M. (1987). Model reference adaptive control of a two axis direct drive manipulator arm. In *Proceedings of 1987 IEEE International Conference on Robotics and Automation*, pp. 1216–1222.

Hsia, T. (1986). Adaptive control of robot manipulators—A review. In *Proceedings of 1986 IEEE International Conference on Robotics and Automation*, pp. 183–189.

Jarnali, H. (1989). Adaptive control methods for mechanical manipulators: A comparative study. Master thesis, Naval Postgraduate School.

Landau, I. D., et al. (2011). "Introduction to Adaptive Control" in adaptive control, communications and control engineering. Springer-Verlag London Limited.

Landau, Y. (1979). *Adaptive control—The model reference approach*. CRC Press.

Neuman, C. P., & Stone, H. W. (1983). MRAC control of robotic manipulators. In K. S. Narendra (Ed.), *3rd Yale Workshop on Applications of Adaptive Systems Theory*. Yale University, New Haven, CT, pp. 203–210.

Sadegh, N. (1987). Adaptive control of mechanical manipulators: Stability and robustness analysis. PhD thesis, University of California, 1987.

Sadegh, N., & Horowitz, R. (1987). Stability analysis of an adaptive controller for robotic manipulators. In *Proceedings of 1987 IEEE International Conference on Robotics and Automation*, pp. 1223–1229.

Sadegh, N., & Horowitz, R. (1990). Stability and robustness analysis of a class of adaptive controllers for robotic manipulators. *International Journal of Robotics Research, 9*(3), 74–92.

Srinivasan, R. (1987). Adaptive control for robotic manipulators. Master thesis, Carleton University.

Sutherland, J. (1987). Model reference adaptive control of a two link manipulator. Master thesis, Carleton University.

Tomizuka, M., & Horowitz, R. (1988). Implementation of adaptive techniques for motion control of robotic manipulators. *Journal of Dynamic Systems, Measurement, and Control, 110*(1), 62–69.

Tomizuka, M., Horowitz, R., & Anwar, G. (1986). Adaptive techniques for motion controls of robotic manipulators, in Japan—USA Symposium on Flexible Automation, Osaka, Japan, pp. 217–224.

Tomizuka, M., Horowitz, R., & Teo, C. L. (1985). Model reference adaptive controller and robust controller for positioning of varying inertia. In *Proceedings of Conference on Applied Motion Control*, University of Minnesota, pp. 191–196.

Valente, A., Carpanzano, E., & Brusaferri, M. (2011). Design and implementation of distributed and adaptive control solutions for reconfigurable manufacturing systems. In *CIRP Sponsored ICMS*. International Conference on Manufacturing Systems.

Valentea, A., Mazzolinib, M., & Carpanzanoa, E. (2015). An approach to design and develop reconfigurable control software for highly automated production systems. *International Journal of Computer Integrated Manufacturing*, *28*(3), 321–336.

Whitaker, H. P., Yamron, J., & Kezer, A. (1958). *Design of model reference adaptive control systems for aircraft*. Report R-164, Instrumentation Laboratory, M. I. T. Press, Cambridge, Massachusetts.

Wilson, E., & Rock, S. (1995). Reconfigurable control of a free-flying space robot using neural networks. In *Proceedings of the 1995 American Control Conference*, *2*, 1355–1359.

Stiffness Analysis of a Planar 3-RPS Parallel Manipulator

Bo Hu, Chunxiao Song and Bo Li

Abstract This paper studied the stiffness model and characteristics of a planar 3-RPS PM with 3-DOF. The 6 × 6 form stiffness matrix of the planar 3-RPS PM is derived with both active and constrained wrenches considered. To characteristic the stiffness of the planer 3-RPS PM, two decomposition methods including the eigenscrew decomposition and the principle axes decomposition are applied to the stiffness matrix. The stiffness matrix decomposition provides a physical interpretation and allows the identification of the compliant axes of the planar 3-RPS PM.

Keywords Planar parallel manipulator · Stiffness · Eigenscrew decomposition · Principle axes decomposition · Compliant aixs

1 Introduction

In recent years, the planar 3 degree of freedom (DOF) parallel manipulators (PMs) have attracted much attention (Angeles 2014). Merlet et al. (1998) presented some definitions such as constant orientation workspace, reachable workspace and dexterous workspace for the planar PMs. Binaud et al. (2010) compared the sensibility of five 3-DOF planar PMs including the 3-RPR, 3-RPR, 3-RRR, 3-RRR and 3-PRR PMs. Mejia et al. (2015) derived a mathematical closed-form solution to obtain the maximum force with a prescribed moment in 3-DOF planar mechanisms. Kucuk (2009) performed dexterity comparison for seven 3-DOF planar PMs with two kinematic chains using genetic algorithms and indicated that the PPR planar

B. Hu (✉) · C. Song · B. Li
Parallel Robot and Mechatronic System Laboratory of Hebei Province,
Yanshan University, Qinhuangdao 066004, Hebei, China
e-mail: hubo@ysu.edu.cn

B. Hu · C. Song · B. Li
Key Laboratory of Advanced Forging and Stamping Technology and Science
of Ministry of National Education, Yanshan University,
Qinhuangdao 066004, Hebei, China

robot manipulator is the best configuration with the best dexterous maneuverability among the others. Dong et al. (2016) proposed a piezoelectric actuated 3-RPR planar micro-manipulator with orthogonal structure and developed its prototype.

Stiffness analysis plays an important role in design of planar 3-DOF PMs. In this aspect, Gosselin (1990) derived general $n \times n$ stiffness matrix for n-DOF PMs by only considering the elastic deformation of actuator factor. Wu et al. (2010) compared the stiffness performance of 4-RRR, 3-RRR and 2-RRR PMs. Zhao et al. (2007) investigated the stiffness performance of planar parallel 3-RRR mechanism with flexible joints.

Most of the stiffness model of planar PMs only considered the actuator factor while the constraint factors were not considered. Recently, the stiffness model considered both active and constrained wrenches has been established for various spatial lower mobility PMs (Li and Xu 2008; Hu and Lu 2011; Hu et al. 2014). Due to the consideration of constraints, this stiffness model is more suitable for the lower mobility PMs. However, up to now, the stiffness models of planar PMs with both active and constrained wrenches considered have not been studied.

Stiffness characteristic analysis is also an important research content for the planar PMs. To investigated the stiffness characteristics of PMs, some researchers proposed effective approaches for the stiffness matrix decomposition (Loncaric 1987; Huang and Schimmels 2000; Chen et al. 2015). Huang and Schimmels (2000) proposed an alternative synthesis algorithm for realization of an arbitrary spatial stiffness matrices, which has been widely used in stiffness characteristic analysis. Chen et al. (2015) presented an alternative decomposition of stiffness matrices, which can be used in both Plucker's ray and axis coordinates. And the compliant axis proposed by Patterson and Lipkin (1993a) is also a better way to explain the characteristic of stiffness.

For the above reasons, the stiffness model and characteristic of a novel planar 3-RPS PM which have constrained forces is studied in this paper.

2 Stiffness Model of the Planar 3-RPS PM

2.1 Kinematics Description

The planar 3-RPS PM includes a base B, a moving platform m, three identical RPS (revolute joint-active prismatic joint-spherical joint)-type leg. Here, B is a regular triangle with O as its center and A_i ($i = 1, 2, 3$) as its three vertices. m is a regular triangle with o as its center and a_i ($i = 1, 2, 3$) as its three vertices. For the planar 3-RPS PM, the three R joints are perpendicular with B (see Fig. 1).

Let $\perp$ be a perpendicular constraint and $\parallel$ be a parallel constraint. Let $\{B\}$ be a frame O-XYZ attached on B at O, $\{m\}$ be a frame o-xyz attached on m at o. Some geometrical conditions ($X \parallel A_1A_3$, $Y \perp A_1A_3$, $Z \perp B$, $x \parallel a_1a_3$, $y \perp a_1a_3$, $z \perp m$) for O-XYZ and o-xyz are satisfied.
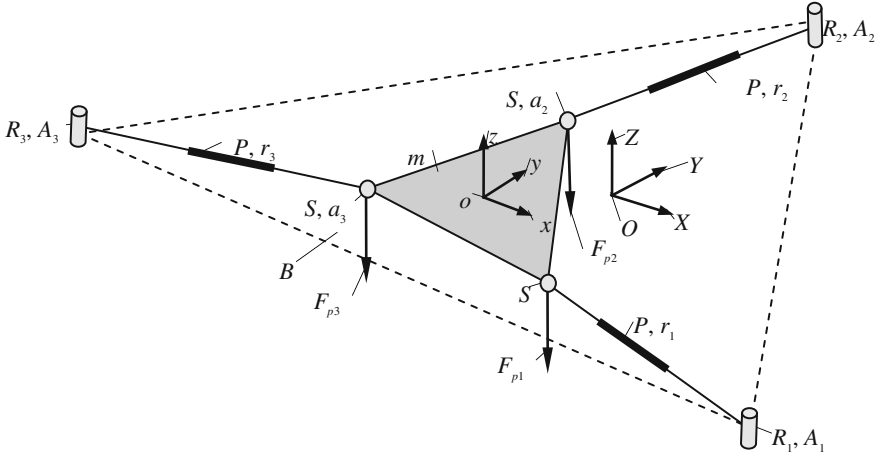
Fig. 1 Sketch of the planar 3-RPS PM

For the planar 3-RPS PM, the unit vectors R_i of R_i ($i = 1, 2, 3$) in $\{B\}$ can be expressed as following:

$$R_1 = R_2 = R_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{1}$$

The position vectors A_i ($i = 1, 2, 3$) of three vertices A_i in $\{B\}$ can be expressed as follows:

$$A_1 = \frac{1}{2}\begin{bmatrix} qL \\ -L \\ 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 \\ L \\ 0 \end{bmatrix}, \quad A_3 = -\frac{1}{2}\begin{bmatrix} qL \\ L \\ 0 \end{bmatrix}, \quad q = \sqrt{3}, \tag{2a}$$

where, L denotes the distance from the center of O to A_i.

The coordinate a_i ($i = 1, 2, 3$) in $\{m\}$ can be expressed as following:

$$^m a_1 = \frac{1}{2}\begin{bmatrix} ql \\ -l \\ 0 \end{bmatrix}, \quad ^m a_2 = \begin{bmatrix} 0 \\ l \\ 0 \end{bmatrix}, \quad ^m a_3 = -\frac{1}{2}\begin{bmatrix} ql \\ l \\ 0 \end{bmatrix} \tag{2b}$$

where, l denotes the distance from the center of o to a_i.

The coordinate a_i in $\{B\}$ can be expressed as following:

$$^B a_i = \begin{bmatrix} X_{ai} \\ Y_{ai} \\ Z_{ai} \end{bmatrix} = {}^B_m\mathbf{R}\,{}^m a_i + o, \quad {}^B_m\mathbf{R} = \begin{bmatrix} c_\alpha & -s_\alpha & 0 \\ s_\alpha & c_\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad {}^B o = \begin{bmatrix} X_o \\ Y_o \\ Z_o \end{bmatrix} \tag{2c}$$

Here, α denotes the angle between B and m.

From Eqs. (2a), (2b) and (2c), the inverse solution can be formulated as following:

$$
\begin{aligned}
r_1^2 &= (qlc_\alpha/2 + ls_\alpha/2 + X_o - qL/2)^2 + (qls_\alpha/2 - lc_\alpha/2 + Y_o + L/2)^2 \\
r_2^2 &= (-ls_\alpha + X_o)^2 + (lc_\alpha + Y_o - L)^2 \\
r_3^2 &= (-qlc_\alpha/2 + ls_\alpha/2 + X_o + qL/2)^2 + (-qls_\alpha/2 - lc_\alpha/2 + Y_o - L)^2
\end{aligned}
\tag{3}
$$

Here r_i (i = 1, 2, 3) is the length of ith leg.

Based on the geometrical approach for determining the constrained forces/torques (Hu et al. 2014), one constrained force F_{pi} (i = 1, 2, 3) which is parallel with R_i and passes through the center of S joint in each RPS type leg can be determined.

As the constrained forces/torques do not work to m, it leads to

$$
\begin{bmatrix} \mathbf{Z}^T & (\mathbf{d}_i \times \mathbf{Z})^T \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \boldsymbol{\omega} \end{bmatrix} = 0, \quad \mathbf{Z} = [\, 0 \ \ 0 \ \ 1\,]^T, \quad \mathbf{d}_i = \mathbf{a}_i - \mathbf{o}
\tag{4a}
$$

where, f_i denotes the unit vector of F_{pi}, $\mathbf{a}_i$ (i = 1, 2, 3) and $\mathbf{o}$ denote the coordinates of a_i and o respect to O, respectively.

From Eq. (4a) and Hu et al. (2014), it leads to

$$
\mathbf{V}_r = \mathbf{J}_{6\times6} \begin{bmatrix} \mathbf{v} \\ \boldsymbol{\omega} \end{bmatrix}, \quad \mathbf{J}_{6\times6} = \begin{bmatrix} \boldsymbol{\delta}_1^T & (\mathbf{d}_1 \times \boldsymbol{\delta}_1)^T \\ \boldsymbol{\delta}_2^T & (\mathbf{d}_2 \times \boldsymbol{\delta}_2)^T \\ \boldsymbol{\delta}_3^T & (\mathbf{d}_3 \times \boldsymbol{\delta}_3)^T \\ \mathbf{Z}^T & (\mathbf{d}_1 \times \mathbf{Z})^T \\ \mathbf{Z}^T & (\mathbf{d}_2 \times \mathbf{Z})^T \\ \mathbf{Z}^T & (\mathbf{d}_3 \times \mathbf{Z})^T \end{bmatrix}, \quad \mathbf{V}_r = \begin{bmatrix} v_{r1} \\ v_{r2} \\ v_{r3} \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{\delta}_i = \frac{\mathbf{a}_i - \mathbf{A}_i}{|\mathbf{a}_i - \mathbf{A}_i|}
\tag{4b}
$$

Here, $\mathbf{v}$ and $\boldsymbol{\omega}$ denote the linear and angular velocities of m, respectively, and $\mathbf{J}_{6\times6}$ is the Jacobian matrix of the planar 3-RPS PM.

2.2 Stiffness Matrix Establishment

Let $\mathbf{F}_o = [F_x \ F_y \ F_z]^T$ and $\mathbf{T}_o = [T_x \ T_y \ T_z]^T$ be the forces and torques applied on m at o, respectively. Let F_{ri} and F_{pi} (i = 1, 2, 3) be the active force and constrained force of r_i, respectively. Using the principle of virtual work, we obtain

$$\mathbf{F}_r^T \mathbf{V}_r + \begin{bmatrix} \mathbf{F}_o^T & \mathbf{T}_o^T \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{\omega} \end{bmatrix} = 0 \tag{5a}$$

Here, $\mathbf{F}_r = [F_{r1} \ F_{r2} \ F_{r3} \ F_{p1} \ F_{p2} \ F_{p3}]^T$.
From Eqs. (4b) and (5a), it leads to

$$\mathbf{F}_r = -(\mathbf{J}_{6\times6}^{-1})^T \begin{bmatrix} \mathbf{F}_o \\ \mathbf{T}_o \end{bmatrix}, \begin{bmatrix} \mathbf{F}_o \\ \mathbf{T}_o \end{bmatrix} = \mathbf{J}_{6\times6}^T \mathbf{F}_r \tag{5b}$$

In the RPS type leg, the active force F_{ri} ($i = 1, 2, 3$) produces a flexibility deformations along r_i and the constrained force F_{pi} ($i = 1, 2, 3$) produces a bending deformation which is perpendicular with r_i.

Let δr_i ($i = 1, 2, 3$) denotes the flexibility deformations along r_i produced by the active force F_{ri}, it leads to

$$F_{r_i} = k_{r_i} \delta r_i, \quad k_{r_i} = \frac{ES_i}{r_i} \tag{6a}$$

Here, E is the modular of elasticity and S_i denotes the ith leg's cross section of RPS type leg.

Let δd_i ($i = 1, 2, 3$) denotes the bending deformation of r_i produced by the constrained forces F_{pi}. It leads to,

$$F_{p_i} = k_{p_i} \delta d_i, \quad k_{p_i} = \frac{3EI}{r_i^3} \tag{6b}$$

where, I is the moment of inertia.
From Eqs. (6a) and (6b), it leads to

$$\mathbf{F}_r = \mathbf{K}_p \begin{bmatrix} \delta\mathbf{r} \\ \delta\mathbf{d} \end{bmatrix}, \quad \delta\mathbf{r} = \begin{bmatrix} \delta r_1 \\ \delta r_2 \\ \delta r_3 \end{bmatrix}, \quad \delta\mathbf{d} = \begin{bmatrix} \delta d_1 \\ \delta d_2 \\ \delta d_3 \end{bmatrix},$$

$$\mathbf{K}_p = \begin{bmatrix} k_{r1} & 0 & 0 & 0 & 0 & 0 \\ 0 & k_{r2} & 0 & 0 & 0 & 0 \\ 0 & 0 & k_{r3} & 0 & 0 & 0 \\ 0 & 0 & 0 & k_{p1} & 0 & 0 \\ 0 & 0 & 0 & 0 & k_{p2} & 0 \\ 0 & 0 & 0 & 0 & 0 & k_{p3} \end{bmatrix} \tag{7}$$

Let $\delta\mathbf{p}$ and $\delta\mathbf{\Phi}$ be the position and orientation deformation of m, respectively. By using the principle of virtual work, the following equation can be derived:

$$F_r^T \begin{bmatrix} \delta r \\ \delta d \end{bmatrix} = -\begin{bmatrix} F_o^T & T_o^T \end{bmatrix} \begin{bmatrix} \delta p \\ \delta \boldsymbol{\Phi} \end{bmatrix} \tag{8}$$

From Eqs. (5b), (7) and (8), it leads to

$$\begin{bmatrix} F_o \\ T_o \end{bmatrix} = \mathbf{K} \begin{bmatrix} \delta p \\ \delta \boldsymbol{\Phi} \end{bmatrix}, \quad \mathbf{K} = \mathbf{J}_{6\times6}^{\mathrm{T}} \mathbf{K}_p \mathbf{J}_{6\times6} \tag{9}$$

Here, $\mathbf{K}$ is the stiffness matrix of the planar 3-RPS PM.

3 Stiffness Characteristics Analysis

To characteristic the stiffness of the planer 3-RPS PM, the eigenscrew decomposition and the principle axes decomposition approaches are applied to the stiffness matrix. Loncaric (1987) proposed that by using the decomposition, the stiffness matrix can be realized by several parallel simple or screw springs, which is a direct correspondence between the mechanism realization and physical appreciation of a spatial stiffness matrix. In addition, the compliant axis of the planer 3-RPS PM are also studied in this section to reversal the characteristic of this PM.

3.1 The Eigenscrew Decomposition of Stiffness Matrix

The eigenscrew problem mentioned by Patterson and Lipkin (1993a) of the spatial stiffness matrix can be expressed as following:

$$\mathbf{K}\Delta e = \lambda e \tag{10}$$

where λ and the corresponding e are the eigenvalue and eigenvector of $\mathbf{K}\Delta$, respectively. The transformation matrix Δ interchanges the first and last three components of a screw, which can be expressed as following:

$$\boldsymbol{\Delta} = \begin{bmatrix} \mathbf{0}_{3\times3} & \mathbf{I}_{3\times3} \\ \mathbf{I}_{3\times3} & \mathbf{0}_{3\times3} \end{bmatrix} \tag{11}$$

The eigenscrew decomposition proposed by Huang and Schimmels (2000) of spatial stiffness matrix can be expressed as:

$$\mathbf{K} = \sum_{i=1}^{6} k_i w_i w_i^T, \quad k_i = \frac{\lambda_i}{2h_i}, \quad h_i = \frac{1}{2} w_i^T \Delta w_i \tag{12}$$

where, spring wrench w_i is the unitization of e_i ($i = 1, \ldots, 6$), h_i is the pitch of w_i and w_i can be defined as:

$$w_i = \begin{bmatrix} n_i \\ \rho_i \times n_i + h_i n_i \end{bmatrix} \tag{13}$$

Here, n_i and ρ_i ($i = 1, \ldots, 6$) are the direction and position vectors of the ith spring, respectively.

3.2 The Principle Axes Decomposition of Spatial Stiffness Matrix

In the principle axes decomposition (Chen et al. 2015), the wrench $\underline{F}$ and $\delta\underline{P}$ are expressed in axis coordinate. The relation between ray and axis coordinate can be expressed as following:

$$\underline{F} = \Delta F, \quad \delta\underline{P} = \Delta\delta P \tag{14}$$

From Eq. (11), it leads to

$$\Delta\Delta = \mathbf{E} \tag{15}$$

here $\mathbf{E}$ is an identity matrix.

The relation of stiffness matrices between these two systems can be derived from Eqs. (14) and (15) as following,

$$\underline{\mathbf{K}} = \Delta \mathbf{K} \Delta = \begin{bmatrix} \underline{\mathbf{A}} & \underline{\mathbf{B}} \\ \underline{\mathbf{B}}^T & \underline{\mathbf{C}} \end{bmatrix} \tag{16}$$

where the symmetric 3×3 block matrices $\underline{\mathbf{A}}$ and $\underline{\mathbf{C}}$ denote the rotational and translational parts, and $\underline{\mathbf{B}}$ denote the coupling part.

$\underline{\mathbf{K}}$ can be represented in a reduced form $\underline{\mathbf{K}}_O$ by applying a pure rotation $\mathbf{R} = \mathbf{Q}^T$ to the current frame in order to translate $\underline{\mathbf{C}}$ to a diagonal form $\underline{\mathbf{C}}_O$, where $\mathbf{Q}$ represents a 3×3 orthogonal matrix whose columns are just the eigenvectors of $\underline{\mathbf{C}}$. Then the stiffness matrix can be decomposed into two sets of rank-1 symmetric stiffness matrices as following (Chen et al. 2015):

$$\underline{\mathbf{K}}_O = \begin{bmatrix} \underline{\mathbf{A}}_O & \underline{\mathbf{B}}_O \\ \underline{\mathbf{B}}_O^T & \underline{\mathbf{C}}_O \end{bmatrix} = \underline{\mathbf{K}}_{OS} + \underline{\mathbf{K}}_{OT} = \sum_{i=1}^{3} k_i w_i w_i^T + \sum_{j=4}^{6} k_j w_j w_j^T,$$

$$\underline{\mathbf{A}}_O = \mathbf{Q}^T \underline{\mathbf{A}} \mathbf{Q}, \quad \underline{\mathbf{B}}_O = \mathbf{Q}^T \underline{\mathbf{B}} \mathbf{Q}, \quad \underline{\mathbf{A}}_{OT} = \underline{\mathbf{A}}_O - \mathbf{B}_O \underline{\mathbf{C}}_O^{-1} \underline{\mathbf{B}}_O^T, \tag{17}$$

$$w_i = \begin{bmatrix} \frac{1}{k_i} b_i^T & e_i^T \end{bmatrix}^T, \quad w_j = \begin{bmatrix} \alpha_j^T & \mathbf{0}_{3\times1}^T \end{bmatrix}^T,$$

where, $\underline{\mathbf{K}}_{OS}$ and $\underline{\mathbf{K}}_{OT}$ are the principal components corresponding to the screw and torsional springs, respectively. k_i ($i = 1, 2, 3$) and k_j ($j = 1, 2, 3$) are the ith eigenvalue of $\underline{\mathbf{C}}$ and $\underline{\mathbf{A}}_{OT}$, respectively. e_i ($i = 1, 2, 3$) denotes the unit vector associated with the coordinate axis of $\{O\}$, namely $e_1 = [1, 0, 0]^T$, $e_2 = [0, 1, 0]^T$, $e_3 = [0, 0, 1]^T$, b_i represents the ith column of $\underline{\mathbf{B}}_O$, a_i represents the ith eigenvector of $\underline{\mathbf{A}}_{OT}$, and w_i ($i = 1, 2, 3$) is the ith wrench-compliant axis of this elastic system.

From Eq. (17), any spatial stiffness matrix can be uniquely realized by three screw and three torsional springs connected in parallel, and the screw springs and torsional springs are orthogonal to each other, respectively.

Let $\{C\}$ be a frame $C\text{-}X_QY_QZ_Q$ with the direction of X_Q, Y_Q and Z_Q-axis are along each row of Q, respectively. Then $\underline{\mathbf{K}}$ can be expressed in $\{C\}$ as following:

$$\underline{\mathbf{K}}_C = \begin{bmatrix} \underline{\mathbf{A}}_* & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \underline{\mathbf{B}}_*\underline{\mathbf{C}}\underline{\mathbf{B}}_* & \underline{\mathbf{B}}_*\underline{\mathbf{C}} \\ \underline{\mathbf{C}}\underline{\mathbf{B}}_* & \underline{\mathbf{C}} \end{bmatrix},$$

$$\underline{\mathbf{A}}_* = \underline{\mathbf{A}} - \underline{\mathbf{B}}\underline{\mathbf{C}}^{-1}\underline{\mathbf{B}}^T, \quad \underline{\mathbf{B}}_* = \frac{1}{2}(\underline{\mathbf{B}}\underline{\mathbf{C}}^{-1} + \underline{\mathbf{C}}^{-1}\underline{\mathbf{B}}^T)$$

(18)

Equation (18) is referred to as the central principle frame, and C is also called the center of stiffness. $\underline{\mathbf{K}}_C$ is the simplest form of the spatial stiffness matrices, which decouples rotational and translational aspects of stiffness to a certain extent. In (18), there only exists three 3×3 symmetric blocks $\underline{\mathbf{A}}_*$, $\underline{\mathbf{B}}_*$, $\underline{\mathbf{C}}$, which correspond to the rotational, coupling and translational parts, respectively.

The homogeneous transformation matrix is given by,

$$g_K = \begin{bmatrix} \mathbf{Q} & p \\ \mathbf{0}_{3\times1}^T & 1 \end{bmatrix}, \quad \widehat{p} = \frac{1}{2}(\underline{\mathbf{B}}_O\underline{\mathbf{C}}^{-1} - \underline{\mathbf{C}}^{-1}\underline{\mathbf{B}}_O)$$

(19)

where p is the coordinate of C respected to the original reference frame $\{B\}$.

Based on the above analysis, the stiffness matrix of planar 3-RPS PM can be decomposed into two sets of three rank-1 symmetric matrices, which can also identify the elastic system's force-deflection behavior of planar 3-RPS PM.

3.3 Compliant Axis and Center of Compliance

For a compliant axis (Patterson and Lipkin 1993b), a force produces a parallel liner deformation and a rotational deformation produces a parallel couple. The compliant axis exists if and only if there are two collinear eigenscrews with eigenvalues of equal magnitude and opposite sign. Thus, not all the elastic system exhibits compliant axes. Wrench-compliant and twist-compliant axes are the basic of a compliant axis hierarchy, and most elastic systems exhibit the wrench-compliant/twist-compliant axes. Wrench-compliant axis exists when a wrench produces a parallel linear deformation, and a twist-compliant axis exits when a twist produces a parallel couple. Such kinds of the force-deflection behavior can be interpreted as following:

$$\begin{bmatrix} f \\ \boldsymbol{\rho} \times f + hf \end{bmatrix} = \mathbf{K}\Delta \begin{bmatrix} 0 \\ k_f f \end{bmatrix}, \begin{bmatrix} 0 \\ k_\delta \boldsymbol{\theta} \end{bmatrix} = \mathbf{K}\Delta \begin{bmatrix} \boldsymbol{\theta} \\ \boldsymbol{\rho} \times \boldsymbol{\theta} + h\boldsymbol{\theta} \end{bmatrix} \qquad (20)$$

where, f is a force, $\boldsymbol{\theta}$ is a rotational deformation, k_f and k_δ are translational and rotational stiffness, h and $\boldsymbol{\rho}$ are the pitch and position vector of a wrench/twist, respectively.

If h is equal to 0, then Eq. (20) is turn into,

$$\begin{bmatrix} f \\ \boldsymbol{\rho} \times f \end{bmatrix} = \mathbf{K}\Delta \begin{bmatrix} 0 \\ k_f f \end{bmatrix}, \quad \begin{bmatrix} 0 \\ k_\delta \boldsymbol{\theta} \end{bmatrix} = \mathbf{K}\Delta \begin{bmatrix} \boldsymbol{\theta} \\ \boldsymbol{\rho} \times \boldsymbol{\theta} \end{bmatrix} \qquad (21)$$

Then the wrench-compliant and twist-compliant axes are turn into force-compliant and rotation-compliant axes, respectively.

If an elastic system has three linearly independent compliant axes and they intersect at the same point, then this point is called the center of compliance. The center of compliance refers to a very specialized concept. At this point, a force passes through the point produces a collinear translation, and a rotation through the point produces a collinear couple (Patterson and Lipkin 1993a). The planar 3-RPS PM has such a center of compliance, which is verified in the last section.

3.4   The Stiffness Characteristics Along Z-Axis

The Jacobian matrix in (4a), the stiffness matrix in (9) and $\mathbf{K}_P$ can be divided into 4 parts, respectively

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_1 & \mathbf{J}_2 \\ \mathbf{J}_3 & \mathbf{J}_4 \end{bmatrix}, \quad \mathbf{K} = \begin{bmatrix} \mathbf{C} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{A} \end{bmatrix}, \quad \mathbf{K}_P = \begin{bmatrix} \mathbf{K}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_p \end{bmatrix},$$

$$\mathbf{K}_r = \begin{bmatrix} k_{r_1} & 0 & 0 \\ 0 & k_{r_2} & 0 \\ 0 & 0 & k_{r_3} \end{bmatrix}, \quad \mathbf{K}_p = \begin{bmatrix} k_{p_1} & 0 & 0 \\ 0 & k_{p_2} & 0 \\ 0 & 0 & k_{p_3} \end{bmatrix} \qquad (22)$$

Here, $\mathbf{J}_i$ ($i = 1, 2, 3, 4$) are 3×3 form matrices.
The relation of these blockings can be expressed as

$$\mathbf{C} = \mathbf{J}_1^T \mathbf{K}_r \mathbf{J}_1 + \mathbf{J}_3^T \mathbf{K}_p \mathbf{J}_3, \mathbf{B} = \mathbf{J}_1^T \mathbf{K}_r \mathbf{J}_2 + \mathbf{J}_3^T \mathbf{K}_p \mathbf{J}_4,$$
$$\mathbf{A} = \mathbf{J}_2^T \mathbf{K}_r \mathbf{J}_2 + \mathbf{J}_4^T \mathbf{K}_p \mathbf{J}_4 \qquad (23)$$

The components of $\boldsymbol{\delta}_i(i = 1, 2, 3)$ along Z-axis are 0, from Eq. (4b), the matrix $\mathbf{C}$ can be expressed as

$$\mathbf{C} = \begin{bmatrix} k_{11} & k_{12} & 0 \\ k_{13} & k_{14} & 0 \\ 0 & 0 & k_{15} \end{bmatrix} \tag{24}$$

k_{1i} (i = 1, 2) is determined by $\mathbf{J}_1$, k_{ri} and k_{pi} (i = 1, 2, 3). From Eq. (24), it can be seen that $\mathbf{C}$ must have one eigenvalue k_{15} ($k_{15} = k_{p1} + k_{p2} + k_{p3}$), and the corresponding eigenvector is $[0\ 0\ 1]^\mathrm{T}$, which is always along Z-axis.

From Eq. (16), the direction of a wrench-compliant axis is determined by the eigenvector of $\underline{\mathbf{C}}$, which is equal to $\mathbf{C}$ in Eq. (22). Obviously, the planar 3-RPS PM always have a wrench-compliant axis along Z-axis. The pith is determined by $\underline{\mathbf{B}}$, which is determined by the configuration of the PM. However, we can certain that a force along Z-axis only produce a linear deformation and will not affect another directions. Then the planar 3-RPS PM has better operation in Z-axis.

4 Numerical Examples

In this section, a 3D assembly manipulator and the finite element (FE) model of the planar 3-RPS PM is established to verify the stiffness model obtained in Sect. 2. Then, one numerical example is provided to characterize the stiffness matrix of planar 3-RPS PM based on the eigenscrew decomposition and principle axes decomposition. In this process, the stiffness matrix is realized by six springs connected in parallel based on two methods, and the compliant axes are obtained through eigenscrew decomposition, which identified the decoupled stiffness matrix.

Set X_0 = 0 m, Y_0 = 0 m, a = 10°, the corresponding length of legs are solved as: $r_1 = r_2 = r_3$ = 0.8158 m, and the stiffness matrix corresponding to this configuration is obtained as following:

$$\mathbf{K} = \begin{bmatrix} -2.7281 & -0.0000 & 0 & 0 & 0 & -0.0000 \\ -0.0000 & -2.7281 & 0 & 0 & 0 & 0.0000 \\ 0 & 0 & 0.0014 & 0.0000 & -0.0000 & 0 \\ 0 & 0 & 0.0000 & 0.0001 & 0.0000 & 0 \\ 0 & 0 & -0.0000 & 0.0000 & 0.0001 & 0.0000 \\ -0.0000 & 0.0000 & 0 & 0 & 0 & 0.0396 \end{bmatrix} \times 10^8 \tag{25}$$

4.1 Finite Element Analysis

In the Solid model of planar 3-RPS PM, S joint is constructed by three R joints (see Fig. 2). Assume a force $\boldsymbol{F}_o = [-20\ -30\ -60]^\mathrm{T}$ N applied on the center of m. The simulated results based on finite element model for the deformation of m are
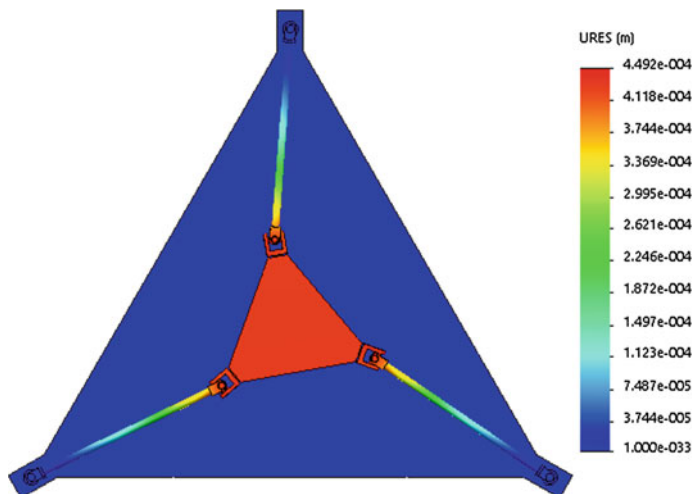
Fig. 2 The simulated result of 3-RPS PM

Table 1 The simulated results based on the finite element model and the theoretical results based on the stiffness model of the deformation of m

The deformation of m (mm)			Error rate (%)
	FE model	Stiffness model	
δx	-7.157×10^{-5}	-7.331×10^{-5}	2.3
δy	-1.113×10^{-4}	-1.110×10^{-4}	0.27
δz	-0.4298	-0.4337	0.89

derived as shown in Fig. 2. The simulated results based on the FE model and the calculated results based on the stiffness model are listed in Table 1.

From Table 1, we can see that the simulated results of FE model are almost equal to the calculated results of the stiffness model. The most error rate is 2.3 %, which is less than 3 %. Then, the FE model verify the correctness of the stiffness model. We also find that the deformation of m in Z-axis is approximate 10^4 times of the deformation in X-axis and Y-axis, which means a small external force in Z-axis would cause relatively large deformation in Z-axis, and this situation has not been mentioned in stiffness analyses of planar PMs in most of previous works. The stiffness model established in this paper presents this situation, which is also appropriate to other planar PMs.

4.2 Decomposition of Planar 3-RPS PM Stiffness Matrix

Applying the eigenscrew decomposition to the stiffness matrix (25), the corresponding six eigenvalue values λ, screw pitches h and the corresponding

Table 2 Parameters of springs based on the eigenscrew decomposition

Spring	k_i	n_i	ρ_i	h_i
1	1.3640×10^8	$[0, 1, 0]^T$	$[0, 0, 0]^T$	0.0055
2	1.3640×10^8	$[0, 1\ 0]^T$	$[0, 0, 0.]^T$	−0.0055
3	1.3640×10^8	$[1, 0, 0]^T$	$[0, 0, 0]^T$	0.0055
4	1.3640×10^8	$[1, 0, 0]^T$	$[0, 0, 0]^T$	−0.0055
5	6.9176×10^4	$[0, 0, 1]^T$	$[0, 0, 0]^T$	5.3470
6	6.9176×10^4	$[0, 0, -]^T$	$[0, 0, 0]^T$	−5.3470

eigenscrews w can be obtained by solving Eq. (12) and the results are shown in Eq. (26) as following:

$$
\begin{aligned}
\lambda &= \mathrm{diag}([\,1.5049 \quad -1.5049 \quad 1.5049 \quad -1.5049 \quad -0.7398 \quad -0.7398\,]) \times 10^6 \\
h &= \mathrm{diag}([\,0.0055 \quad -0.0055 \quad 0.0055 \quad -0.0056 \quad -5.3470 \quad -5.3470\,]) \\
w &=
\begin{bmatrix}
0 & 0 & 1 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 \\
0 & 0 & -0.0055 & 0.0055 & 0 & 0 \\
0.0055 & -0.0055 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 5.3470 & -5.3470
\end{bmatrix}
\end{aligned}
\tag{26}
$$

The parameters of springs based on the eigenscrew decomposition are illustrated in Table 2.

Applying the principle axes decomposition to the stiffness matrix in Eq. (25), the central principal aspects of this spatial stiffness can then be obtained readily and given by,

$$
\underline{\mathbf{A}}_* =
\begin{bmatrix}
0.0083 & 0 & 0 \\
0 & 0.0083 & 0 \\
0 & 0 & 3.9556
\end{bmatrix} \times 10^6, \quad
\underline{\mathbf{B}}_* =
\begin{bmatrix}
0 & 0 & 0.0741 \\
0 & 0 & -0.1672 \\
0.0741 & -0.1672 & 0
\end{bmatrix} \times 10^{-10}
$$

$$
\underline{\mathbf{C}}_* =
\begin{bmatrix}
2.7281 & 0 & 0 \\
0 & 2.7281 & 0 \\
0 & 0 & 0.0014
\end{bmatrix} \times 10^8
\tag{27}
$$

The corresponding homogeneous transformation matrix is given by,

$$
g_K =
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\tag{28}
$$

Table 3 Parameters of springs based on the principle axes decomposition

Spring	k_i	n_i	ρ_i	h_i
1	2.7281×10^8	$[1, 0, 0]^\mathrm{T}$	$[0, 0, 0]^\mathrm{T}$	0
2	2.7281×10^8	$[0, 1, 0]^\mathrm{T}$	$[0, 0, 0]^\mathrm{T}$	0
3	1.3835×10^5	$[0, 0, 1]^\mathrm{T}$	$[0, 0, 0]^\mathrm{T}$	0
4	3.9555×10^6	$[0, 0, 1]^\mathrm{T}$	/	∞
5	8.3012×10^3	$[0, 0, 1]^\mathrm{T}$	/	∞
6	8.3012×10^3	$[0, 1, 0]^\mathrm{T}$	/	∞
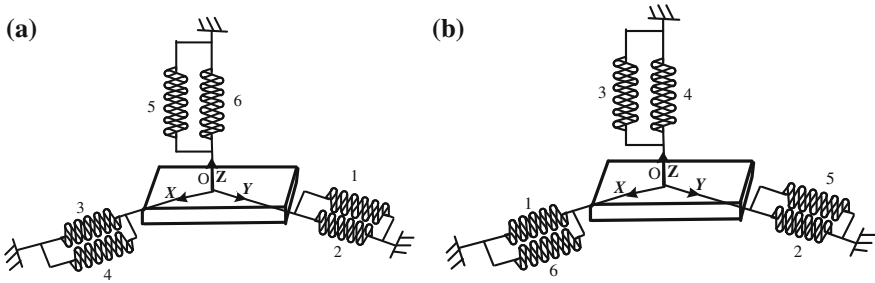
Fig. 3 Physical interpretation of the stiffness matrix based on the eigenscrew decomposition (**a**) and principle axes decomposition (**b**)

In Eq. (27), $\underline{\mathbf{C}}_*$ is a diagonal matrix and the coupling matrix $\underline{\mathbf{B}}$ is turn into a symmetric matrix $\underline{\mathbf{B}}_*$, which is almost equal to null matrix.

The parameters of springs based on the principle axes decomposition are illustrated in Table 3.

The physical interpretation of this stiffness matrix realized by springs based on the eigenscrew decomposition and the principle axes decomposition are shown in Fig. 3a, b.

4.3 Compliant Axes of Planar 3-RPS PM

It can be seen from Fig. 3a that the stiffness matrix is realized by six screw springs. The six screw springs intersect at the coordinate center O. They can be divided into three groups, and each group has two springs. These two springs in each group are collinear and have the same stiffness constants, while opposite in sign. These three group springs are three compliant axes actually, which means the force and deformation about the compliant axes would not affect any other directions. These compliant axes can be expressed as following:

$$w_c = \begin{bmatrix} 0 & 1 & 0 & 0 & 0.0055 & 0 \\ 1 & 0 & 0 & 0.0055 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 5.3470 \end{bmatrix}^T \qquad (29)$$

From Eq. (29), it can be seen that the three compliants are orthogonality and along the direction of Y-, X-, Z-axes, respectively. The three compliant axes also intersect at the same point O, which means that the O is the center of the compliance of the elastic system. It also can be observed that the stiffness matrix is diagonal, which means the stiffness is decoupled in this configuration. In this situation, the stiffness matrix is identify with Class 3b presented by Patterson and Lipkin (1993b). In Class 3b, the elastic system has a pencil of compliant axes and a single compliant axis perpendicular to the pencil. Since $\lambda_1 = \lambda_3$, $h_1 = h_3$, and the eigenscrews corresponding to λ_1 and λ_3 are distributed in X-Y plane, it means that in the X-Y plane, a force (rotational deformation) through the origin produces a linear deformation(couple) parallel to the X-Y plane, and such kind of the force-deflection behavior can be interpreted by

$$[F_x \quad F_y \quad 0 \quad 0 \quad 0 \quad 0] = \lambda_i \mathbf{K\Delta} [0 \quad 0 \quad 0 \quad F_x \quad F_y \quad 0],$$
$$[0 \quad 0 \quad 0 \quad \delta\Phi_x \quad \delta\Phi_y \quad 0] = \lambda_i \mathbf{K\Delta} [\delta\Phi_x \quad \delta\Phi_y \quad 0 \quad 0 \quad 0 \quad 0], (i = 1,3)$$

In Table 3, the pitches of the first three springs are equal to 0. There are three force-compliant axes which are correspond with Eq. (29), they can be expressed as follow:

$$w = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}^T \qquad (30)$$

From Fig. 3b, it can be seen that the stiffness matrix is realized by six simple springs. The first three springs are perpendicularity mutually and intersect at O. The last three springs are perpendicularity mutually and intersect at O, and O is also the center of stiffness of this elastic system. In this configuration, the center of stiffness is degenerate to the center of compliance, which verifies the decoupled characteristic of the stiffness matrix in another way. There are four springs in the X-Y plane and two springs along the Z-axis, which is in accordance with the distribution of screw springs displayed in Fig. 3a. The three pitches of the first three springs are equal to 0, which means the three wrench-compliant axes degenerate to three force-compliant axes. The third spring is along Z-axis which indicated that a force act along Z-axis on the elastic system always only produce a collinear deformation.

5 Conclusions

The main contribution of this paper consists in analyzing the forces/torque situation, deformation and stiffness by considering active forces and constrained torques factors for the planar 3-RPS PM. By considering the constrained forces in each RPS leg, a 6×6 form Jacobian matrix is derived for a planar 3-RPS PM. This 6×6 form Jacobian matrix is used in the stiffness model which leads to a 6×6 form stiffness matrix.

A FE model is established to verify the stiffness model presented in this paper and the comparison results show that the stiffness model is applicable to such kind of planar PMs. And the results also show that the stiffness in Z-axis is much larger than X-axis and Y-axis which cannot be ignored in practical application.

A numeral example is analyzed to reveal the stiffness characteristic of the planar 3-RPS PM by eigenscrew decomposition and principle axes decomposition. The three compliant axes obtained by eigenscrew decomposition show that the stiffness matrix is decoupled in *X-Y* plane. And the compliant axis along Z-axis obtained by eigenscrew decomposition and the force-deformation axis along Z-axis obtained by principle axes decomposition show that a force act along Z-axis on the elastic system always only produce a collinear deformation without affect another direction.

The stiffness analysis modeling of the planar 3-RPS PM in this paper is fit for other planar PMs. This research provides a good reference for the stiffness analysis of the planar PMs.

References

Angeles, J. (2014). *Fundamentals of robotic mechanical systems*. Springer.

Binaud, N., Caro, S., & Wenger, P. (2010). Sensitivity comparison of planar parallel manipulators. *Mechanism and Machine Theory, 45*(11), 1477–1490.

Chen, G. L., Wang, H., Lin, Z. Q., et al. (2015). The principle axes decomposition of spatial stiffness matrices. *IEEE Transactions on Robotics, 31*(1), 191–207.

Dong, Y., Gao, F., & Yue, Y. (2016). Modeling and experimental study of a novel 3-RPR parallel micro-manipulator. *Robotics and Computer-Integrated Manufacturing*, *37*, 115–124.

Gosselin, C. M. (1990). Stiffness mapping for parallel manipulators. *IEEE Transactions on Robotics and Automation*, *b3*, *6*, 377–382.

Hu, B., & Lu, Y. (2011). Solving stiffness and deformation of a 3-UPU parallel manipulator with one translation and two rotations. *Robotica, 29*(6), 815–822.

Hu, B., Mao, B., et al. (2014). Unified stiffness model of lower mobility parallel manipulators with linear active legs. *International Journal of Robotics and Automation, 29*(1), 58–66.

Huang, S., & Schimmels, J. M. (2000). The eigenscrew decomposition of spatial stiffness matrices. *IEEE Transaction on Robotics and Automation, 16*(2), 146–156.

Kucuk, S. (2009). A dexterity comparison for 3-DOF planar parallel manipulators with two kinematic chains using genetic algorithms. *Mechatronics, 19*(6), 868–877.

Li, Y., & Xu, Q. (2008). Stiffness analysis for a 3-PUU parallel kinematic machine. *Mechanism and Machine Theory, 43*(2), 186–200.

Loncaric, J. (1987). Normal forms of stiffness and compliance matrices. *IEEE Journal of Robotics and Automation, 3*(6), 567–572.

Mejia, L., Simas, H., & Martins, D. (2015). Force capability in general 3 DoF planar mechanisms. *Mechanism and Machine Theory, 91*, 120–134.

Merlet, J. P., Gosselin, C. M., et al. (1998). Workspace of planar parallel manipulators. *Mechanism and Machine Theory, 33*(1–2), 7–20.

Patterson, T., & Lipkin, H. (1993a). Structure of robot compliance. *Transactions of the ASME, 115*, 576–580.

Patterson, T., & Lipkin, H. (1993b). A classification of robot compliance. *Journal of Mechanical Design, 115*(3), 581–584.

Wu, J., Wang, J., Wang, L., & You, Z. (2010). Performance comparison of three planar 3-DOF parallel manipulators with 4-RRR, 3-RRR and 2-RRR structures. *Mechatronics, 20*(4), 510–517.

Zhao, T., Zhao, Y., & Shi, L. (2007). Stiffness characteristics and kinematics analysis of parallel 3-DOF mechanism with flexible joints. In *The Proceeding of the IEEE International Conference on Mechatronics and Automation*, pp. 1822–1827.

Overview of an Engineering Teaching Module on Robotics Safety

Dan Zhang, Bin Wei and Marc Rosen

Abstract Robots are widely used in industry. They can perform unsafe, hazardous, highly repetitive and unpleasant tasks for humans. Safety is a very high priority in engineering and engineering education. In this paper, an overview is provided of engineering teaching module on robotic safety developed by the authors. The module covers types of robots, types and sources of robotics hazards, robot safety requirements, robot safeguards and robot safety standards. The importance of safety is highlighted throughout, especially for practical industrial applications. Some new emerging engineering trends and features safety are discussed.

Keywords Robotics · Safety · Engineering · Teaching module

1 Introduction

Industrial robots, unlike humans, can perform complex or mundane tasks without tiring, and they can work in hazardous conditions that would pose risks to humans. Nowadays, industrial robots have been widely introduced to production lines and are expected to find more applications in the future. This is primarily due to the many merits of industrial robots that conventional machines do not possess. For example, robots are increasingly being used in industry to perform such tasks as material handling and welding, and there are around one million robots in use worldwide (Dhillon 2003). However, robots can pose hazardous risks to humans if sufficient precautions are not provided.

Safety is a key factor in industrial and service robot applications, making robotics safety an important subject for engineers. For instance, around 12–17 % of accidents in industries using advanced manufacturing technologies have been reported to be related to automated production equipment, including robots. Robot safety may be interpreted in various ways, including preventing the robot from

D. Zhang · B. Wei (✉) · M. Rosen
University of Ontario Institute of Technology, Oshawa, ON, Canada
e-mail: Bin.Wei@uoit.ca

damaging its environment, particularly the human element of that environment, and simply preventing damage to the robot itself. Without proper precautions, a robot experiencing a fault or failure can cause serious injuries to people and damage equipment in or around a work cell.

Industrial robots are programmable units designed to form expected movements but, unfortunately, the movements of people who work with robots cannot be predicted, making robot safety very important. Most robot-related accidents occur during programming, maintenance, repair, setup and testing. All of these tasks involve human interaction, necessitating proper safety training for employees and the proper use of appropriate safeguards. Note that robots, depending on the task, may generate paint mist, welding fumes, plastic fumes, etc. Also, robots, on occasion, are used in environments or tasks too dangerous for workers, and as such creates hazards not specific to the robot but specific to the task.

Robotics safety operates under a set of principles, primarily related to how to protect humans from robot motions. The principles of robotics safety and the systems to be used when working with robotics are covered in this engineering teaching module.

2 Types of Robots and Industrial Robots

A robot is a mechanical or virtual intelligent agent that can perform tasks automatically or with guidance by remote control. A robot typically has the capacity for sensory input (vision, touch, etc.), recognition and movement, which means a robot should at least have sensors, motors and controllers. There are several types of robots, often differentiated based on function, axis, degree of freedom, workspace, etc. The main types of robots today include, but are not limited to, industrial robots, military robots, medical robots (Speich and Rosen 2004), mobile robots, service robots, and micro and nano robots.

Industrial robots are, multifunctional, mechanical devices, programmable in three or more axes, designed to move material, parts, tools or specialized devices through variable programmed motions to perform a variety of tasks. They have many functions such as material handling, assembly, arc welding, resistance welding, machine tool loading and unloading, etc. An industrial robot system includes not only industrial robots but also related devices and/or sensors required for the robot to perform its tasks, as well as sequencing and monitoring communication interfaces.

2.1 Classifications of Robots

Robots can be classified according to various aspects, such as design configuration, control system, path generation, and others.
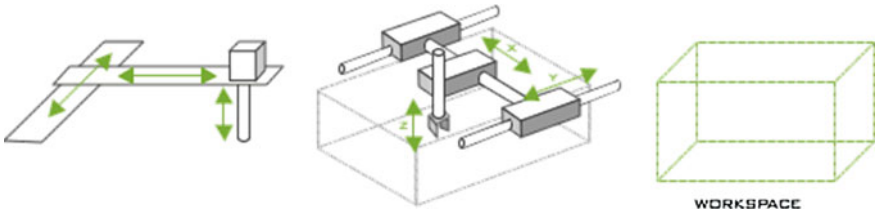
Fig. 1 Cartesian coordinate robot

Fig. 2 Cylindrical robot
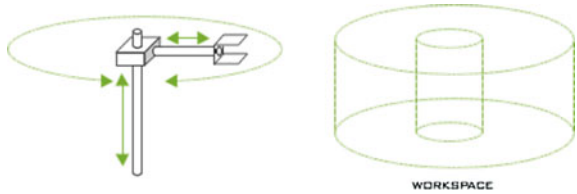
Fig. 3 Spherical/polar robot
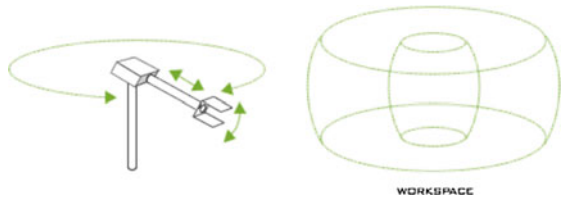
1. Classification based on design configuration

Various robot design configurations, which are included in this classification, are shown in Figs. 1, 2, 3 and 4.

2.2 Cartesian Coordinate Robots

A Cartesian coordinate robot has three linear axes of control (x, y, z), as shown in Fig. 1. Cartesian coordinate robots with the horizontal member supported at both ends are sometimes called gantry robots.

2.3 Cylindrical Robots

A cylindrical robot has two linear axes and one rotary axis around its origin, as shown in Fig. 2.
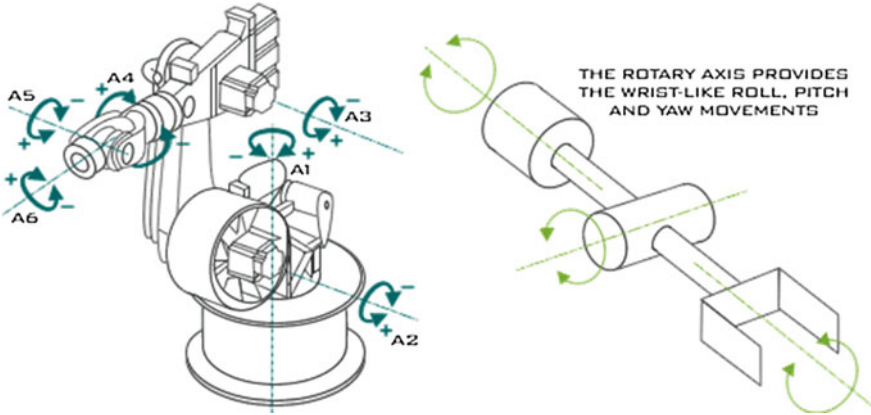
Fig. 4 Articulated/jointed-arm robot

2.4 Spherical/Polar Robots

Spherical and polar robots have one linear axis and two rotary axes, as shown in Fig. 3.

2.5 Articulated/Joint-Arm Robots

Articulated or joint-arm robots are the most versatile of robots. They closely mimic the natural form of the human arm, and can be treated as a shoulder, a bicep and a forearm, but often without a wrist, as shown in Fig. 4. A six-axis robot includes a wrist, providing pitch, roll and yaw movements. With the extra axes of motion provided by the wrist, this robot can deliver the end effector to any point in space in any orientation.

2. Classification based on control systems

All industrial robots are either servo or non-servo controlled, and they can be classified based on their method of control.

(a) Servo robots

Servo robots are controlled through the use of sensors that continually monitor the robot's axes and associated components for position and velocity. This feedback is compared to the desired information which has been programmed and stored in the robot's memory. The motors are actuated to move the actuator to the desired position/speed.

(b) Non-servo robots

Non-servo robots do not have feedback capability, and their axes are controlled through a system of mechanical stops and limit switches.

3. Classification based on path generation

   (a) Point-to-Point path

      Such robots are programmed and controlled to move from one discrete point to another within the robot's working envelope. In the automatic mode of operation, the exact path taken by the robot may vary slightly due to variations in velocity, joint geometries and point spatial locations. This difference in paths is difficult to predict and therefore can create a potential hazards to personnel and equipment.

   (b) Controlled path

      The path or mode of movement ensures that the end of the robot's arm will follow a predictable (controlled) path and orientation as the robot travels from point to point. The coordinate transformations required for this hardware management are calculated by the robot's control system computer. Robots that utilize this type of programming are less likely to present a hazard to personnel and equipment.

   (c) Continuous path

      A robot whose path is controlled by storing a large number or close succession of spatial points in memory during a teaching sequence is a continuous path controlled robot. During this time, the coordinate points in space of each axis are continually monitored on a fixed time base while the robot is being moved, e.g., 60 or more times per second. When the robot is placed in an automatic mode of operation, the program is replayed from memory and a duplicate path is generated.

2.6 Industrial Robot Components

Industrial robots have four major components: mechanical unit, power source, control system and tooling.

   (a) Mechanical unit

      The robot's manipulative arm is the mechanical unit. This mechanical unit is comprised of a fabricated structural frame with provisions for supporting mechanical linkages, joints, guides, actuators (linear or rotary), control valves and sensors. The physical dimensions, design and weight-carrying ability depend on application requirements.

   (b) Power source

      Energy is provided to various robot actuators and their controllers as pneumatic, hydraulic or electrical power. The robot's drives are usually mechanical combinations, and the selection of a power source is usually based on the requirements of a given application, as described below:

      Pneumatic power (usually low-pressure air) is generally used for low weight carrying robots.

Hydraulic power transmission (usually high-pressure oil) is often used for medium to high force or weight applications, or where smoother motion control can be achieved than with pneumatics. Consideration should be given to potential hazards caused by leaks if petroleum-based oils are used.

Electrically powered robots are the most prevalent in industry. Either AC or DC electrical power is used to supply energy to electromechanical actuating mechanisms and their respective control systems. Motion control is much better, and in an emergency an electrically powered robot can be stopped or powered down more safely and quickly than robots driven by either pneumatic or hydraulic power.

(c) Control system

Either auxiliary computers or embedded microprocessors are used for the control of almost all industrial robots today. These perform all of required computational functions, as well as interface with and control associated sensors, grippers, tooling and other associated peripheral equipment. The control system performs the necessary sequencing and memory functions for on-line sensing, branching and integration of other equipment. Programming of the controllers can be done on-line or at remote off-line control stations with electronic data transfer of programs. Self-diagnostic capability for troubleshooting and maintenance greatly reduces robot system downtime. Some robot controllers have sufficient capability in terms of computational ability, memory capacity and input-output capability to serve also as system controllers and handle many other machines and processes. Programming of robot controllers and systems has not been standardized by the robotics industry, so manufacturers usually use their own proprietary programming languages which require special training of personnel.

3 Types and Sources of Robotics Hazards

3.1 Types of Robot Accidents

Robotic safety incidents typically occur when a robotic arm or controlled tool causes an accident and/or places an individual in a risk circumstance, an accessory of the robot's mechanical parts fails, or the power supplies to the robot are uncontrolled (Alvarez 2002). Robotic incidents can be grouped into four categories (Alvarez 2002):

- Impact or collision accidents: Unpredicted movements, component malfunctions or unpredicted program changes related to the robot's arm or peripheral equipment can result in contact accidents.

- Crushing and trapping accidents: A worker's limb or other body part can be trapped between a robot's arm and other peripheral equipment, or the individual may be physically driven into and crushed by other peripheral equipment.
- Mechanical part accidents: A breakdown of the robot's drive components, tooling or end-effector, peripheral equipment, or power source can lead to a mechanical accident. The release of parts, failure of gripper mechanism, or the failure of end-effector power tools (e.g., grinding wheels, buffing wheels, deburring tools, power screwdrivers and nut runners) are a few types of mechanical failures.
- Other accidents: Other accidents can result from working with robots. Equipment that supplies robot power and control represents potential electrical and pressurized fluid hazards. Ruptured hydraulic lines can create dangerous high-pressure cutting streams or whipping hose hazards. Environmental accidents from arc ash, metal spatter, dust, and electromagnetic or radio-frequency interference can also occur. In addition, equipment and power cables on the floor present tripping hazards.

3.2 Sources of Hazards

The potential hazards posed by machines to humans can be assessed and broken down by source/cause, as follows: human interaction, control error, unauthorized access, mechanical failure, and environmental source. Each of these is examined below:

- Human interaction: Hazards from human interaction associated with programming, interfacing activated peripheral equipment, or connecting live input-output sensors to a microprocessor or a peripheral device, can cause dangerous, unpredicted movement or action by a robot. The incorrect activation of the "teach pendant" or control panel is an example of human interaction. One of the greatest hazards is over-familiarity with the robot's redundant motions, so that an individual places herself/himself in a hazardous position while programming the robot or performing maintenance on it.
- Control errors: Intrinsic faults within the control system of the robot, errors in software, and electromagnetic interference are possible control errors. These errors can also occur due to faults in the hydraulic, pneumatic or electrical sub-controls associated with the robot or robot system.
- Mechanical failures: Operating programs may not account for cumulative mechanical part failure, which can allow faulty or unexpected operation to occur.

4 Robot Safety Requirements

Statistics from investigations indicate that the majority of industrial robot incidents and accidents do not take place during normal operation. When working with industrial robots and installations, there is an emphasis on special operation modes, such as commissioning, setting up, programming, testing, checking, troubleshooting and maintenance. In these operating modes, personnel are usually in a danger zone, so safety measures are required to protect them from hazardous events.

4.1 Requirements and Safety Measures in Normal Operation

The use of robot technology necessitates hazard analysis, risk assessment and safety measures. The following examples and suggestions can serve as guidelines (Marty et al. 2011):

- Possible measures to prevent physical access to hazardous areas involving automatic movements include the following: Prevent manual or physical access into danger zones by means of mechanical barriers; use safety devices that respond when approached (light barriers, safety mats), and take care to switch off machinery safely when accessed or entered; permit manual or physical access only when the entire system is in a safe state. For example, this can be achieved by the use of interlocking devices with closure mechanisms on access doors.
- Possible measures to prevent injuries as a result of the release of energy include the following: designs should prevent any release of energy; measures can be added to prevent the release of energy from the danger zone, for example, by a correspondingly dimensioned safety hood; proper lock out and/or tag out procedures need to be utilized to ensure all hazardous energy sources have been isolated.

4.2 Demands and Safety Measures in Special Operation Modes

Certain special operation modes (e.g., setting up, programming) of an industrial robot require movements which must be assessed directly at the site of operation. None of these movements should endanger the persons involved. With this objective in mind, the movements should be only of the scheduled type and speed, prolonged only as long as instructed, and performed only if it can be guaranteed that no parts of the human body are in the danger zone (Marty et al. 2011). One

suggested measure to support this goal involves the use of special operating control systems which permit only controllable and manageable movements using acknowledgeable controls. The speed of movement is thus safely reduced (energy reduction by the connection of an isolation transformer or the use of fail-safe state monitoring equipment) and the safe condition is acknowledged before the control is allowed to activate.

4.3 Demands on Safety Control Systems

An important feature of a safety control system is that the required safety function should be guaranteed as much as possible to work whenever any faults arise. Industrial robots should be almost instantaneously directed by such controllers from a hazardous state to a safe state. Safety control measures should be designed with the following safety objectives in mind (Marty et al. 2011): a fault in the safety control system should not trigger a hazardous state, and a fault in the safety control system should be identified (immediately or at intervals). Suggested measures to provide reliable safety control systems include the following: redundant and diverse layouts of electromechanical control systems, including test circuits; redundant and diverse set-ups of microprocessor control systems developed by different teams. This modern approach is considered state-of-the-art, and often includes safety light barriers; and redundant control systems that take in account mechanical as well as electrical failures.

5 Robot Safeguards

Robot safeguards from design to operation

(a) Risk assessment
   The first step in designing a safe robot system is to understand the hazards that exist in the system (Kelly 2003). The hazards can be classified based on the following criteria: severity, potential injury, frequency of access to the hazard and the possibility of avoidance. Different systems and personnel safeguarding requirements exist at each stage in the development of a robot and robot system. At each stage, a risk assessment should be performed. A robot system design concept, which can account for safety, should include the following elements: limits of the robot system, task identification, user considerations, and hazardous energy control.
(b) Robot safety begins with the design process
   Safeguards should be designed into and around the robotic cell early in the design process to maximize the inherent safety of the overall system. System designers must understand current safeguarding technology. It is also worth

appreciating how this technology can save time and money both now and in the future. To keep unnecessary personnel out of the restricted space of a robot cell, the following two safeguarding methods are often used (Kelly 2003): hard-guarding and optical perimeter guards. Optical perimeter guards are often used in combination with hard-guards. An additional requirement for perimeter safety is that operator control needs to be located outside the safe-guarded area. If there is a danger to the operator, maintenance personnel or other personnel from robotic motion within the restricted or operating space, this area must also be safeguarded. Area safety scanners and light curtains are often used in these areas, as the scanner coverage area is wider and more flexibly programmed than with other devices. Again, these safeguarding devices must be located at a distance that provides adequate stopping time of the system and accounts for the speed of approach from the personnel in the area as well as a depth penetration factor, as defined in the ANSI/RIA R15.06-1999 standard.

For the planning stage, installation and subsequent operation of a robot or robot system, the following need to be considered:

1. Safeguarding devices
   Personnel should be safeguarded from hazards associated with the restricted envelope (space) through use of one or more safeguarding devices, such as: mechanical limiting devices, non-mechanical limiting devices, presence-sensing safeguarding devices, fixed barriers (which prevent contact with moving parts), and interlocked barrier guards.
2. Awareness devices
   Typical awareness devices include chain or rope barriers with supporting stanchions or flashing lights, signs, whistles and horns. These are usually used in conjunction with other safeguarding devices.
3. Maintenance and repair personnel
   When maintenance and repair are being performed, the robot should be placed in the manual or teach mode, and the maintenance personnel should perform their work within the safeguarded area and outside the robot's restricted enve-lope. Additional safeguarding techniques and procedures to protect maintenance and repair personnel are listed in the ANSI/RIA R15.06-1992 standard (Sect. 6.1).

Robot safeguards and engineering applications
The measures taken to safeguard a robot depend on the circumstances of its operation and the surrounding environment.

(a) Today's safeguarding methods

Fences are used to prevent entry to a robot's working space. The fence also pro-vides a preventive barrier against losing parts from the gripper. Barriers are a different type of protection preventing the worker from entering the cell through the

load/unload area. In special cases, like laser cutting, arc welding and water jet applications, the system should be completely encapsulated (Behnisch 2008).

Perimeter fencing

A fixed barrier guard is a fence that requires tools for removal. Barrier guards are appropriate safeguards for full-revolution and part-revolution mechanical power presses. Barrier guards are designed to keep the operator's hands and arms from entering the "danger zone" as prescribed by the particular machine. Barrier guards are usually the first point-of-operation safeguard considered for machines.

Presence sensing devices

Presence detectors are commonly used in robotics safety, and are usually pressure mats and light curtains. Floor mats (pressure-sensitive mats), light curtains (similar to arrays of photocells) and laser scanners can be used to detect a person stepping into a hazardous area near a robot. Effective presence sensing devices stop all motion of the robot if any part of a worker's body enters the protected zone. They are also designed to be fail-safe so that the occurrence of a failure within the device will leave it unaffected or convert it to a mode in which its failed state would not result in an accident.

Manipulator position indication and limiting

In every use of a robot system there is a connection to external safety devices. These connections can control the robot system from external devices like a PLC. But in all cases, a robot system should have a connection to external safety devices to ensure that the robot can be stopped in a safe way.

Other safeguard devices

Various other safeguards can be employed, including the following: trip devices, positive stops, brakes, emergency stop buttons, and general stops.

(b) Instruction to improve robot safety

According to the Occupational Safety and Health Administration (OSHA), most accidents with robots occur during programming, maintenance, repair, setup and testing, all of which involve human interaction. To reduce accidents and improve robot safety, it is important to implement the following industrial robot safety tips: use boundary warning devices, barriers and interlocks around robot systems, provide annual robot safety training for employees working on the floor with robots, provide work cell operators with training geared toward their particular robot, create and implement a preventive maintenance program for robots and work cells, ensure operators read and understand robot system documentation, including that related to robot safety, and allow only capable employees who know the safety requirements for working with a robot to operate robot systems.

   Along with implementing industrial robot safety practices for a facility and its personnel, it is important to ensure the robotic work cell satisfies the following requirements: the maximum reach of a robot should be marked on the floor with safety tape or paint, a flashing warning device must be visible from any point

around the work cell, safety curtains, fences or work cell equipment should be used as barriers around the cell to protect employees, and emergency stop buttons should be located around the cell.

The ISO published new safety standards on July 1, 2011. One notable addition is the new standard for risk assessment. Risk assessment is a process in which one identifies hazards, analyzes or evaluates the risks associated with the hazards, and determines appropriate ways to eliminate or control the hazards. Risk assessment must now be completed when planning and integrating robot systems.

(c) Typical engineering applications

ABB SafeMove

SafeMove is intended to be a major step in removing the restrictions placed on regulated industrial robots that operate in isolated settings, and to represent the next generation in robot safety. Developed and tested to comply with international safety standards, SafeMove is an electronics- and software-based safety approach that ensures safe and predictable robot motion. It also permits operation that is more economic, flexible and lean. Some attributes of SafeMove are as follows:

- Increase man-machine collaboration: SafeMove permit operators and robots to work together more closely, without compromising safety. It uses geometrical and speed restrictions to maintain automatic operation, combining the flexibility of human interaction with the precision and handling capacity of robots.
- Reduce costs of safety devices: SafeMove reduces the need for many safety devices, such as light curtains, safety relays, mechanical position switches and protective barriers, which can in turn reduce installation and maintenance costs.

SafeMove also incorporates electronic position switches, programmable safe zones, safe speed limits, safe standstill positions and an automatic brake test, which allow more flexible safety setups. Programmable safe zones ensure that the robot stays out of protective, three-dimensional zones, which can have complex shapes depending on the installation. Alternatively, the robot can be confined within three-dimensional geometric spaces, reducing robot installation size and floor space and permitting fences to be moved closer to the robot (Kock et al. 2006).

6 Robot Safety Standards

To ensure safety in the workplace, much effort has been expended, especially in the United States and Europe, to codify the safety requirements for humans working around industrial robots. In the U.S., the Robotic Industries Association (RIA) developed the R15.06 robot safety standard through the American National Standards Institute (ANSI). In Europe, ISO brought forth the first edition of ISO 10218 in 1992, which was subsequently adopted by the European Committee for Standardization (CEN) as EN 775. The American documents provide more detailed

information for the integration and use of the robots, while the ISO documents place more emphasis on requirements for robot manufacturers.

Various types of safety regulations apply to robot systems. Today, three different standards for robot safety are used.

6.1 Standard ANSI/RIA R15.06/ANSI/RIA/ISO 10218/RIA TR R15.206

The ANSI/RIA R15.06/ANSI/RIA/ISO 10218/RIA TR R15.206—Industrial Robots Safety Package provides the fundamentals for industrial robots and systems, as it pertains to the safety requirements. The safety requirements are applicable to manufacturers, integrators, installers and personnel. This industrial robots safety package includes: ANSI/RIA R15.06-2012, ANSI/RIA/ISO 10218-1-2007, and RIA TR R15.206-2008.

6.2 Standard CAN/CSA-Z434-03 (R2013)

This safety standard applies to the manufacture, remanufacture, rebuild, installation, safeguarding, maintenance and repair, testing and start-up, and personnel training requirements for industrial robots and robot systems. It was published on 2003-02-01 and reaffirmed on 2013-05-09, and it supersedes standard CAN/CSA-Z434-94,

6.3 Standards ISO 10218-1:2011 and ISO 10218-2:2011

The ISO 10218-1 standard for the robot, and the ISO 10218-2 standard for robot systems and integration, were both published 1 July 2011. ISO 10218-1 was adopted as an ANSI standard. New features in ISO 10218 (not available before) include: cableless pendants—wireless operation, collaborative robots, simultaneous motion control, and synchronous robots. In the United States, ISO 10218 and ANSI RIA 15.06.1999 are both valid. These standards are voluntary, and they provide the integrator and the robot supplier with practical advice on how a setup should be performed in a safe way. For instance, by fulfilling the requirements of ISO 10218, an installation will also fulfill the machinery directive. In other countries, only local bodies exist for the inspection of machine safety and the integration of machines. Standards such as ANSI should be used as a guideline for machine integration. The Robotic Industries Association (RIA) and the Canadian Standards Association now are cooperating to publish a single harmonized standard for the

United States and Canada. The new standard—ANSI/RIA R15.06 in the United States and CAN/CSA Z434 in Canada—will be a "four-in-one" document that includes ISO 10218-1:2011, ISO 10218-2:2011, and the unique requirements of both countries.

7 Conclusion

A teaching module on robotic safety developed by the authors for engineering students is described. Different types of robots, types and sources of robotics hazards, robot safety requirements, robot safeguards and robot safety standards are covered in detail. The importance of safety is highlighted throughout, especially for practical industrial applications. Some new emerging engineering safety trends and features related to robotics are also discussed.

References

Alvarez, M. (2002). Working safely around industrial robots. In *Gateway for safety and health information resources*, http://www.osh.net/articles/archive/osh_basics_2002_may24.htm

ANSI/RIA R15.06-1999, American National Standard for Industrial Robots and Robot Systems—Safety Requirements, National Standards Institute, Inc.

Behnisch, K. (2008). White paper: Safe collaboration with ABB robots electronic position switch and SafeMove.

Changes coming in the new industrial robot safety standard. In Precision Metalforming Association, http://www.metalformingmagazine.com/enterprise-zones/article.asp?aid=6417

Dhillon, B. S. (2003). Robot safety analysis methods. In *Proceedings of the 11th National Conference on Machines and Mechanics*. India, pp. 86–93.

Guidelines for Robotics Safety. In: Occupational Safety and Health Administration (OSHA). U.S. Department of Labor, Washington, DC. https://www.osha.gov/pls/oshaweb/owadisp.show_document?p_table=DIRECTIVES&p_id=1703

Industrial robots and robot system safety. (2013). In OSHA Technical Manual (OTM), Section IV: Chapter 4, Occupational Safety and Health Administration (OSHA), U.S. Department of Labor, Washington, DC. https://www.osha.gov/dts/osta/otm/otm_iv/otm_iv_4.html

Kelly, S. (2003). Robot safety begins with the design process. In *Robotics* Online. http://www.robotics.org/content-detail.cfm/Industrial-Robotics-Featured-Articles/Robot-Safety-Begins-with-the-Design-Process/content_id/1120

Kock, S., Bredahl, J., & Eriksson, P. J. (2006). Taming the robot—Better safety without higher fences. ABB Review 4.

Marty, A., Retsch T., & Schmitter, G. (2011). Safety principles for industrial robots. In 58. Safety Applications, Encyclopedia of Occupational Health and Safety, International Labor

Organization, Geneva, http://www.ilo.org/oshenc/part-viii/safety-applications/item/972-safety-principles-for-industrial-robots?tmpl=component&print=1. Accessed July 2013.

ROBOTS. (2013). http://hopetotheend.com/robots.html

SafeMove-Next generation in robot safety. In ABB. http://www.abb.com/product/seitp327/ec6cfad87f69dd2dc12572d300775f5b.aspx

Speich, J. E., & Rosen, J. (2004). Medical robotics. In *Encyclopedia of biomaterials and biomedical engineering* (pp. 983–993). Marcel Dekker.

Types of Robots. (2013). In loop technology. *Robotics*. http://www.looptechnology.com/robotic-robot-types.asp

Robot Safety. In Loop Technology, *Robotics*. http://www.looptechnology.com/robotic-robot-safety.asp.

Robot Safety. In RobotWorx, http://www.robots.com/articles/viewing/robot-safety

Mobile Robot Applied to QR Landmark Localization Based on the Keystone Effect

Vibekananda Dutta

Abstract This paper proposes a method for the localization of QR (Quick Respond) landmark based on the Keystone Effect (KE) for real-time positioning applications. This is an important aspect, whose solution is still necessary for improving performance in autonomous landmark-based navigation tasks. The paper presents an experimental analysis of the localization of QR landmark and discuss the results which minimize the keystone distortion related to the precision of estimating the landmark's position in a given environment. Through the experiment, the use of proposed algorithm in performing landmark-based robotics action tasks by a Seekur Jr mobile robot is demonstrated.

Keywords QR code · Localization · Keystone effect · Image processing · Mobile robotics action · Landmark recognition

1 Introduction

Keystone distortion produces the effect of virtual twist in an image plane. When the vision system records the image at an oblique angle, it produces the so-called keystone effect (Kang and Lee 2009). In particular, the keystone effect not only produces the virtual rotation in the stationary image plane but also for images of QR landmarks in various autonomous landmark-based robot applications. Mapping the landmark's position is considered one of the foremost navigation tasks in mobile robotics. Recently, QR landmarks have been used extensively because of their capability to store large amounts of information and their resistance against distortion and damage. Naturally, QR landmarks are popular for robots to detect and peruse the information. Although a variety of landmark recognition applications have been proposed which provide users with high recognition speed (Ahn and Lee

V. Dutta (✉)
Institute of Aeronautics and Applied Mechanics, Warsaw University of Technology, ul. Nowowiejska 24, 00-665 Warsaw, Poland
e-mail: vibek@meil.pw.edu.pl

Fig. 1 **a** Source RGB-D
image, **b** thresholded image

2014; Belussi and Hirata 2013). The existing methods typically involve problems
with the localization of the QR landmark, which is appears virtually twisted in an
image plane due to keystone distortion.

Our proposed method improves the localization by evaluating the virtual ori-
entation based on the keystone distortion information, while mapping the land-
mark's position in an environment. Images received from a RGB-D camera
detected landmarks and approximate in their position relative to the robot.
Information received from the robot used to estimate the landmark's position (x, y)
and orientation (roll, pitch, yaw) according to the robot's relative position in the
environment (Li et al. 2011) (Fig. 1).

The method consists of the following steps: (1) adaptive local thresholding is
applied to the RGB-D image to enhance the readability of QR code, (2) the
resulting binary image is processed by the QR code decoding library, (3) the
coordinates of the corners from both left and right edge of the QR code area in an
image are used to evaluate the orientation error of the twisted landmark due to
keystone distortion, (4) the robot pose is then estimated based on the robot current
odometry position, and, (5) the landmark's position with current relative position of
the mobile robot in an environment is mapped (Lee et al. 2011). The experimental
evaluation demonstrates the improvement of the results in indoor mobile robot
action tasks.

In summary, our contribution in this paper are five-fold:

– provide a methodology of QR landmark localization.
– address the problem of Keystone Effect (KE) in an image plane during real-time
  processing of data from a RGB-D camera.
– propose a method to compensate the discrepancy of virtual twist in an image
  plane in real-time applications by using basic geometrical formulations and
  evaluate the rotation angle, which often effects by keystone distortion.
– introduce an algorithm for mobile robot actions in visual-landmarks scenario
  using the proposed method.
– provide extensive analysis of our algorithm on the suggested QR landmarks
  framework and demonstrate how our algorithm can be used by mobile robots.

The rest of the paper is organized as follows. Start with a review of the related work in Sect. 2. Then describe the overview of our methodology in Sect. 3 and describe the scenario in Sect. 4. Introduction of an algorithm for mobile robot actions in Sect. 5. Later on, the paper present the experimental results along with robotics demonstrations in Sect. 6 and finally conclude the paper in Sect. 7.

2 Related Work

Collecting information about the selected features in the environment is a requirement for localization, which can be obtained using various sensors (Skrzypczyński 2009), among which the most popular is vision systems. Nowadays cameras are considered the most appropriate sensors for mapping QR landmarks in the environment for mobile robot navigation tasks, and thus widely used in many research and practical applications (Yudin et al. 2013). However, common visual feature extraction algorithms are very sensitive to unpredictable changes in the environment (Ismail et al. 2012) and thus, the task mobile robot actions based on feature extraction in the environments can fail easily.

Therefore, many practical localization methods use some form of artificial landmarks that provide easily recognizable visual features due to high data compressibility. The simplest approach is to use some predefined basic geometrical shapes (e.g. squares, circles) that can be quickly extracted from the images (Bkaczyk and Kasinski 2010). An example of using barcodes to label many kind of objects (e.g. goods, parcels, parts) in the environments is discussed in (Rostkowska and Topolski 2015). The main advantage of this approach is that the barcodes can encode a lot of additional information. This approach creates an opportunity to embed even more data in the landmark, for example information about the objects the landmarks are attached to (e.g. chair, door, table), properties (e.g. dimension) of these objects is an important aspect that can be explored in order to easily create a semantic description of the environment, which in turn may be very useful information for mobile robot navigation tasks (Borkowski et al. 2010).

3 Localization of QR Landmark

Localization is the key step of QR code recognition. Accuracy in mapping the landmark's position is an important factor of localization (Li et al. 2012). Although many experts and scholars have studied this problem extensively, the problem of virtual rotation, which is visible in the image plane remain (Wang et al. 2009). Therefore, our method focuses on the localization of QR landmark, which is seen in the image as the tilt. First, it is collect the RGB-D images obtained from the vision sensor. The image is then preprocessed using the intensity value as a threshold to get the binary image (Bernsen 1986) and then QR landmark is decoded. Next,

evaluate the virtual orientation error of the rotated QR landmark using keystone distortion information. In this paper, the following subsections are organized as follows: in Sect. 3.1, the paper presents an overview of binarization of QR code. In Sect. 3.2, shows how to extract the QR code area in an image. Finally in Sect. 3.3, using geometrical formulation, the orientation of QR code in an image plane using keystone distortion information is performed.

3.1 Binarization

Binarization is used to improve the recognition of QR code images with non-uniform background and light. In this work, an adaptive local thresholding T^I has been chosen and make the module (this work refer the black and white squares of the QR code as module) into 1, when the value is higher than T^I, otherwise 0, as given below,

$$g^I = \begin{cases} 1, & if \ I(x,y) > T^I \\ 0 & otherwise \end{cases} \tag{1}$$

The g^I stands for the output image I, $I(x, y)$ is the QR code module (x, y) in input image and T^I is the threshold applied in image I. The threshold T^I can be selected as following:

$$T^I = \begin{cases} constant \\ f\left(M_{max}^I, M_{min}^I\right) \\ f\left(M^I\right) \\ f\left(M^I, \delta^I\right) \end{cases} \tag{2}$$

where f is the thresholding function of: (1) $\left(M_{max}^I, M_{min}^I\right)$ is the max. and min. pixel intensity in image I, (2) M^I is the mean value of pixels intensity in image and δ^I is the standard deviation of pixels intensity in image I. In this work,

$$T^I = f\left(M^I, \delta^I\right) \tag{3}$$

is taken as a local threshold (Yang and Shang 2008), where

$$M^I = \frac{1}{b^2} \sum_{(x,y)\epsilon W_{b^2}} I_{(x,y)} \tag{4}$$

$$\delta^I = \sqrt{\frac{1}{b^2} \sum_{(x,y)\epsilon W_{b^2}} \left(I_{(x,y)} - M^I\right)^2} \tag{5}$$
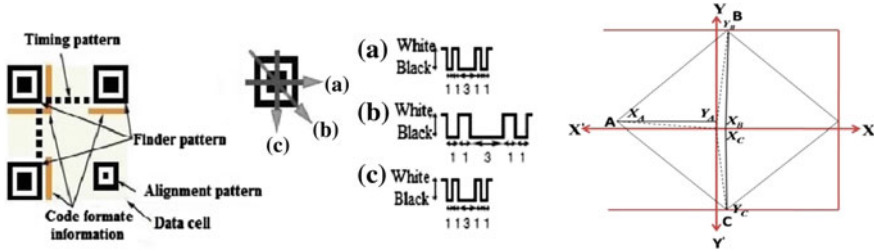
Fig. 2 QR code structure: **a** overview, **b** finder pattern, **c** notation used for QR code extraction method (Ahn and Lee 2014)

The b^2 is the number of pixels in the $b \times b$ square neighborhood region W_{b^2} centered at (x, y).

3.2 QR Code Extraction

QR code extraction in this work is investigates the QR code area in an image. QR extraction consists of: (1) detecting the finder patterns, (2) estimating the code format information, and, (3) detecting of the alignment pattern. It is shown in Fig. 2a, QR landmark contains three finder patterns with same size, one alignment pattern and code format information.

The finder pattern is composed of three overlapped concentric module's, the black and white of the module's ratio is 1:1:3:1:1 along the (A) vertical, (B) diagonal, and (C) horizontal axis (Fig. 2b). First, the QR code symbol must be extracted, and then capture the image region to ensure that the center's coordinate (A, B, C) of the QR code is in the center of the new image (see Fig. 2c).

3.3 Evaluation of Rotation Angle Based on Keystone Effect

Keystone distortion brings trapezoidal distortion in an image plane, which causes both vertical and horizontal parallax in the stereoscopic image (Kang and Lee 2009). Although several solutions for automatic keystone correction have been proposed, these approaches typically only deal with stationary image and do not work for real-time continuous processing of data from a moving camera. The aim of this paper is to introduce a novel approach for evaluating rotation angle using keystone distortion information in real-time. To achieve distortion in real-time, the proposed RGBD-camera system is first calibrated offline, then uses the depth images from the RGBD camera and compute it's relative pose with respect to the camera. Finally remove the keystone distortion by compensating the discrepancy of relative orientation between the camera and the image plane (Petrozzo and Singer 2000).
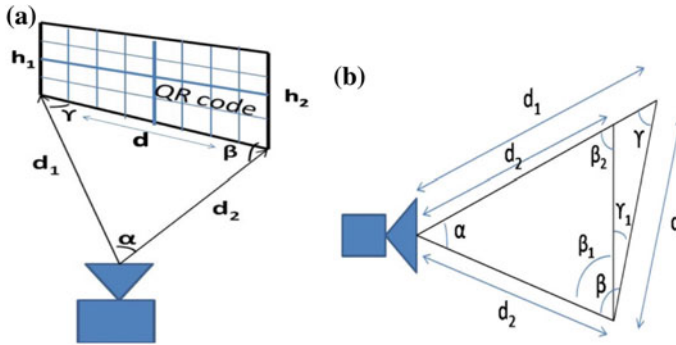
Fig. 3 Geometric model of estimating distortion angle based on keystone effect

Table 1 Distance estimating depends on rotation angle

Angle	Distance at 100 cm		Distance at 150 cm		Distance at 200 cm	
	d_1 (cm)	d_2 (cm)	d_1 (cm)	d_2 (cm)	d_1 (cm)	d_2 (cm)
10°	102	98.2	151.2	148.9	201.2	198.9
15°	103.5	97.8	151.9	148.3	201.9	198.3
20°	104.2	97	152.5	148	202.5	198
25°	105	96.6	152.8	147.6	202.8	197.6
30°	105.3	96.1	153.1	147	203.1	197
35°	106.5	95.4	153.6	146.4	203.6	196.4
40°	107	95.9	154	146	204	196
45°	107.8	94.3	154.4	145.7	204.4	195.7
50°	108.2	93.6	155	145.1	205	195.1

The proposed method consider distances between camera and left/right edge of the QR landmark is to estimate the keystone distortion angle. The method and the geometrical transformation is then used to evaluate the rotation angle of QR landmark in an image plane.[1]

Let us denote the left and right edge by h_1 and h_2 respectively (Fig. 3), and d_1 and d_2 are the distances between left and right edge and the camera plan. In this work (Fig. 3a), it is shown that the edge h_1 is shorter and edge h_2 is longer due to the rotation of the image plane. Using these edges h_1 and h_2, estimate the distances between left/right edge of the QR code and the camera plan in order to obtain the rotation angle γ_1. Table 1 illustrates the experimental analysis of estimating distances (d_1, d_2) between the left/right edge of QR code and the camera plane at a distance 100, 150 and 200 cm respectively (Dutta 2015). Where d = Length of QR code (Fig. 3b). Using distances d_1 and d_2, it is possible to evaluate the rotation angle (γ_1).

---

[1]The idea about the geometrical derivation has been taken from article (Dutta 2015).

$$\alpha = \arccos \frac{d_1^2 + d_2 -^2 d^2}{2.d_1.d_2} \tag{6}$$

Similarly, calculate the angle β. Then, γ is expressed by,

$$\gamma = 180° - \alpha - \beta \tag{7}$$

Thereafter, it is obtained β_1 and β_2 (as shown in Fig. 3 $\beta_1 = \beta_2$). Using Eq. 7, β_1 can be evaluated as,

$$\beta_1 = \frac{180° - \alpha}{2} \tag{8}$$

Finally, β_1 and β are determined, therefore γ_1 can be expressed by,

$$\gamma_1 = \beta - \beta_1 \tag{9}$$

Based on the above calculation, the rotation angle (γ_1) is evaluate, which often appears virtually twisted in an image plane due to keystone distortion.

4 The Necessity of QR Landmark Localization for Mobile Robot Applications

QR landmark localization allows a robot to perform an action in a visual landmarks framework, where a robot is expected to execute a given task. In our experiment, the QR landmarks provide semantic information, which can help to overcome the complexities and limitation of recognition and assist the scene understanding. Semantic knowledge extended by the QR landmarks is obtained in human form of representations in the environment and the robot's actions are studied to show the context-based robot actions practicability.

Aiming at a mobile robot action task, an environment cognition method named as QR landmark-based robotic actions is proposed. The semantic planning approximates the human point of view of robot environments, which enable high-level and more intelligent robot development (Wu et al. 2014). However, the QR code has the merits of storing object's name, other attribute information, high performance of keeping secret and anti-counterfeiting, etc.[2]

The robot reads the QR code and decodes the information immediately in the observed image and establishes a correlation between the observed QR landmark's (i.e., node) and the corresponding robot's relative position (i.e., odometry position). This foundation guarantees the management of environments (see Fig. 4). Thereby

---

[2]A first version of this work was published in "10th Young Scientists Conference", Warsaw, Poland, September 21–23, 2015.

Fig. 4 Environment with QR landmarks

the robot can determine the relation between the QR landmark and the robot relative position in the environment. The author's apply the following formula to the visual landmarks scenario for mobile robot actions,

$$S_{O_i} = (O_1, O_2, O_3, \ldots, O_n) \tag{10}$$

where

$$O_i \in V \tag{11}$$

$$O_i = O_i(g) \tag{12}$$

$$v_j \in V, \quad v_j = v_j(w, g) \tag{13}$$

S is the set of objects, O_i is the set of object's information, V is the set of marked QR codes in the environment, v_j is the number of QR code marked objects, g is the specific property of an object, w is the QR landmark's position in the environment.

5   Mobile Robot Actions Based on QR Landmarks

The action of robot in a visual landmark scenario, a path from the robot's relative position to a considerable QR landmark has to be known that they are considered to be reachable in a given time.

Figure 5 illustrates the geometrical representation of robot's relative position with QR landmark. Assuming, (x_R, y_R) is the position within the world reference

Fig. 5 Geometrical representation of robot position with QR landmark
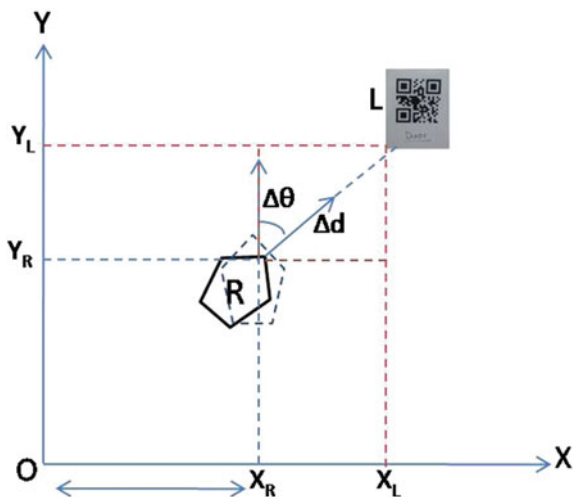
frame in the *xy* plane and θ_R is the orientation with respect to the x-axis, $\theta \in [0, 2\pi]$ (Dutta and Kesswani 2014).

In this work, the author's introduce a minimum distance $d_{\min}$ = 0.3 m, where the robot must stop once it reaches the destination. The proposed algorithm is used to perform the given task in complex indoor laboratory environment with landmarks in different areas, assuming, knowledge of robot's relative position in the environment from the odometry. Upon using robot's relative position, the QR landmark is mapped in the given environment. If (x_L, y_L) is presumed to be landmark's location in the environment and (x_R, y_R) is the robot's current position, then,

$$x_L = x_R + (\sin \theta . \Delta d) \tag{14}$$

$$y_L = y_R + (\cos \theta . \Delta d) \tag{15}$$

Computer vision is used to measure the distance assuming that the size of the QR code is fixed. From the size of the QR landmark seen in the image, the distance Δd between the robot and the QR landmark is obtained.

$$\Delta d = \frac{f * R_h}{I_h} \tag{16}$$

where f is the focal length (cm), R_h is the real height of the QR code (cm), and I_h is the QR code height in the image (pixels). However, in this work, the proportional size of the QR landmark depends on the distances relative to the robot (see Fig. 6). Figure 6 illustrates the images taken by the camera mounted on the robot during the real-time distance measurement (Fig. 7).

Fig. 6 Image of QR code observed at a distance: **a** 1.5 m, **b** 1.0 m, **c** 0.7 m, **d** 0.5 m

Algorithm:
1: **Input**: Image I taken from the corresponding point (current robot position).
2: **Initialization:** d_k = object's information, flag = 0 {specifies non visited landmarks}, Q = 0 {specifies landmarks in the environment}, d_{min}, R = coordinate of the robot reference frame, G = Robot destination point when reaches the target T, L = landmark, T = target {QR code currently targeted}, M = number of landmarks in the field of view.
3: **Output:** Required heading needed to reach G.
4: **Processing Method:** First it is computed the distance between robot (R) and each considered landmark (L) and calculated the needed change of the robot heading $\Delta\theta_i$ respectively.
5: **for** i = 0 to M **do**
6: $\Delta\theta_i$ = evaluated basis on the twisted angle of the rotated QR landmark Eq. 9
7: L = $\min_{i=\{1,...,M\}} |\Delta\theta_i|$ for flag = 0
8: T = L
9: Δd_i = evaluated from the Eq. 15
10: G = Δd_i - d_{min}
11: rotated by $\Delta\theta_i$.
12: move by Δd_i
13: extract(d_k)
14: $flag_L$ = 1
15: introduce robot orientation θ_R clockwise with flag = 0
16: **end for**
17: **while** Q ≠ 0 **do**
18: **for** i = 0 to 2 **do**
19: move forward
20: introduce robot orientation θ_R
21: **end for**
22: **end while**
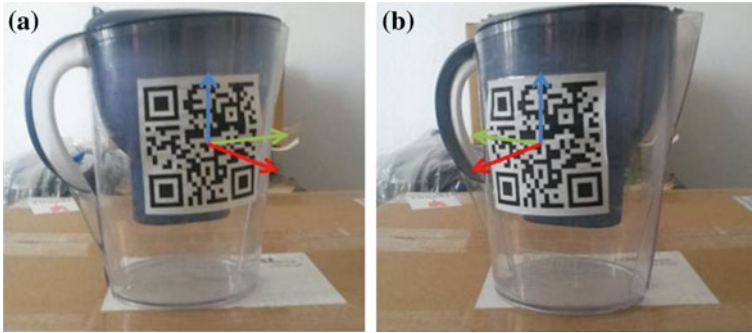
Fig. 7 Image of QR code observed at an angle: **a** 15°, **b** 18°

The proposed algorithm improves the efficiency in landmark recognition with obtaining the information about the objects, which finally speeds up in mobile robot action tasks.

6 Experimental Works and Evaluation

The author's tested 45 samples of QR Code to evaluate the performance of the proposed method. This experiment is carried out under ROS (Robot Operating System) with vision sensor (kinect), using Zbar-library (Brown 2014) to process the QR landmarks for both encoding and decoding.

Table 2 presents the results of angular error testing for the landmark's position with improvement of keystone distortion for distances (d_i) equal to 100, 150 and 200 cm respectively.

Table 2 Average angular errors without keystone effect

Code type	Angle	Orientation error (γ_1)		
		d_{100}	d_{150}	d_{200}
QR	10°	0.3°	0.3°	0.5°
QR	15°	0.3°	0.8°	0.5°
QR	20°	0.5°	0.6°	0.8°
QR	25°	0.4°	0.9°	0.9°
QR	30°	0.55°	0.7°	1.12°
QR	35°	0.5°	0.9°	1.12°
QR	40°	0.45°	0.8°	1.17°
QR	45°	0.7°	1.0°	1.3°
QR	50°	0.6°	0.7°	1.1°
QR	55°	–	–	–
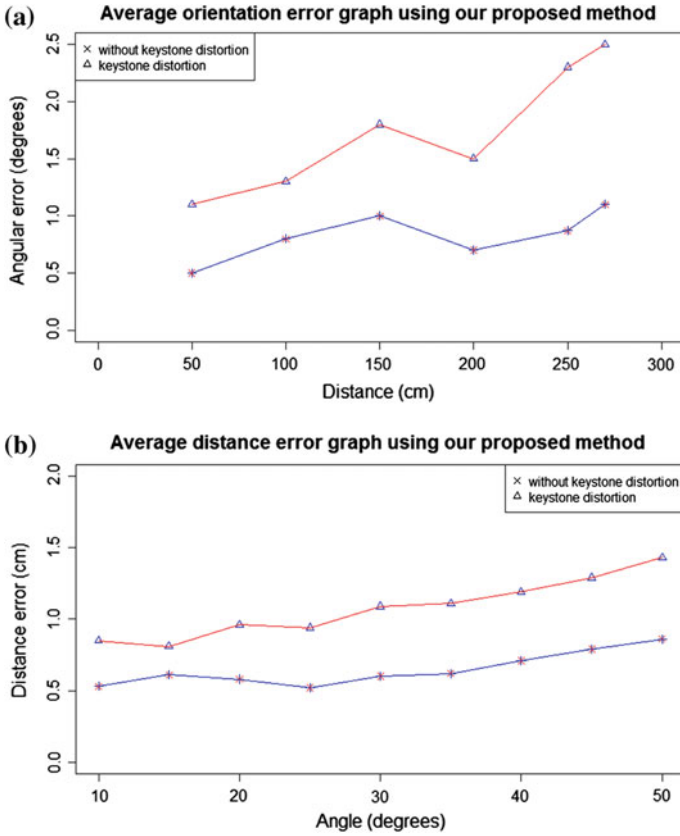
Fig. 8 Result's graph of estimating error in angular orientation and distance measurement based on proposed method. **a** Result's graph of average orientation error. **b** Result's graph of average distance error

Variety range of QR landmarks are collected to evaluate the proposed algorithm on mobile robot action tasks based on visual landmarks framework and an experiment is carried out in the laboratory environment using Seekur Jr mobile robot (LLC 2014). For each action, the localization duration, landmark's position (based on the robot's relative position) and distance has been recorded. Figure 8 illustrates average angular error and distance measurement error during experiment at distances (50, 100, 150, 200, 270 cm).

The experimental results are expressed in terms of robot trajectory to reach the destination (Fig. 10). The distance to reachable way points during the mobile robot action is a parameter as important as the duration of the localization of the QR
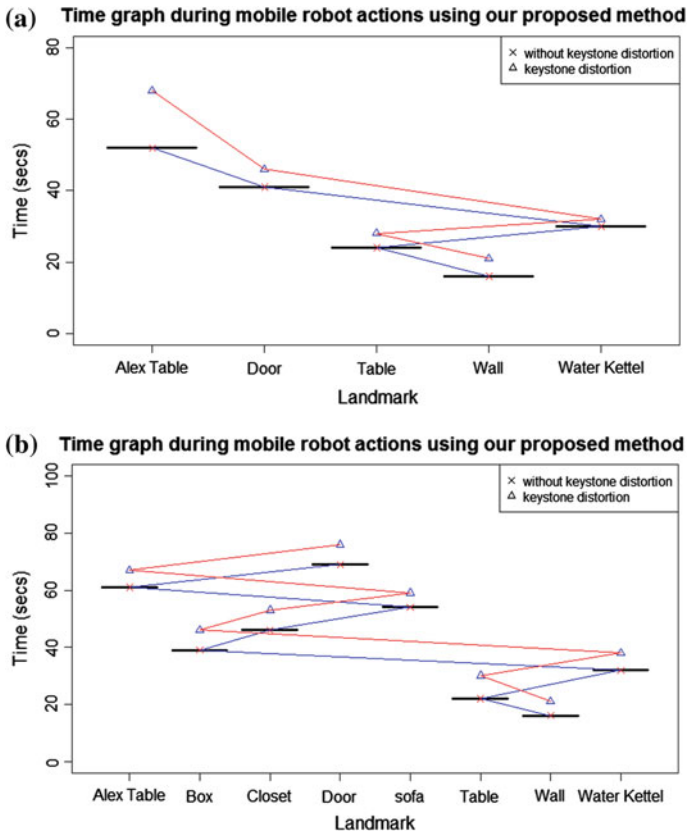
Fig. 9 Navigation duration during the indoor mobile robot action tasks based on the proposed method. **a** Result's graph of duration with 5 objects during mobile robot actions. **b** Result's graph of duration with 8 objects during mobile robot actions

landmark. Indeed if the robot moves far away from the QR landmark reachable points, it increases the risk of collision and falls. To evaluate moves his distance, the actual best path (straight path from robot's current position to QR landmark) followed by the robot is computed. Figures 8 and 9 shows that, the proposed method not only minimizes the angular error but also speeds up in the navigation while mapping the landmark's location in the environment.

The experiment is carried out in an imitation domestic environment in our laboratory. The whole process of mobile robot actions and mapping all the QR landmarks are illustrated in Fig. 11.
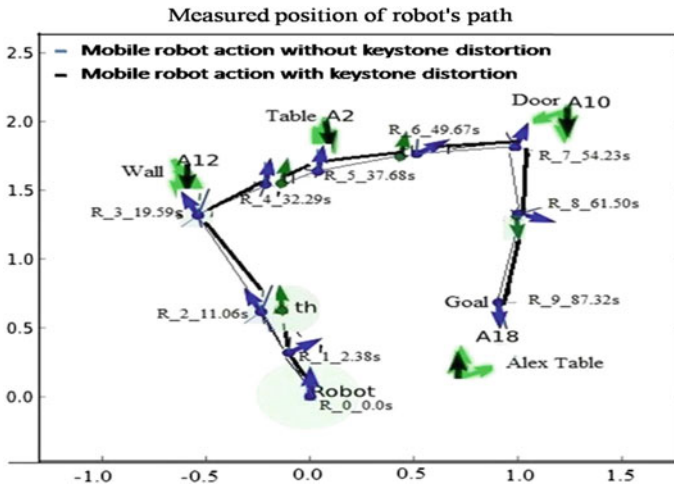
Fig. 10 Robot's path during mobile robot actions. *Black* and *green arrows* indicate the mean position and orientation of the QR landmarks stuck on the objects. *Green circle* defines the observation point. *Black line with green arrows* indicates the mobile robot actions without estimating keystone distortion angle, where as *blue line with arrows* defines the robot trajectory in mobile robot actions using proposed method. All units are in meters

Fig. 11 The whole process of mobile robot actions and mapping QR landmark's location in a given environment

7 Conclusion

This paper introduced the method of QR code localization with evaluating the virtual rotation towards the image plan and the distance to the mobile robot. In this paper, localization is performed with a fully calibrated camera. During the experiment, it was estimated that the maximum angle and the distance from which the

robot can see the QR code are respectively 51° and 270 cm. The experiment with mobile robot in Table 2 and Figs. 8 and 9 confirm reliable localization of QR landmark within the above-stated limits. It is shown that, being able to infer the Keystone Effect (KE) enables the robot to perform the tasks in a more meaningful way. In extensive experiments over a challenging scenario, the results indicate that through the proposed method, QR landmark can be mapped by the mobile robot and serial landmarks can facilitate robot actions in a complex area. The experimental results shows in Fig. 11 is just an one scenario. The experiment with the proposed method has been carried out in different possible complex scenarios and the results shows the improvement in indoor environments.

References

Ahn, M., & Lee, S. (2014). A research on QR code recognition enhancement using pre-constructed image matching scheme. In *2014 International Conference on Information and Communication Technology Convergence (ICTC)* (pp. 82–83). IEEE.

Aleksandrovich, Y. D., Gennadievich, P. G., Stepanovich, K. A., & Zalmanovich, M. V. (2013). Mobile robot navigation based on artificial landmarks with machine vision system. *World Applied Sciences Journal, 24*(11), 1467–1472.

Belussi, L. F., & Hirata, N. S. (2013). Fast component-based QR code detection in arbitrarily acquired images. *Journal of Mathematical Imaging and Vision, 45*(3), 277–292.

Bernsen, J. (1986). Dynamic thresholding of grey-level images. In *International conference on pattern recognition* (pp. 1251–1255).

Bkaczyk, R., & Kasinski, A. (2010). Visual simultaneous localization and map-building supported by structured landmarks. *International Journal of Applied Mathematics and Computer Science, 20*(2), 281–293.

Borkowski, A., Siemiatkowska, B., & Szklarski, J. (2010). Towards semantic navigation in mobile robotics. In *Graph transformations and model-driven engineering* (pp. 719–748). Berlin: Springer.

Brown, J. (2014). Zbar Software. http://zbar.sourceforge.net/. (Online) Accessed November 19, 2014.

Dutta, V. (2015). QR code localization using keystone effect. In *Young scientists towards the challenges of modern technology*, September (pp. 21–23).

Dutta, V., & Kesswani, N. (2014). Designing vision based autonomous docile-x mobile robot for real-time application to soccer behaviors. *Journal of Automation Mobile Robotics and Intelligent Systems, 8*(4), 40–50.

Ismail, A., Jamil, S., Ismail, A. H., Ayob, M., & Rahim, N. A. (2012). A comprehensive study of using 2d barcode for multi robot labelling and communication. *International Journal on Advanced Science, Engineering and Information Technology, 2*(1), 80–84.

Kang, W., & Lee, S. (2009). Horizontal parallax distortion correction method in toed-in camera with wide-angle lens. In *3DTV conference: The true vision-capture, transmission and display of 3D video* (pp. 1–4). IEEE.

Lee, Y.-C., Chae, H., Kim, S.-H., et al. (2011). Applications of robot navigation based on artificial landmark in large scale public space. In 2011 IEEE international conference on robotics and biomimetics (ROBIO) (pp. 721–726). IEEE.

Li, W., Duan, F., Chen, B., Yuan, J., Tan, J. T. C., & Xu, B. (2012). Mobile robot action based on QR code identification. In *2012 IEEE international conference on robotics and biomimetics (ROBIO)* (pp. 860–865). IEEE.

Li, Z., Wong, K.-H., Gong, Y., & Chang, M.-Y. (2011). An effective method for movable projector keystone correction. *IEEE Transactions on Multimedia, 13*(1), 155–160.

LLC, A. M. (2014). Seekur Jr mobile robot. http://www.mobilerobots.com/ResearchRobots/SeekurJr.aspx. (Online) Accessed March 03, 2014.

Petrozzo, R. A., & Singer, S. W. (2000). Cinema projection distortion. *SMPTE Journal, 109*(8), 648–651.

Rostkowska, M., & Topolski, M. (2015). On the application of QR codes for robust self-localization of mobile robots in various application scenarios. In *Progress in automation, robotics and measuring techniques* (pp. 243–252). Berlin: Springer.

Skrzypczyński, P. (2009). Simultaneous localization and mapping: A feature-based probabilistic approach. *International Journal of Applied Mathematics and Computer Science*, *19*(4), 575–588.

Wang, D., Jia, Y., & Cao, Z. (2009). A Unique Multi-functional Landmark for Autonomous Navigation. In *2009 Fifth International Conference on Natural Computation (ICNC'09)* (Vol. 5). IEEE.

Wu, H., Tian, G. H., Li, Y., Zhou, F. Y., & Duan, P. (2014). Spatial semantic hybrid map building and application of mobile service robot. *Robotics and Autonomous Systems*, *62*(6), 923–941.

YANG, S., & SHANG, Z. H. (2008). A New Binarization Algorithm for 2D Bar Code Image [J]. *Journal of Kunming University of Science and Technology (Science and Technology)*, *1*, 011.

A Collective Behaviour Framework for Multi-agent Systems

Mehmet Serdar Güzel and Hakan Kayakökü

Abstract This paper addresses a novel framework that employs a decentralized strategy for collective behaviours of multi-agent systems. The framework proposes a new aggregation behaviour that focusses on letting agents on the swarm agree on attending a group and allocating a leader for each group. As the leader starts moving towards a specific goal in a particularly cluttered environment, other members are enabled to move while keeping themselves coordinated with the leader and the centre of gravity of the group.

Keywords Multi-agent systems · Decentralized architecture · Collective behaviour · Swarm intelligence

1 Introduction

Swarm robotics is a scientific discipline to collective robotics, inspired from the behaviours of social animals. Multi-agent systems, considered to be aggregations of autonomous agents, resembles swarm robotics concept in a certain way (Brambilla et al. 2013). Accordingly, both concepts will be considered together in this study. Collective behaviours of multi-agent systems can be classified into three main groups namely, collective decision making, navigation behaviours and spatially organizing behaviours (Brambilla et al. 2013). Aggregation is one of the fundamental and critical spatial organization that allows a group of robots to get close one other, providing interaction and collective movements (Camazine et al. 2001).

Aggregation behaviour can be observed in nature frequently, such as bacteria, bees, fish and etc. (Camazine et al. 2001; Jeanson et al. 2005). Probabilistic finite state machines (PFSMs) are the main methodology used in aggregation ensuring that finally only a sole aggregate is formed. Each robot starts exploring the environment so as to find other robots. Once other robots are found, it decides whether

M.S. Güzel (✉) · H. Kayakökü
Computer Engineering Department, Ankara University, Ankara, Turkey
e-mail: mguzel@ankara.edu.tr

to join or leave the aggregate in a stochastic manner (Garnier et al. 2005; Soysal and
Şahin 2005, 2007). Alternatively, artificial evolution approach has been employed
to automatically select aggregation behaviour (Soysal et al. 2007). Coordinated
motion, flocking, is a navigation behaviour inspired from fish or flock of birds
(Kaminka et al. 2008). In multi-agent system, coordinated motion approach pro-
vides safer navigation for a group of robots while keeping a constant distance from
one another based on virtual physic-based design. One of the popular studies in this
area proposes a virtual heading sensor allowing each robot to be able to sense the
heading direction of the other robots without requiring a goal direction. Within this
sensor, the swarm could provide coordinated motion while avoiding obstacles
(Turgut et al. 2008). This study was extended and revealed that it is possible to
insert some "informed" robots, knowing the goal direction, in the swarm so as to
lead the other "non-informed" robots towards the goal direction (Ferrante et al.
2010). A novel and recent study in coordinated motion field allows robots to change
both angular and forward speed according to the computed vector without requiring
an explicit alignment rule (Ferrante et al. 2012).

This paper proposes a novel approach based on PFSMs and Virtual
physics-based design to assign each robot into a group and allocate a leader for each
group in a decentralized manner. This is different from the conventional aggrega-
tion behaviour of social animals that instead of forming a single aggregate, robots
are grouped according to their distance to each other and a leader is selected in a
complete decentralized manner. Besides, each group relies on a centre of gravity
(COG) based algorithm and navigates in a coordinated motion towards a specific
goal while avoiding obstacles placed on their paths.

This paper is organized as follows. In Sect. 2, the proposed framework and
corresponding algorithms for multi-agent systems are presented, whereas Sect. 3
focuses on implementation and evaluation of the system. The study is concluded in
Sect. 4.

2 Collective Behavior Framework

This section details the features of the proposed framework used for collective
behaviour of multi-agent systems in a decentralized manner that comprises a
generic grouping algorithms and a leader assignment procedure, followed by a
coordinated navigation strategy. Flowchart of the proposed collective behaviour
framework is illustrated in Fig. 1. As it can be seen from the corresponding figure,
the framework consists of three main modules and two sub-modules, which will be
detailed in the following sections. Essentially each robot possesses a map based
navigation strategy that each robot has the 2-D map of the environment and can
navigate individually towards a specific goal using a local navigation strategy.
Nevertheless, robots are not allowed to communicate each other while performing
grouping, leader selection and coordinated motion behaviors. Accordingly, all these
tasks are achieved in a decentralized manner.

Fig. 1 Flowchart of the collective behaviour framework
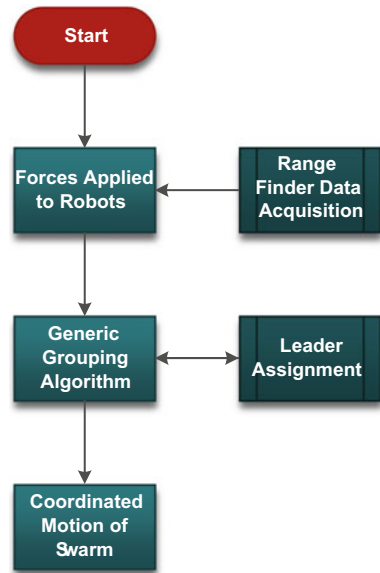
2.1 Forces Applied to Robots

It is assumed that all robots in the swarm are randomly located and there is no direct communication between robots during the execution of the algorithm. Each robot is equipped with a standard range finder, which has 360° field of view. Newton's law of universal gravitation (1) is calculated for all robots in the swarm so as to estimate attractive force applied to each of them. According to the law, it is stated that there occurs a force between two agents that is proportional to the product of their masses and inversely proportional to the square of the distance between them as illustrated in (1).

$$F = G \times \frac{m_1 \times m_2}{r^2} \tag{1}$$

where F is the attractive force, m_1 and m_2 are the masses of each agent, G is the constant and r is the distance between centres of two masses. Total attractive force applied on the ith robot (R_iA) in the swarm is defined as follows:

$$R_iA = \sum_{j=1}^{n} \vec{F}_j \tag{2}$$

2.2 Generic Grouping Algorithm

Randomly located robots in the swarm apply force and attract each other based on (2), as illustrated in Fig. 2. Afterwards, robots start moving in the direction of resultant force in order to approach each other that each robot continues moving until the applied attractive force exceeds a certain limit, as shown in Figs. 2 and 3. Stationary robots wait other robots to approach them until the predefined grouping number, '3' for this example, is reached. This, in essence, provides a decentralized aggregation and grouping behaviour for robots without requiring any direct communication. Once the robots get close each other as request, the proposed algorithm achieves leader assignment in a decentralized manner. According to which, first, each robot calculates the relative position of each of the surrounding robots using range finder sensor. Next, each robot considers itself as the origin of its local

Fig. 2 Resultant force vectors of robots
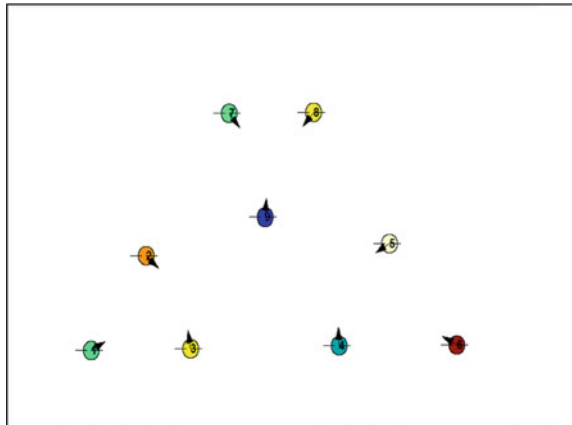
Fig. 3 Robots approach to each other using force vectors
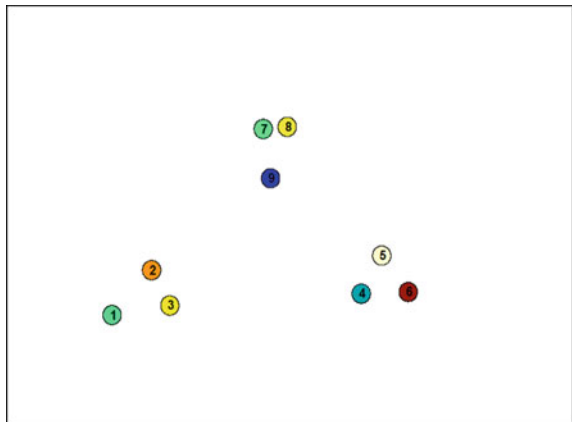
Fig. 4 Example scenario for grouping algorithm: **a** initial scenario illustrating applied forces, **b** generic grouping algorithm
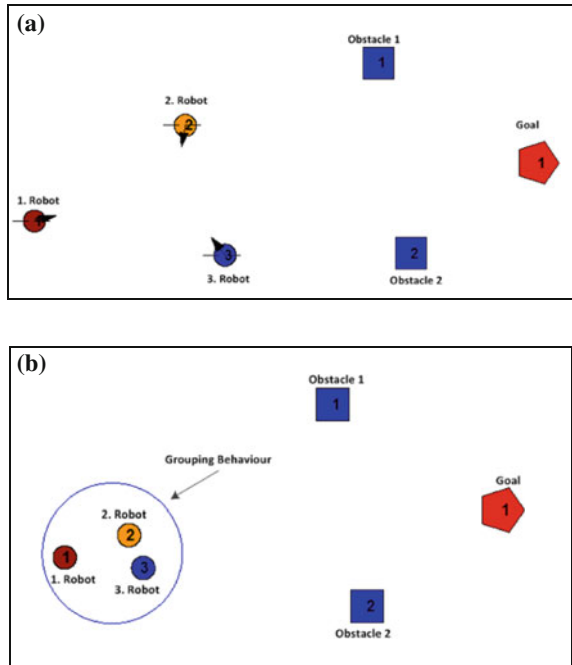
coordinate system and observes other robots' location to estimate at which quadrants they are placed in the coordinated system. Consequently, a robot is accepted as the leader of its group once other members are located in the same quadrants. An example scenario is shown in Figs. 4 and 5a, illustrating the leader assignment procedure. According to scenario shown in Fig. 4, each robot applies the given algorithm to select a leader for the group without performing any direct communication or employing a central decision mechanism. The decision mechanism of the algorithm for the given scenario works as follows:

- For the first robot (shown in Fig. 5a), the second robot is at 1st quadrant but the third one is placed at the 4th quadrant.
- For the second robot (shown in Fig. 5a), the first robot is at the 3rd quadrant but the third one is placed at the 4th quadrant.
- For the third robot (shown in Fig. 5a), the first and the second robots are shown at the 2nd quadrant.

The results of the algorithm indicate that the third robot is the only one to be able to detect other robots at the same quadrant (shown in Fig. 5a), which makes it the leader of the group according to the aforementioned algorithm; whereas the first and the second robots realise that they cannot be the leader. In addition, the algorithm also enables members (other robots in the swarm) to directly detect the location of the leader. Accordingly, once the leader assignment algorithm is executed, each

Fig. 5 Example scenario for
COG based navigation:
a leader is calculated for the
group, **b** coordinated motion
of swarm based on COG
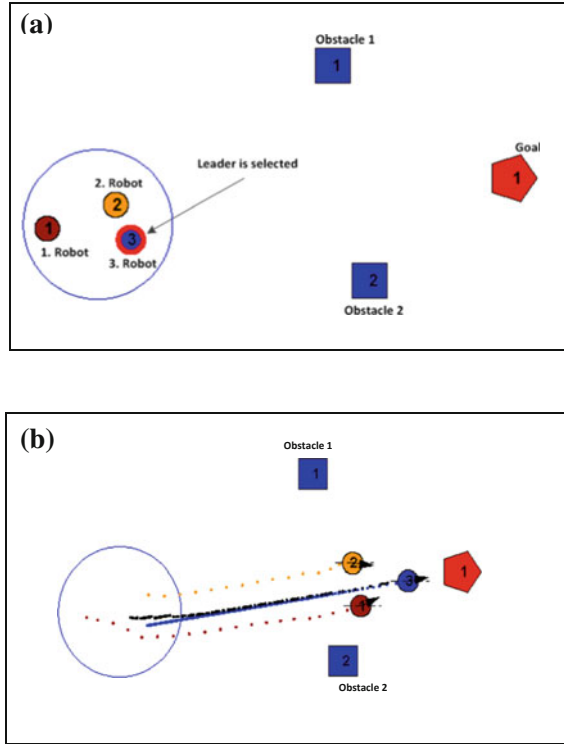
robot realises other robots' role and location in a decentralized approach. The only
exception to this algorithm occurs when all robots lie on the same line. In order to
prevent this, an assumption is made that robots cannot be located initially on the
same line.

2.3 Coordinated Motion of Swarm

A centre of gravity based navigation approach is addressed to execute the coor-
dinated motion of the robots, from an initial position to the goal position. Once the
leader assignment task is completed, the leader robot moves towards a specific goal
while avoiding obstacles based on the potential field method (Khatib 1985). The
method proposes an elegant solution to the challenging path finding problem. The
attractive potential is assigned for the goal, whereas a repulsive potential is assigned
for each of the obstacles in the environment. In general the scalar potential field
P can be defined as:

$$P = P_{att} + P_{rep} \tag{3}$$

where P_{att} and P_{rep} represent attractive and repulsive forces correspondingly. The vector field of forces F(q) is given by the gradient of U:

$$F_q = -\nabla P_{att} + \nabla P_{rep} \tag{4}$$

The details of the potential filed method used in this study can be seen in Koren and Borenstein (1991) and Tang et al. (2010). As it is expected, one of the most critical issues in this framework is to be able to move robots in a coordinated manner.

As soon as the leader is selected, it is steered towards a specific goal according to the potential field method as aforementioned. Member robots, on the other hand, tend to stay in the group while the leader moves to a goal. Consequently, a simple but efficient method was proposed. According to which, the point referring the centre of gravity of the group is calculated by member robots, allowing them to both stay in the group and navigate along the direction of the leader.

The main idea lying behind the proposed navigation strategy is to assign an artificial mass to the COG point that applies a gravitational force to member robots. The COG point is updated continuously during the navigation behaviour of the leader robot, as illustrated in Fig. 5. The COG point is updated in each iteration of leader's search procedure, resulting in instant change in applied gravitational force on member robots. Member robots are eager to move the direction of the COG point so as to keep the force applied on them at the limit value. This COG based navigation strategy allows member robots to stay in the group and prevent them colliding with obstacle. Coordination motion strategy is also designed in a decentralized manner. However it is assumed that robots' known each other's mass, which is the only limitation of the algorithm. Results of preliminary experiments revealed that assigning smaller mass to member robots than the leader provides more accurate coordinated motion behavior.

3 Evaluation of the System

In order to estimate the capability of the proposed work, the system has been evaluated using a simulator. The simulator was implemented by authors based on the Matlab toolbox. A series of simulations has been conducted to verify that the robots are able to navigate towards a goal in a coordinated and safe manner. The interface of the simulator allows users to add robots, obstacles and goals deliberately or randomly into the working environment. Algorithms of the proposed system can be visualized as a whole or individually. Table 1 summarizes the parameters employed for experiments that include initial parameters assigned for grouping and coordinated motion behaviors respectively. For instance, 'G_l' refers

Table 1 Parameter table for experiments

Parameters for experiments		
Parameter	Value	Explanation
Parameters before grouping		
G_l	0.78	Maximum gravitational limit between robots while grouping
R_{mass}	2.0	Robots 'initial mass
O_{mass}	−0.24	Obstacle mass
G_{bg}	−1.5	Target mass before grouping
G_c	3.0	Number of robots belong to each group
Parameters after grouping		
A_{mass}	−0.85	Attraction mass between robots after grouping
G_{ag}	3.0	Target mass after grouping
C_{lmass}	3.0	Leader mass for COG calculation
M_{COG}	1.0	Member mass for COG calculation
C_{mass}	2.0	COG initial mass

Fig. 6 Grouping and leader assignment task
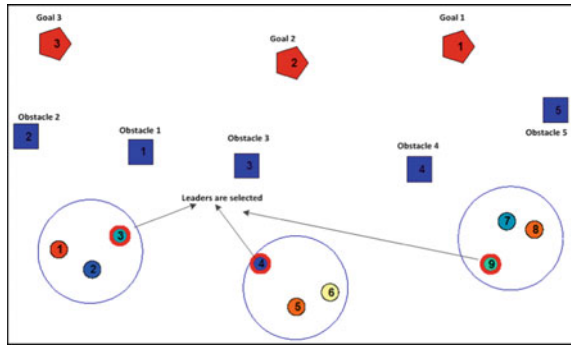
attraction force, applied between robots whereas 'G_{bg}', 'O_{mass}' and 'R'_{mass}' refer mass values employed during the grouping behavior. On the other hand, 'A_{mass}' is considered as the most critical parameter for coordinated motion behavior, which denotes repulsive force applied between robots. This is in essence responsible from preventing collision between robots in the same group and allows those robots to move in a coordinated manner. Initial mass value, assigned to calculate COG is also given in the corresponding table that the mass of the leader and member robots are adjusted for the coordinated motion behavior.

A comprehensive example is shown and detailed in this section. Figures 6, 7 and 8 illustrate this example where there exists 9 robots, 5 obstacles and 3 goals in the corresponding scenario, as shown in Fig. 7, which also illustrates the critical grouping and leader selection behaviours for this example. Once the leader is assigned to each group, COG point is calculated (see Fig. 7) based on the parameters, shown in Table 1. Afterwards, the leader starts moving towards the

Fig. 7 COG are calculated for each group
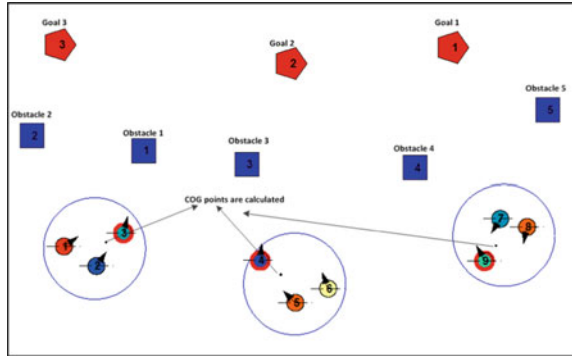
Fig. 8 Groups achieve to reach the goal while preventing collision
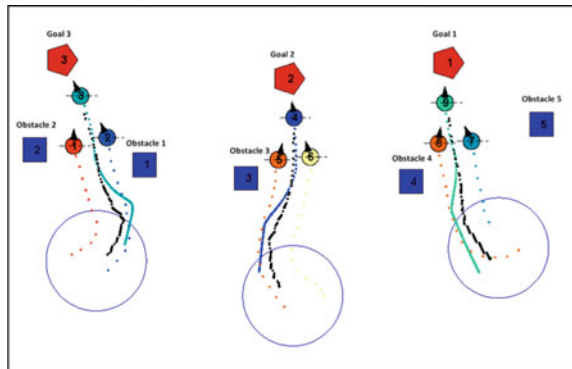
goal with the potential field method. Shortly after the leader starts navigation, member robots are enabled and starts following the COG point which is modified continuously within the maneuver of the leader as illustrated in Fig. 8. This figure also demonstrates the paths followed by each robot that all robots achieve to reach the goal while avoiding collision with obstacles.

4   Conclusions

Decentralized control of multi-agent systems is a critical engineering field and has recently gained a lot of attention from researchers. In this manner, this paper proposes a new framework for aggregation and coordinated motion of swarm robots in a decentralized manner. The framework presents a generic grouping approach that groups robots in a hierarchical manner and allocates a leader for the group. The approach does not require any direct communication between robots and is designed using virtual physic-based algorithms. The leader robot of each group, possessing a map based navigation strategy, tends to move towards a

specific goal while avoiding obstacles placed on its path. On the other hand, member robots, utilizing a COG based algorithm, move the direction of the leader while keeping themselves in the group with a safe and coordinated motion.

Simulation experiments reveal that the proposed framework achieves high degree of accuracy in complex scenarios. These results encouraged authors to conduct real word experiments with the proposed framework in order to assess the overall performance of the framework. Besides, an intelligent patter recognition module will be integrated into the proposed framework to separate obstacles from robots during the grouping and leader selection tasks. This will allow robots to ignore obstacles at the initial state and enhance the overall performance of the grouping algorithm.

References

Brambilla, M., Ferrante, E., Birattari, M., & Dorigo, M. (2013, March). Swarm robotics: A review from the swarm engineering perspective. *Swarm Intelligence*, *7*, 1–41.

Camazine, S., Deneubourg, J.-L., Franks, N. R., Sneyd, J., Theraulaz, G., & Bonabeau, E. (2001). *Selforganization in biological systems. Princeton studies in complexity*. Princeton: Princeton University Press.

Ferrante, E., Turgut, A. E., Huepe, C., Stranieri, A., inciroli, C., & Dorigo, M. (2012). Self-organized flocking with a mobile robot swarm: a novel motion control method. *Adaptive Behavior*.

Ferrante, E., Turgut, A. E., Mathews, N., Birattari, M., & Dorigo, M. (2010). Flocking in stationary and non-stationary environments: A novel communication strategy for heading alignment. In Lecture notes in computer science (pp. 331–340). Berlin: Springer.

Garnier, S., Jost, C., Jeanson, R., Gautrais, J., Asadpour, M., Caprari, G., et al. (2005). Aggregation behaviour as a source of collective decision in a group of cockroach-like robots. In *Lecture notes in artificial intelligence: Vol. 3630. Advances in artificial life* (pp. 169–178). Berlin: Springer.

Jeanson, R., Rivault, C., Deneubourg, J.-L., Blanco, S., Fournier, R., Jost, C., et al. (2005). Selforganized aggregation in cockroaches. *Animal Behaviour, 69*(1), 169–180.

Kaminka, G. A., Schechter-Glick, R., & Sadov, V. (2008). Using sensor morphology for multirobot formations. *IEEE Transactions on Robotics, 24*(2), 271–282.

Khatib, O. (1985). Real-time obstacle avoidance for manipulators and mobile robots. IEEE International Conference on Robotics and Automation (pp. 500–505), St. Louis.

Koren, Y., & Borenstein, J. (1991). Potential field methods and their inherent limitations for mobile robot navigation. In IEEE International Conference in Robotics and Automation (Vol. 2, pp. 1398–1404), April 9–11, 1991.

Soysal, O., Bahçeci, E., & Şahin, E. (2007). Aggregation in swarm robotic systems: Evolution and probabilistic control. *Turkish Journal of Electrical Engineering and Computer Sciences*, *15*(2), 199–225.

Soysal, O., & Şahin, E. (2005). Probabilistic aggregation strategies in swarm robotic systems. In *Proceedings of the IEEE swarm intelligence symposium* (pp. 325–332). Piscataway: IEEE Press.

Soysal, O., & Şahin, E. (2007). A macroscopic model for self-organized aggregation in swarm robotic systems. In *Lecture notes in computer science: Vol. 4433. Swarm robotics* (pp. 27–42). Berlin: Springer.

Tang, L., Dian, S., Gu, G., Zhou, K., Wang, S., & Feng, X. (2010). A novel potential field method for obstacle avoidance and path planning of mobile robot. In 2010 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT) (pp. 633–637).

Turgut, A. E., Çelikkanat, H., Gökçe, F., & Şahin, E. (2008). Self-organized flocking in mobile robot swarms. *Swarm Intelligence, 2*(2–4), 97–120.

Kinematic Performance Analysis of a Hybrid-Driven Waist Rehabilitation Robot

Bin Zi, Guangcai Yin, Yuan Li and Dan Zhang

Abstract A hybrid-driven waist rehabilitation robot (HWRR) is designed for human waist training. According to the analysis of human body structure, the structure of the rehabilitation robot that employs the double parallel mechanism is presented. The cable-driven parallel mechanism is chosen to drive the human lower limb, and pneumatic artificial muscle (PAM) is chosen to twist the human waist. Based on the synthesis of the two types of actuators, the adaptability and safety of the HWRR are enhanced. The inverse kinematics and statics are analyzed, and the Jacobian matrix is established to analyze the rehabilitation robotic features. In addition, on the basis of the Jacobian matrix and the statics equilibrium equation, the kinematic performances of the HWRR in terms of workspace, dexterity and stiffness are calculated and analyzed.

Keywords Waist rehabilitation robot · Hybrid-driven · Kinematic performance analysis

B. Zi (✉) · G. Yin · Y. Li
School of Mechanical and Automotive Engineering, Hefei University of Technology,
193 Tunxi Road, Hefei, China
e-mail: zibinhfut@163.com

B. Zi
The Key Laboratory of Ministry of Education for Electronic Equipment Structure Design,
Xidian University, South Taibai Road, Xi'an, China

D. Zhang
Faculty of Engineering and Applied Science, University of Ontario Institute of Technology,
Oshawa, ON L1H 7K4, Canada

1 Introduction

Rehabilitation robots were studied and developed for restoration of limb function which is caused by human disease (such as stroke), old age, and nature disaster, etc. In order to reduce the discomfort, the flexibility and safety of rehabilitation robots were pursued by using soft actuators. The soft actuator was applied to drive directly to obtain the whole mechanical stiffness. A cable-driven arm rehabilitation robot is designed by Zanotto et al. (2014). Jamwal et al. (2014) presented an adaptive wearable parallel robot actuated by pneumatic artificial muscle (PAM) for the treatment of ankle injuries.

According to Zi et al. (2008) and Trevisani (2010), the cables driven parallel robots (CDPRs) features high dynamics due to small moving mass, large workspace, and low cost. But for more degrees of freedom (DOFs) of the end-effector, there are some major drawbacks: given the higher number of cables and actuators, the CDPRs are often more difficult to design (e.g. it may be difficult to avoid cable-interference) and more expensive to build and maintain. On the other hand, the arrangement of cables should be more difficult in view of the recover patient's safety in this work according to Mao et al. (2015).

Based on the works of Zhao and Zi (2013) and Dwivedy and Eberhard (2006), it is obvious that there are some advantages for PAM such as high force-to-weight ratio, variable installation possibilities, no mechanical parts, lower compressed air consumption, and low cost, but the change of PAM length is quite limited and it only can be operated in the contractile direction.

For achieving the goal which is developed a waist rehabilitation robot suitable for task-oriented rehabilitation therapy, and on the papers of Yoon et al. (2010) and Kong et al. (2010) studied, the main factors should be considered in the process of mechanical structure design are the human body structure, the waist movement characteristics and the actuators' drive characteristics.

For these reasons aforementioned, in this paper a new mechanism for waist rehabilitation actuated by PAM and cable is presented. The PAM system which is make up of PAM and supporting spring drive and support the upper parallel mechanism, and the lower platform was driven by four cables in order to obtain quite large workspace and the flexibility of the robot. These two parallel mechanisms driven human body respective and synergistic which connected by the recover patient's body. The double parallel mechanism also provides unique features which other mechanism is not easy obtain wholly. These features include: (1) the cable parallel mechanism makes the HWRR's workspace larger and the PAM systems make the force exerted on the waist bigger; (2) the role of the human body in rehabilitation robot was considered as two serial links, and the safety of HWRR was obtained by using soft actuator; (3) considering the difference of human character, the hybrid mechanism was designed for the HWRR's adaptability; (4) the full symmetry structure makes the analysis of the HWRR easier.

This organization of the paper is as follows: Sect. 2 introduces the novel concept of the HWRR with hybrid-driven double parallel mechanism. Section 3 briefly

describes the unified planar parametric kinematic equations and static analysis. Section 4 shows the kinematic performance of the HWRR and some analysis results were displayed.

2 Structure Design

For the purpose of designing a cost-effective, safe, flexible, and well-adapted waist rehabilitation robot, the human waist structure and its complex motions are studied first (see the Fig. 1). According to Liang and Ceccarelli (2012), the waist has 3 DOFs relative to the pelvis, three direction of rotation, and the rotation of the Z axis direction can be neglected because of it is small. So the motion between the waist and pelvis can be considered as 2 DOFs, i.e. the connection between the waist and the pelvis can be simplified to a universal joint. So do the lower limb.

The HWRR architecture introduced in this paper is shown in Fig. 2. As mentioned before, the type of actuator, cable and PAM are chosen, then the human body structure and the waist movement character are considered in the structure design to make the HWRR safety, well-adapted and flexible.

The belt is actuated by four surrounded PAM systems with the movement of the vertical direction. The standing platform is driven by four cables, and each connecting points are distributed uniformly. The other end of each cable is twined on

Fig. 1 The human body structure and movement analysis
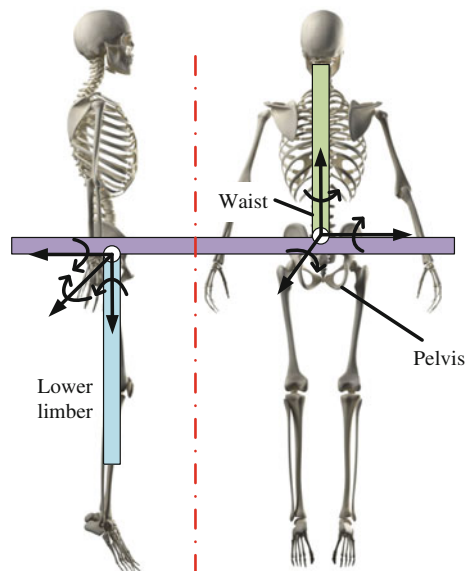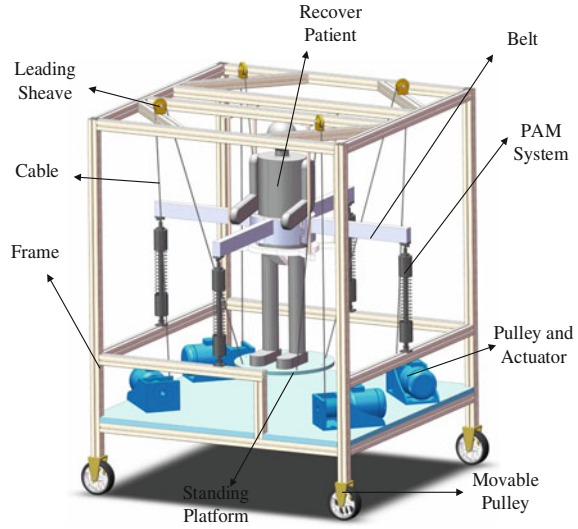
the pulley by round the leading sheave on the upper frame. The pulley is fixed on the ground. Obviously, the recover patient's particular height is adapted through the standing platform's rising and dropping. In this work, the waist and the lower limb are treated as two serial links which connect the two parallel structures. The HWRR's rehabilitation strategy is executed by changing the posture of the two serial links.

3   Inverse Kinematics and Statics

3.1   Inverse Kinematics

The planar inverse kinematics is studied because of the structure is strictly symmetrical. The inverse kinematics problem is stated as: given the desired posture $\{y, z, \alpha, \beta\}$, where α is the angle between the waist and the vertical, and β is the angle between the legs and the vertical, we need to calculate the cable lengths L_i and the PAM lengths H_i. In this paper, the cable length is the actual cable's projection in the plane showed in Fig. 3.

In order to describe the kinematic model of the HWRR, the coordinate system is shown in Fig. 3. A coordinate frame OYZ is attached to the ground, with the origin at the bottom-left corner of the frame. The y-axis is along OB_4, and the z-axis is perpendicular to the ground. Given the posture of pelvis (the central point is O_p), the coordinates of the cross point of the waist and the belt O_c and the cross point O_B of the legs and the standing platform can be calculated:
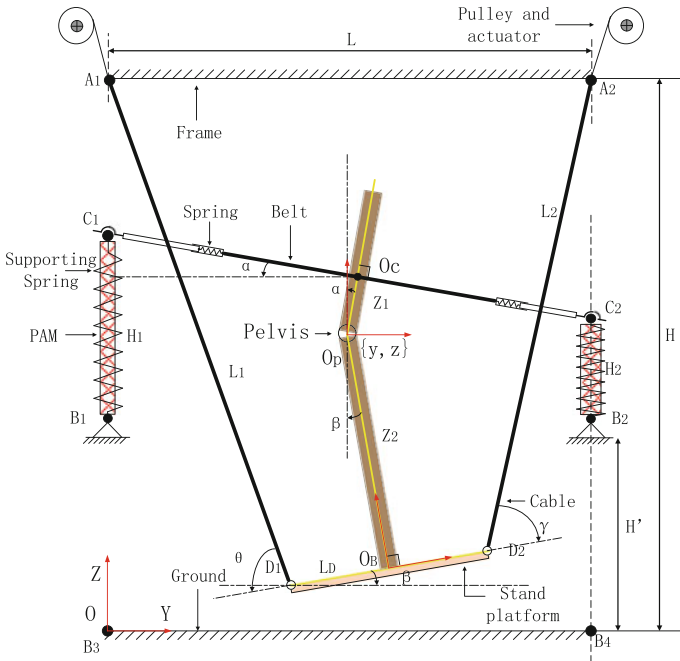
Fig. 3 Coordinates and planar kinematic analysis of the HWRR

$$O_c : (y + Z_1 \sin \alpha, z + Z_1 \cos \alpha)$$
$$O_B : (y + Z_2 \sin \beta, z - Z_2 \cos \beta)$$

The lengths of $O_c O_p$ and $O_p O_B$ are described as Z_1 and Z_2. Assuming the side length and the height of the frame and the height of the bottom of the PAM are L, H, and H'. The coordinates of points A_1, A_2, B_1 and B_2 are $(0, H)$, (L, H), $(0, H')$ and (L, H'). Based on geometrical relationship, the coordinates of points C_i and $D_i (i = 1, 2)$ can be calculated as follows:

$$C_1 : (0, z + Z_1 \cos \alpha + (y + Z_1 \sin \alpha) \tan \alpha),$$
$$C_2 : (L, z + Z_1 \cos \alpha - (L - y - Z_1 \sin \alpha) \tan \alpha),$$
$$D_1 : \left(y + Z_2 \sin \beta - \frac{L_D}{2} \cos \beta, z - Z_2 \cos \beta - \frac{L_D}{2} \sin \beta \right),$$
$$D_2 : \left(y + Z_2 \sin \beta + \frac{L_D}{2} \cos \beta, z - Z_2 \cos \beta + \frac{L_D}{2} \sin \beta \right),$$

where L_D is the side length of the standing platform.

The lengths of each cable and PAM can be obtained according to the coordinates of each point:

$$L_1 = \sqrt{\left(y + Z_2 \sin\beta - \frac{L_D}{2}\cos\beta\right)^2 + \left(z - Z_2 \cos\beta - \frac{L_D}{2}\sin\beta - H\right)^2}$$

$$L_2 = \sqrt{\left(y + Z_2 \sin\beta + \frac{L_D}{2}\cos\beta - L\right)^2 + \left(z - Z_2 \cos\beta + \frac{L_D}{2}\sin\beta - H\right)^2}. \quad (1)$$

$$H_1 = \sqrt{(z + Z_1 \cos\alpha + (y + Z_1 \sin\alpha)\tan\alpha - H')^2}$$

$$H_2 = \sqrt{(z + Z_1 \cos\alpha - (L - y - Z_1 \sin\alpha)\tan\alpha - H')^2}$$

The inverse kinematics has been used to obtain the Jacobian matrix. Performing the differential in both sides of aforementioned equations, the linear relationship is obtained as

$$\begin{pmatrix} \dot{L}_1 \\ \dot{L}_2 \\ \dot{H}_1 \\ \dot{H}_2 \end{pmatrix} = A \times \begin{pmatrix} \dot{y} \\ \dot{z} \\ \dot{\alpha} \\ \dot{\beta} \end{pmatrix}, \quad A = \begin{pmatrix} \frac{\partial L_1}{\partial y} & \frac{\partial L_1}{\partial z} & \frac{\partial L_1}{\partial \alpha} & \frac{\partial L_1}{\partial \beta} \\ \frac{\partial L_2}{\partial y} & \frac{\partial L_2}{\partial z} & \frac{\partial L_2}{\partial \alpha} & \frac{\partial L_2}{\partial \beta} \\ \frac{\partial H_1}{\partial y} & \frac{\partial H_1}{\partial z} & \frac{\partial H_1}{\partial \alpha} & \frac{\partial H_1}{\partial \beta} \\ \frac{\partial H_2}{\partial y} & \frac{\partial H_2}{\partial z} & \frac{\partial H_2}{\partial \alpha} & \frac{\partial H_2}{\partial \beta} \end{pmatrix}. \quad (2)$$

Therefore, the Jacobian matrix J_{jac}, which is the inverse of the matrix A, is expressed as

$$\begin{pmatrix} \dot{y} \\ \dot{z} \\ \dot{\alpha} \\ \dot{\beta} \end{pmatrix} = J_{jac} \times \begin{pmatrix} \dot{L}_1 \\ \dot{L}_2 \\ \dot{H}_1 \\ \dot{H}_2 \end{pmatrix}, \quad J_{jac} = A^{-1}. \quad (3)$$

3.2   Static Analysis

As the recover patient plays a very important role in the HWRR, this section focuses on the static analysis. As mentioned in Sect. 2, the waist and the lower limb are simplified to the two serial links, so forces acting on the point of junction and forces along the direction of cables and PAMs can be calculated. These are obtained by means of the Newton's method, as follows:

$$W = M^{-1}B. \quad (4)$$

The matrix M is

$$M = \begin{bmatrix} \sin\theta & \sin\gamma & -\sin\beta & -\cos\beta & 0 & 0 \\ \cos\theta & -\cos\gamma & \cos\beta & \sin\beta & 0 & 0 \\ \frac{Z_2}{2}\cos\theta + \frac{L_D}{2}\sin\theta & -\frac{Z_2}{2}\cos\gamma - \frac{L_D}{2}\sin\gamma & \frac{Z_2}{2}\cos\beta & \frac{Z_2}{2}\sin\beta & 0 & 0 \\ 0 & 0 & \cos\alpha & \sin\alpha & 0 & 0 \\ 0 & 0 & \sin\alpha & \cos\alpha & -\sin\alpha & -\sin\alpha \\ 0 & 0 & \cos\alpha & \sin\alpha & dx_1 & dx_2 \end{bmatrix}, \quad (5)$$

where dx_1 and dx_2 are

$$\begin{aligned} dx_1 &= y + Z_1\sin\alpha \\ dx_2 &= L - (y + Z_1\sin\alpha) \end{aligned}, \quad (6)$$

and W and B are

$$W = \begin{pmatrix} T_1 \\ T_2 \\ Y \\ Z \\ F_1 \\ F_2 \end{pmatrix}, \quad B = \begin{pmatrix} G_2\cos\beta \\ -G_2\sin\beta \\ 0 \\ G_1\sin\alpha - F_{k1} - F_{k2} \\ G_1\cos\alpha \\ 0 \end{pmatrix}. \quad (7)$$

Note that θ and γ are the angles between each cables and the standing platform, respectively. Y and Z are the horizontal force and the vertical forces acting on node O_p, respectively. $T_i(i = 1, 2)$ is the force acting on the standing platform by cable. F_i is the vertical force acting on the belt by PAM. Assuming that the center of gravity of the legs is coincided with geometric center point and the center of gravity of the trunk is coincided with the point of intersection with belt. G_1 and G_2 are the gravity of the lower limb and the trunk, respectively. The gravity of cables, standing platform and belt are ignored. F_{ki} is the force generated by the deformation of the springs which are set up in the each links along the belt.

4 Kinematic Performance Analysis

4.1 Workspace Analysis

Having defined the inverse kinematics and static analysis of the HWRR, we can easily calculate the forces acting on the nodes for different sets of design parameters. And due to the cables can only pull the end-effector but not push on it, the HWRR's workspace can be simulated using Matlab. The parameters of the HWRR are specified in Table 1.

Table 1 The HWRR's structure parameters

Definition	Value (m)
The length of the lower limb Z_1	1.000
The length between O_p and O_c Z_2	0.200
The height of the upper frame H	2.500
The side length L	2.000
The height of the base point of PAM H'	0.500
The side length of the standing platform L_D	0.400

For preventing the standing platform from crashing the ground when working, the value of y is limited in the simulation. Notably the vertical coordinate of the center point of the pelvis meets the following function:

$$S = Z_1 \cos \beta + \frac{L_D}{2} \sin \beta. \tag{8}$$

According to Zhao and Zi (2013), the ranges of α and β are specified, the maximum value of the above formula is 1.02 m.

The workspace of the HWRR is defined as the subset of space where the tensions on cables are both non-negative based on Seriani et al. (2015). As mentioned on Sect. 3.2, it is possible to define the following relation for the workspace despite the working range limitations of each PAMs:

$$W = \{E \in P | T_1 > 0 \wedge T_2 > 0\}. \tag{9}$$

The workspace is calculated when the α and β are equal and range from $0°$ to $20°$. The value of y is limited in [0.4, 1.6], the value of z is limited in [1.02, 2]. Three certain postures are chosen and the results are shown in the Figs. 4, 5 and 6, respectively. Figure 4 shows the workspace when $\alpha = \beta = 5°$. Figure 5 shows the workspace when $\alpha = \beta = 10°$. Figure 6 shows the workspace when $\alpha = \beta = 20°$.

Figures 4, 5 and 6 show that the shape of each workspace presents right trapezoid when the angles are less than $10°$, and the area increased with the angles. The area of workspace is large enough to meet the change of different recover patient's lower limb length. And most region of the workspace locate at the left of $y = 1$ m when the angles are positive. The workspace is symmetrical when the absolute value of different angles are equal due to the symmetry of the structure.

4.2 Dexterity Analysis

Dexterity is one of the basic measure of kinematic performance. It depicts the ability to arbitrarily change its posture, or apply force and torques in arbitrary directions during the movement. In this paper, the condition number of Jacobian

Fig. 4 The workspace when
$\alpha = \beta = 5°$
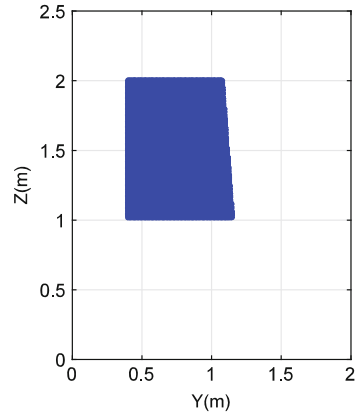
Fig. 5 The workspace when
$\alpha = \beta = 10°$
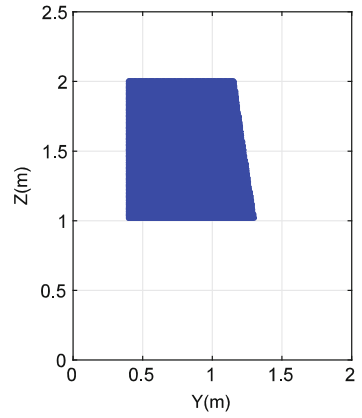
Fig. 6. The workspace when
$\alpha = \beta = 20°$
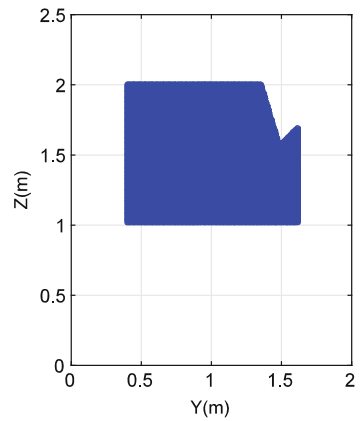
matrix is used to describe the HWRR's dexterity because of the Jacobian matrix describes the internal relationship of both velocities and forces between the end effector and the actuators.

The condition number of Jacobian matrix is

$$\kappa = \left\| J_{jac}^{-1} \right\| \cdot \left\| J_{jac} \right\|. \tag{10}$$

By studying on the works of Wang et al. (2014), the mechanism's dexterity is best when the value of the condition number is 1, and at this time mechanism is referred to as isotropic; on the contrary, mechanism tend to be singular position when the value is tend to be infinity. The condition numbers are calculated when the angles $\alpha = \beta$, and the results are shown in Figs. 7, 8 and 9, from the left to the right $\alpha = \beta = 5°$, $\alpha = \beta = 10°$ and $\alpha = \beta = 20°$, respectively.

Comparing the values of z-axis of Figs. 7, 8 and 9, it is shown that the condition number of Jacobian matrix decreases with the increasing angles. Because of the values are oversize, the mechanism optimization should be contained in the future works.

Fig. 7 The dexterity when $\alpha = \beta = 5°$
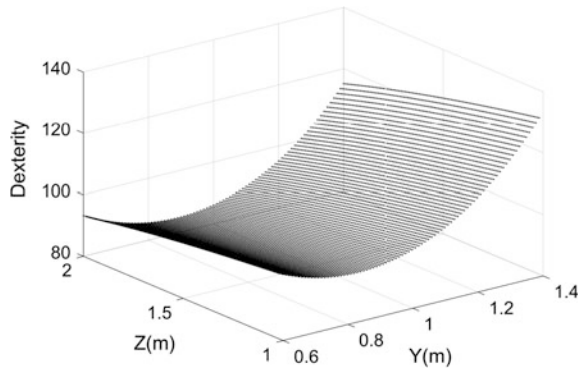
Fig. 8 The dexterity when $\alpha = \beta = 10°$
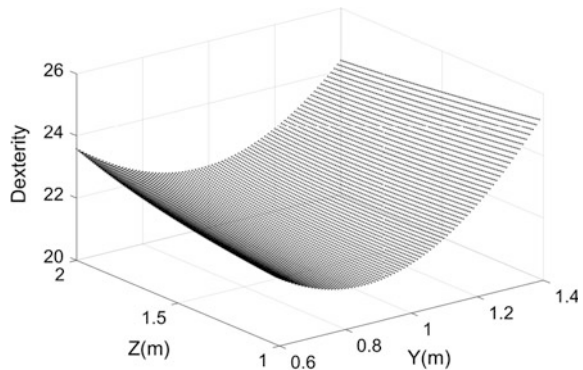
Fig. 9 The dexterity when $\alpha = \beta = 20°$
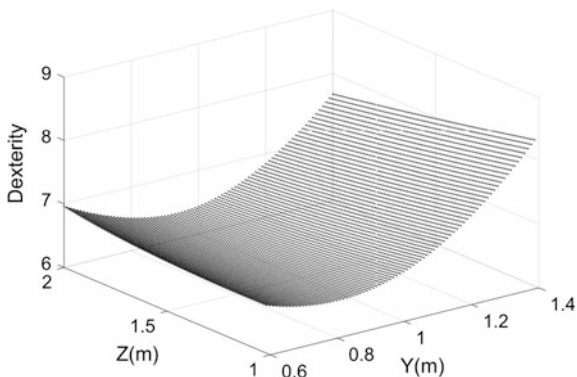
4.3 Stiffness Analysis

During practical operating, the external forces acted on the pelvis will induce a slight deviation, when the deviation is usually related with the stiffness of the proposed manipulator. For the sake of safety and accuracy, the stiffness is a critical performance of the HWRR.

Cables are treated as massless. The cable axial stiffness for unit length is

$$k_{io} = EA. \tag{11}$$

where E and A are the elasticity modulus and cross sectional area of cable, respectively.

According to Sarosi et al. (2015) and Chou and Hannaford (1996), the dragging force F in the PAM is a function of H_0 and P

$$F(P, H_0) = r_0^2 \cdot \pi \cdot P \cdot \left(\frac{3H_0^2}{H^2 tg^2 \phi} - \frac{1}{\sin^2 \phi} \right), \tag{12}$$

where r_0 is the inner radius of the tube before contraction, H_0 and H are the nominal length and the length after contraction, ϕ is the angle between the folded thread and the axle before contraction of PAM. P is the air pressure in the muscle.

Based on Eq. (12), the stiffness of PAM for the constant pressure is calculated as

$$k = \frac{\partial F(P, H_0)}{\partial H_0} = \frac{6\pi r_0^2}{H^2 tg^2 \phi} \cdot H_0 \cdot P. \tag{13}$$

The PAM can be seen as a variable stiffness spring according to Eq. (13).

According to Gosselin (1990), the stiffness matrix on the center point of pelvis can be calculated by

$$K = J_{jac}^T \Omega J_{jac},\qquad(14)$$

where $\Omega = \begin{bmatrix} k_i & 0 \\ 0 & k_j + k_s \end{bmatrix} (i,j = 1,2)$, and $k_i = \frac{k_{io}}{L_i}, k_j = \frac{k}{H_0}, k_s$ is the supporting spring's stiffness. The stiffness of belt and the spring in belt are neglected.

Based on the previous analysis, the stiffness are calculated when $\alpha = \beta$ by using Matlab, and the results are showed in the Figs. 10, 11 and 12, respectively.

Acting as rehabilitation device, the stiffness of the HWRR cannot be too high, which may cause injury to the recover patient wearing it. In this section, Figs. 10, 11 and 12 show that the stiffness is really low when $\alpha = \beta$, and the stiffness of some special points is equal to 0. Comparing Figs. 10, 11 and 12, it is obvious that the stiffness increased with the angles.

Fig. 10 The stiffness when $\alpha = \beta = 5°$
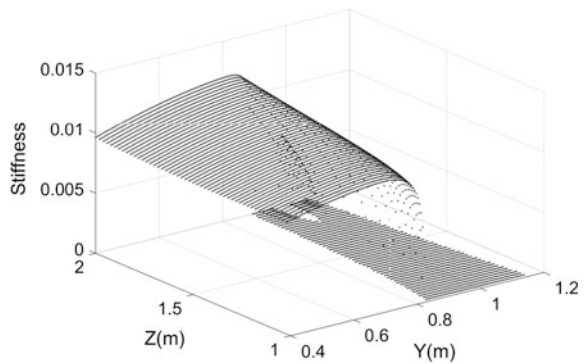
Fig. 11 The stiffness when $\alpha = \beta = 10°$
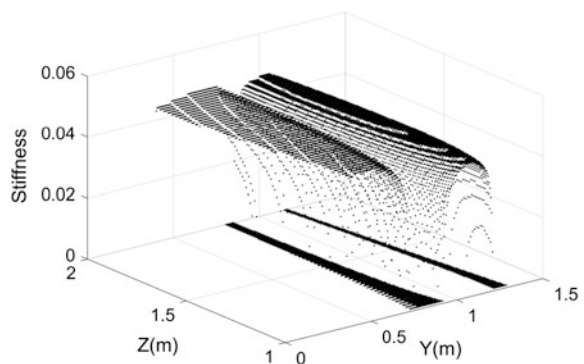
Fig. 12 The stiffness when
$\alpha = \beta = 20°$
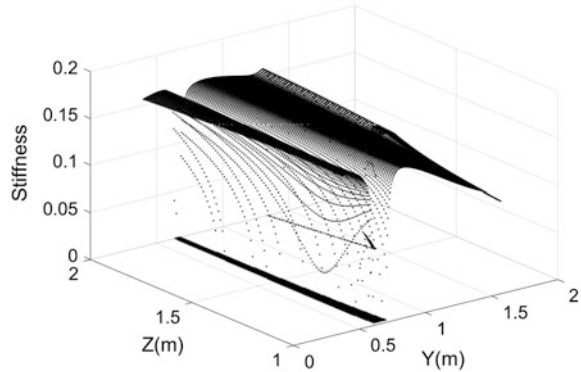
5   Conclusion

In this paper, the HWRR actuated by cable and PAM has been presented for waist rehabilitation therapy. The planar mechanism static model was established to analyze mechanical performance in terms of workspace, dexterity, and stiffness. It is shown that the CDPRs employed in the HWRR enlarged the workspace, which enhance the adaptability of the HWRR. The results present that the dexterity and stiffness of the HWRR can meet the requirement and be further improved. Compared with the expected features, it's possible to build up a cost-effective, safe, flexible, and task-oriented waist rehabilitation robot.

As a topic of the future work, the parameters optimization of the HWRR is the first thing should be done. In the near future, four research issues need to be expanded: (1) the comfort and safety pursing by parameters optimization; (2) the trajectory plan based on the dynamics; (3) coordinated control method of cables and PAMs; (4) the prototype establish for experimental verification.

References

Chou, C. P., & Hannaford, B. (1996). Measurement and modeling of McKibben pneumatic artificial muscles. *IEEE Transactions on Robotics and Automation, 12*(1), 90–102.
Dwivedy, S. K., & Eberhard, P. (2006). Dynamic analysis of flexible manipulators, a literature review. *Mechanism and Machine Theory, 41*(7), 749–777.
Gosselin, C. (1990). Stiffness mapping for parallel manipulators. *IEEE Transactions on Robotics and Automation, 6*(3), 377–382.
Jamwal, P. K., Xie, S. Q., Hussain, S., et al. (2014). An adaptive wearable parallel robot for the treatment of ankle injuries. *IEEE/ASME Transactions on Mechatronics, 19*(1), 64–75.

Kong, K., Moon, H., Jeon, D., & Tomizuka, M. (2010). Control of an exoskeleton for realization of aquatic therapy effects. *IEEE/ASME Transactions on Mechatronics, 15*(2), 191–200.

Liang, C., & Ceccarelli, M. (2012). Design and simulation of a waist–trunk system for a humanoid robot. *Mechanism and Machine Theory, 53*, 50–65.

Mao, Y., Jin, X., Dutta, G. G., Scholz, J. P., & Agrawal, S. K. (2015). Human movement training with a cable driven arm exoskeleton (CAREX). *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 23*(1), 84–92.

Sarosi, J., Biro, I., Nemeth, J., & Cveticanin, L. (2015). Dynamic modeling of a pneumatic muscle actuator with two-direction motion. *Mechanism and Machine Theory, 85*, 25–34.

Seriani, S., Seriani, M., & Gallina, P. (2015). Workspace optimization for a planar cable-suspended direct-driven robot. *Robotics and Computer-Integrated Manufacturing, 34*, 1–7.

Trevisani, A. (2010). Underconstrained planar cable-direct-driven robots: A trajectory planning method ensuring positive and bounded cable tensions. *Mechatronics, 20*(1), 113–127.

Wang, D., Fan, R., & Chen, W. (2014). Performance enhancement of a three-degree-of-freedom parallel tool head via actuation redundancy. *Mechanism and Machine Theory, 71*, 142–162.

Yoon, J., Novandy, B., Yoon, C. H., & Park, K. J. (2010). A 6-DOF gait rehabilitation robot with upper and lower limb connections that allows walking velocity updates on various terrains. *IEEE/ASME Transactions on Mechatronics, 15*(2), 201–215.

Zanotto, D., Rosati, G., Minto, S., et al. (2014). Sophia-3: A semiadaptive cable-driven rehabilitation device with a tilting working plane. *IEEE Transactions on Robotics, 30*(4), 974–979.

Zhao, X., & Zi, B. (2013). Design and analysis of a pneumatic muscle driven parallel mechanism for imitating human pelvis. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science.* 0954406213489410.

Zi, B., Duan, B. Y., Du, J. L., et al. (2008). Dynamic modeling and active control of a cable-suspended parallel robot. *Mechatronics, 18*(1), 1–12.

Admittance Filter Parameter Adjustment of a Robot-Assisted Rehabilitation System (RehabRoby)

Fatih Ozkul, Duygun Erol Barkana and Engin Maşazade

Abstract The goal of this study is to develop an adaptive adjustment system for robot-assisted rehabilitation system (RehabRoby) so that subjects can control the robot in an easy manner to complete a given rehabilitation task. A system that determines the proper admittance parameters for each subject has been integrated into RehabRoby. A survey is used to verify that executing the rehabilitation tasks under proper admittance parameters reduces the stress level of subjects, improves the controllability of the robot, and improves engagement of the subjects.

Keywords Admittance filter parameter adjustment · Robot-assisted rehabilitation systems · Admittance control

1 Introduction

Robot-assisted rehabilitation systems have first been used in large scaled clinical tests in 1998, and since then several robot-aided rehabilitation systems have been developed. There are two kinds of robot-assisted rehabilitation systems for the upper extremities in terms of mechanical design, which are end-effector-based, and exoskeleton type rehabilitation robots. MIT-MANUS (Krebs et al. 2004), MIME (Lum et al. 2006), GENTLE/S (Loureiro et al. 2003), and NeReBot are end-effector-based, and ARMin (Nef et al. 2009), and Salford Rehabilitation Exoskeleton are exoskeleton type robot-assisted rehabilitation systems. Exoskeleton type robots resemble the anatomy of the human arm and each of the robot's joints can be controlled separately, which reduces the complexity of issuing controls. Recently, we have developed an exoskeleton type upper-extremity robot-assisted rehabilitation system, which is called RehabRoby (Ozkul and Erol Barkana 2013a, b).

F. Ozkul · D.E. Barkana (✉) · E. Maşazade
Department of Electrical and Electronics Engineering,
Yeditepe University, 26 Agustos Yerlesimi, Istanbul, Turkey
e-mail: duygunerol@yeditepe.edu.tr

The robot-assisted rehabilitation system interacts with the human thus it is necessary to consider the robot, and the human as a coupled mechanical system. Application of a position control may not be enough to ensure safe dynamic interaction between the robot and the human, and use of a force feedback for the control loops in the robot-assisted rehabilitation system can be a solution (Carmichael and Liu 2013). Force based control methods are suitable where compliance with the environment is necessary for safe behavior (Augugliaro and D'Andrea 2013). Direct force control is difficult to implement in practice since it requires a model of the environment to apply a closed force feedback loop (Carmichael and Liu 2013). Indirect force control methods such as impedance control and admittance control are developed to control force through the motion of the robot. In the indirect force control a mass-spring-damper type dynamic system is modeled to relate robot interaction forces to its motion (Carmichael and Liu 2013). Actuator forces are controlled using the measurements of the position change in the joints of the robot in the impedance control (Carmichael and Liu 2013). Thus impedance control needs precise force actuation which means the actuators should be lightweight, and have low friction and low impedance (Carmichael and Liu 2013). On the other hand, measurements of the interaction forces between the robot and the environment are necessary in admittance control because desired motion trajectory is generated using these force measurements. Thus, a robot with high power actuator, and a stiff construction are more convenient when admittance control is used (Carmichael and Liu 2013). Admittance control provides smooth motion which does not oppose the external forces to enable the robot to comply with any forces applied by the human at a predefined mechanical behavior (Yamada et al. 2013; Dimeas and Aspragathos 2014). In this work, we choose an admittance control because it is convenient for RehabRoby in terms of electromechanical structure, and human-robot interaction requirements. The admittance controller has been integrated with a robust position controller, which consists of a linear discrete Kalman filter to compensate for the effects of the parameter variations, and nonlinearities in the inherent dynamic model of RehabRoby (Ozkul and Erol Barkana 2013a, b).

Note that three parameters; inertia, damping and stiffness coefficients are used to specify the virtual admittance characteristics of the robotic system in the admittance control. A large admittance causes a rapid response while a small admittance provides a slow reaction to the applied forces (Zhu et al. 2010). The inertia coefficient affects the response time since it determines the time constant of the virtual dynamic system. The damping coefficient is related to the perceived burden, and the stiffness coefficient adds a spring characteristic to the virtual dynamics. Generally inertia and damping coefficients are set as low as possible to start the motion of the robot smoothly (Zhu et al. 2010). However, too small values of inertia and damping coefficients may cause the system to be sensitive to the interaction forces. Additionally, small values for the admittance parameters may cause oscillations at the resonance frequencies depending on the type of interaction, and the dynamics of the robotic system (Okunev et al. 2012). It has previously shown that the most dominant parameter of admittance control in human-robot collaboration is the

damping coefficient (Dimeas and Aspragathos 2014). Additionally, nonzero values of the stiffness parameter can cause counterforces which are not desired in the control of the human-robot collaborative systems. We have used the inertia and damping coefficients in the admittance control of RehabRoby while the stiffness coefficient is taken as zero.

In this study, we ask each subject to perform three tasks with increased difficulty levels where RehabRoby is first operated with the default admittance parameters. Then, each subject performs the same three tasks where RehabRoby uses the admittance parameters specifically adopted for each subject. Later, a survey is used to discriminate whether the subjects can control or not control the robot with the default or updated admittance parameters.

The rest of the paper is organized as follows. The details of the Robot Assisted Rehabilitation System, RehabRoby are introduced in Sect. 2. The tasks that the subjects are executing with RehabRoby and the experimental set-up details are given in Sect. 3. Then in Sect. 4, the details of the admittance filter parameter adjustment method are presented. The results of the experiments are discussed in Sect. 5. Finally, Sect. 6 is devoted to our conclusions and future research directions.

2 RehabRoby

In this study, an exoskeleton type upper-extremity robot- assisted rehabilitation system, which is called RehabRoby, has been used (Fig. 1). RehabRoby has been designed in such a way that (a) it can implement passive mode therapy, active-assisted mode therapy and resistive mode therapy, (b) it can be easily adjusted for people of different heights and with different arm lengths, (c) it can be used for both right and left arm rehabilitation. RehabRoby is designed to provide extension, flexion, abduction, adduction, rotation, pronation and supination upper-extremity movements and the combinations of these movements that are necessary for the tasks and activities of daily living (ADL). An arm splint, which has humeral and forearm thermoplastic supports with Velcro straps, and a single axis free elbow joint, has been designed and attached to RehabRoby (Fig. 1). The thermoplastic arm splint designed for the RehabRoby has humeral and forearm supports with Velcro straps, and a thermoplastic inner layer that is covered by a soft material (Plastazote). Note that there are two force sensors (Kistler—9313AA1; Kistler France, Les Ulis, France), one of them is used to measure the applied forces during elbow flexion, and the other one is used for measuring the applied forces during elbow extension (Fig. 1). Ensuring the safety of the subject is a critical issue for a robot-assisted rehabilitation system. Thus, in the event of an emergency situation, the therapist can press an emergency button to stop the RehabRoby (Fig. 1), and the motor drivers of RehabRoby are disabled separately or together by pressing the driver enable/disable buttons without turning off RehabRoby. The system is
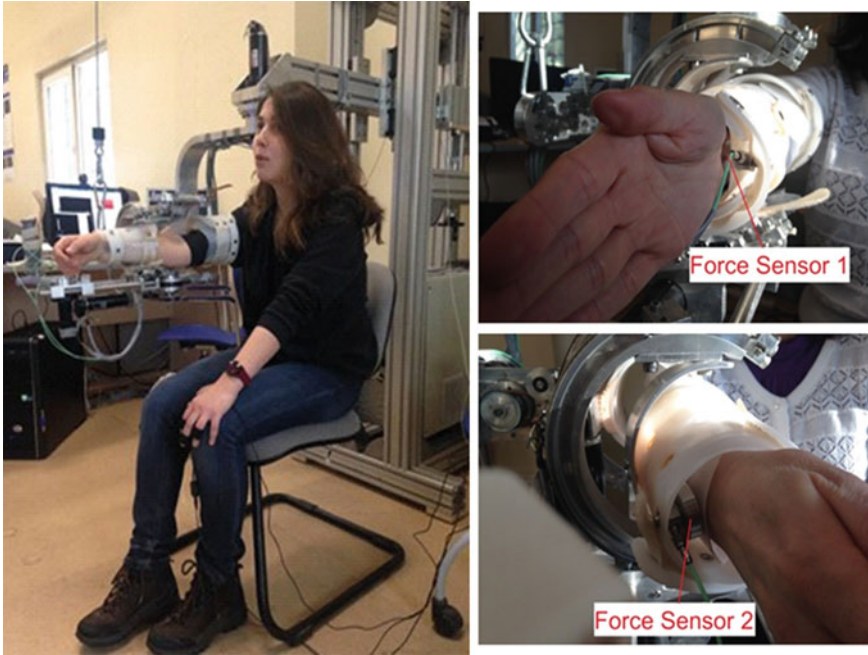
Fig. 1 *Left* RehabRoby with the subject. *Right* Locations of the force sensors

powered by an uninterruptible power supply, thus, there can be no power loss, and RehabRoby will not collapse at any time.

3 Admittance Filter Parameter Adjustment

The admittance control with an inner robust position control loop, which is used as the low-level controller of RehabRoby is shown in Fig. 2. In this study, each force axis has independent parameters (Gain; I; b) that are needed to be calculated. The forces that are applied by the subject, which are measured using Force Sensor 1 (FS1) and Force Sensor 2 (FS2), are converted into torque values (one for each force sensor) using the robot Jacobian matrix. The torque values are then passed through an admittance filter, which is used to define the characteristics of the motion of RehabRoby against the applied forces, to generate the reference motion for the position controller. The applied force/torques are the input and reference position is the output of the admittance filter. Subjects are asked to perform 90° elbow flexion and extension movement 3 times continuously while RehabRoby is kept passive to determine the proper admittance filter parameters (I and b), and the gain (Gain) for each subject.
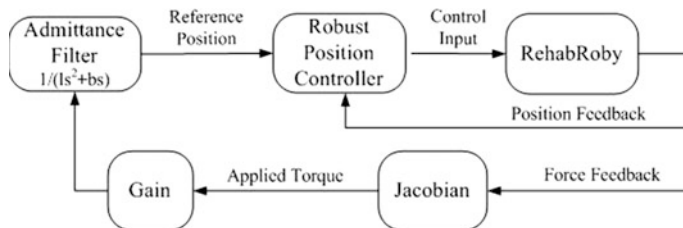
Fig. 2 Reference position generation using admittance filter in the low-level controller of RehabRoby

A suitable adaptive control scheme, which adapts the admittance parameters dynamically, has previously been obtained with a trained frequency domain classifier using adaptive boosting algorithm (Okunev et al. 2012). Furthermore, a method has been proposed to achieve an optimal variable control scheme by combining human decision making process emulated using an on-line fuzzy inference system (FIS) with a fuzzy model reference learning (FMRL) controller (Dimeas and Aspragathos 2014). FRML controller is used to adapt the FIS according to the minimum jerk trajectory model. The heuristically created FIS assists both rapid movements of the human, and the accurate positioning during lower velocities by determining the desired damping of the admittance controller. A FMRL controller adapts the FIS for optimal cooperation towards the minimum jerk trajectory model. Online Fast Fourier transform (FFT) of the manipulator end-effector forces are used to detect oscillations that disturb the haptic experience. In this study, the algorithm, which is used to find the suitable parameters in the admittance filter is executed for each of the motion regions using MATLAB/Simulink, is given below:

- Each subject has been asked to perform 90° elbow flexion and extension movement 3 times continuously while RehabRoby has been kept passive. The actual angular position $x[n]$, and applied torque $\tau[n]$ during this passive motion have been sampled with the frequency of 500 Hz using Simulink Real Time Windows Target tool. An example of one of the subject's position data is shown in Fig. 3.
- The maximum applicable torque is known as 6 Nm. The parameter *Gain* is used to scale the measured torque data during real-time experiments and found as Gain = 6/max{$\tau[n]$}, where max{$\tau[n]$} is the maximum value of the measured torque data during passive motion. A threshold is applied to the measured torque data during real-time experiments because of the noise in the measurement data. This value is chosen as 2 Nm, which is one third of the maximum applicable torque value (6 Nm).
- We are interested in the velocity characteristics of the reference motion. The slopes of $x[n]$ and $\hat{x}[n]$ are defined as $m[n] = x[n] - x[n-1]$ and $\hat{m}[n] = \hat{x}[n] - \hat{x}[n-1]$, respectively. $m[n]$ and $\hat{m}[n]$ are the backward differences between the adjacent elements of $x[n]$ and $\hat{x}[n]$.

Fig. 3 Actual position of one the subject (S1)
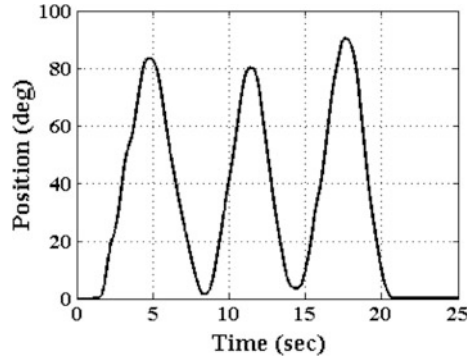
Fig. 4 Sum of the square differences at each iteration
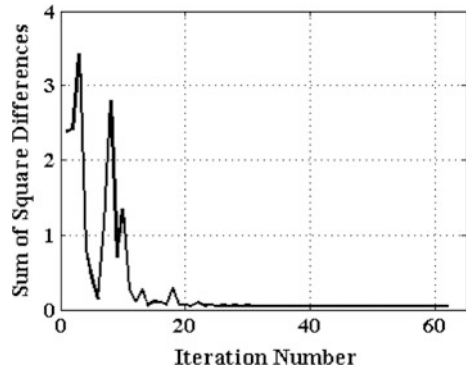
- Set $j = 0$, set b^0; I^0 to their default values, and set $T_{iterations} = 2000$
- WHILE $j \leq T_{iterations}$

  – (Step 1) $f_{sum} = \sum_{n=0}^{N-1} (m[n] - \hat{m}[n])^2$
  – (Step 2) $\{b^{j+1}, I^{j+1}\} = \arg \min_{b,I} f_{sum}$.

  – $j = j + 1$.

- END WHILE where N is the number of samples.
- Above in Step (2), optimization is performed using the *fminsearch* function of MATLAB. For the one motion region, Fig. 4 shows that the sum of the square differences f_{sum} converge to minimum while the iteration number increases. The final values of the parameters b; I are obtained by taking the mean of the parameter values calculated in each of three active motion regions shown in Fig. 3.

4 Task and Experimental Set-up

15 subjects (seven female and eight male), whose ages were in the range of 20–35, participated in this study. All subjects were right handed. All the subjects were healthy and had no background of any diseases that might have affected the study. The subjects were informed of the experiment protocol, and an orientation was given to each subject. The study was approved by the Institutional Review Board of Sabanci University.

Subjects were seated in the chair, and their arms are placed in the splint and tightly secured with Velcro straps. The height of the RehabRoby was adjusted for each subject so that they would start the task in the same arm configuration. Initially, the subjects shoulder was positioned at an extension of 90°, the elbow was in the neutral position, the lower arm is at a pronation of 90°, and the hand and the wrist were free in the neutral position. Initially, subjects were informed about the experiment and were asked to sign consent form.

A survey was given to the subject to learn how he/she feels (comfort, stress) at the beginning of the experiment. A single player pong game, which has previously been used to detect emotions such as anxiety and performance, has been selected as the task (Brown et al. 1997). The pong game contains a ball and racket. The ball moves continuously and the goal of the game is to keep the ball in the game by shifting the racket as much as possible. When the subject hits the ball, s/he gets point. When the subject misses the ball, s/he loose the point. The racket moves according to the elbow flexion and extension movements. When the subject performs elbow flexion, the racket moves through left direction. When the subject performs extension, the racket moves through right direction. At the beginning, beginning, the racket is at the middle bottom of the monitor, therefore each subject starts within 45° elbow flexion position and the game has ±45° working range as acceptable for the human anatomy. The difficulty level of the game is determined according to the speed of ball and length of racket. The ball has 3 different level speeds which are slow, medium and fast.

The racket has 3 different level length which are short, medium and long. Easy level game (Level 1—L1) considers a slow ball and long racket, medium level game (Level 2—L2) contains medium speed ball and medium length racket, difficult level game (Level 3—L3) contains fast ball and short racket. A computer monitor is placed in front of the subject to provide visual feedback about his/her motion during the execution of the task. Survey questions are asked again to subjects to learn how they feel (comfort, stress) during the task executions. Initially, the tasks L1, L2 and L3 are executed with default admittance parameters.

After executing the tasks with the default admittance parameters, subjects are asked to perform 90° elbow flexion and extension movement 3 times continuously while RehabRoby is kept passive to determine the proper admittance filter parameters (*Gain*, *I*, *b*) for that subject. Finally the tasks L1, L2 and L3 are repeated under the updated admittance filter parameters obtained individually for each subject.

Table 1 Calculated admittance parameters for each subject

Subject	I_1	b_1	Gain$_1$	I_2	b_2	Gain$_2$
S1	1.25	6.00	0.73	0.68	6.00	1.62
S2	1.35	10.67	0.89	0.84	6.00	1.21
S3	1.73	6.00	1.55	0.81	6.00	1.02
S4	1.50	10.67	1.15	0.92	6.00	0.80
S5	1.10	6.00	0.80	0.60	6.00	2.19
S6	1.07	6.31	1.19	0.78	6.00	1.32
S7	1.13	6.64	1.13	0.60	6.00	1.07
S8	1.75	6.10	0.98	1.30	6.00	0.95
S9	1.40	6.00	1.22	0.67	6.51	1.13
S10	1.44	6.00	1.16	0.65	6.00	0.78
S11	1.13	6.00	0.95	0.60	6.00	1.09
S12	1.01	6.00	0.90	0.83	6.00	0.71
S13	1.54	6.00	1.54	1.07	6.00	1.75
S14	1.26	6.00	1.37	1.85	6.00	1.50
S15	0.60	6.00	1.02	0.60	6.00	0.92
Mean	1.28	6.69	1.10	0.85	6.03	1.20

I_1, b_1, Gain$_1$ parameters for FS1; I_2, b_2, Gain$_2$ parameters for FS2

5   Results

The admittance parameter values (I, b) and the value of the *Gain* parameter that are found for both FS1 and FS2 are given in Table 1. Note that, the minimum and maximum values of I are 0.6 and 25, the initial selected value is 2, the minimum and maximum values of b are 6 and 250, the initial selected value is 20. These limit values are found by experiments and they are used to keep RehabRoby in stable region. As can be seen from Table 1 that I_1 values are above 1 with small margin except for S15 and I_2 values are below 1 again with small margin except for S8, S13 and S14. We can say that the inertia is kept small with these I values. b_1 and b_2 are at the minimum values or very close to the minimum except for S2 and S4 which means that the damping effect is minimized. The *Gain* values are above but also close to 1 which means that a small amount of torque amplification is needed.

6   Conclusions and Future Work

The goal of this study is to adjust the admittance parameters of a robot-assisted rehabilitation system (RehabRoby) in such a manner that it is possible for the subjects to control the robotic system in an easy manner to complete a given rehabilitation task. Thus, in this study adjustment of the admittance parameter is

important since dealing both with the tasks and the robot control may cause anxiety and stress to the subjects. A parameter adjustment system has been developed to determine the proper admittance parameters for each subject.

Subjects can perform flexion/extension motion without much effort when RehabRoby is passive since the mechanical design of the elbow joint of RehabRoby provides a dynamic structure with low friction and low inertia. Additionally, the flexion/extension motion has not been performed against gravity. Thus, we use an algorithm to ensure the convergence of the dynamic response of the elbow joint of RehabRoby (when admittance control mode is active) for passive dynamic characteristics. Therefore the calculated admittance parameters helps each subject to perform rehabilitation tasks in an easy manner.

Survey results indicate that by using the default admittance parameters the controllability of the robot is difficult and the subjects score poorly at each task. The controllability of the robot becomes easier, subjects score better.

As a future work, we plan to develop a real time system where we first plan to predict the emotional state of the subject using the biofeedback sensor data, then increase/decrease the difficulty level of the task which best suits to subject's emotion state. Additionally the mean of the calculated admittance parameters for each subject will be used as default admittance parameter values for the future experiments, and the difficulty level of the rehabilitation tasks can also be changed by updating these default admittance parameters.

References

Augugliaro, F., & D'Andrea, R. (2013). Admittance control for physical human-quadrocopter interaction. In *Proceedings of European Control Conference* (*ECC*), (pp. 1805–1810).

Brown, R. M., Hall, L. R., Holtzer, R., Brown, S. L., & Brown, N. L. (1997). Gender and video game performance. *Sex Roles A Journal Research, 36*(11), 793–812.

Carmichael, M. G., & Liu, D. (2013). Admittance control scheme for implementing model-based assistance-as-needed on a robot. In *Proceedings of the IEEE International Conference on Engineering in Medicine and Biology Society* (pp. 870–873), Osaka, Japan.

Dimeas, F., & Aspragathos, N. (2014) Fuzzy learning variable admittance control for human-robot cooperation. In *Proceedings of Intelligent Robots and Systems* (pp. 4770—4775).

Krebs, H. I., Ferraro, M., Buerger, S. P., Newbery, M. J., Makiyama, A., Sandmann, M., et al. (2004). Rehabilitation robotics: Pilot trial of a spatial extension for MIT-manus. *Journal of NeuroEngineering and Rehabilitation, 1*(1), 5.

Loureiro, R., Amirabdollahian, F., Topping, M., Driessen, B., & Harwin, W. (2003). Upper limb robot mediated stroke therapy GENTLE/s approach. *Autonomous Robots, 15*(1), 35–51.

Lum, P., Burgar, C. G., Van der Loos, M., Shor, P. C., Majmundar, M., & Yap, R. (2006). MIME robotic device for upper-limb neurorehabilitation in subacute stroke subjects: A follow-up study. *Journal of Rehabilitation Research and Development, 43*(5), 631.

Nef, T., Guidali, M., & Riener, R. (2009). ARMin III–arm therapy exoskeleton with an ergonomic shoulder actuation. *Applied Bionics and Biomechanics, 6*(2), 127–142.

Okunev, V., Nierhoff, T., & Hirche, S. (2012). Human-preference-based control design: Adaptive robot admittance control for physical human-robot interaction. In *2012 IEEE RO-MAN* (pp. 443–448).

Ozkul, F., & Erol Barkana, D. (2013a). Upper-extremity rehabilitation robot RehabRoby: Methodology, design, usability and validation. *International Journal of Advanced Robotic Systems, 10*(401), 1–13.

Ozkul, F., & Erol Barkana, D. (2013b). A robot-assisted rehabilitation system–RehabRoby. In M. K. Habib & J. Paulo Davim (Eds.), *Interdisciplinary mechatronics: Engineering science and research development* (pp. 145–162). New York: Wiley.

Yamada, D., Sano, Y., Hori, R., Huang, J., & Yabuta, T. (2013). Dynamic movement by admittance control of a multi-finger-arm robot with manipulability control of fingers. In *2013 IEEE/SICE International Symposium on System Integration (SII)* (pp. 521–526).

Zhu, C., Oda, M., Yoshioka, M., Nishikawa, T., Shimazu, S., & Luo, X. (2010). *Admittance control based walking support and power assistance of an omnidirectional wheelchair typed robot*. In *2010 IEEE international conference on robotics and biomimetics* (*ROBIO*) (pp. 381–386).

Continuum Robot Surfaces: Smart Saddles and Seats

Ian D. Walker

Abstract We introduce a novel form of human-robot interaction, based on an emerging new form of robot. Specifically, we analyze and discuss the ability of continuum surface robots to support the human posterior, adapting their shape to the changing needs of the human. We illustrate the concept via several examples, focusing on the application of horse riding, in which adaptive robot saddles could increase rider comfort, versatility, and safety.

Keywords Robotics · Continuum robots · Robot surfaces · Equestrian robotics

1 Introduction

Traditionally, robots have been designed to *do things apart from people*. Deployed in factory environments, industrial robots have typically operated within carefully designed workcells, isolated from humans by barriers (Hagele et al. 2008). Other traditionally successful application areas for robotics have been in environments dangerous or inaccessible to humans (Space, security, nuclear, underwater, etc.) (Trevelyan et al. 2008). In these situations, the robot is an alternative to humans, and is inherently separated from them.

Recently however, robot research has begun to focus on sharing their environment with humans. The new emphasis is on "human co-robots" (Riek 2013), or robots *doing things alongside people*. Typical examples have involved humans and robots dividing up aspects of an activity, for example assembly tasks in manufacturing applications (Hayes and Scassellati 2013). The co-robot scenario aims to exploit complementary strengths of robots (precision, repeatability) and people (perception, reasoning). However in these situations the robots and humans are typically sharing the task, but still *doing different things*.

I.D. Walker (✉)
Department of Electrical/Computer Engineering,
Clemson University, Clemson, SC 29634, USA
e-mail: iwalker@clemson.edu

In this paper, we consider a novel form of symbiotic human-robot interaction, robots *doing things intimately with people*. Specifically, we motivate and discuss the use of continuum robot surfaces which actively conform to the human body to support activities of the human. We illustrate the concept via two examples of posterior support: robot saddles and active seats.

The paper is organized as follows. In the next section, the state of the art in robotic surfaces is introduced and discussed, and the key enabling technology of continuum robots is reviewed. Example scenarios and the implications for development of the robotics are detailed in Sect. 3. Conclusions are presented in Sect. 4.

2 Robot Surfaces and Continuum Robotics

In this section, we discuss progress to date in the development of robotic surfaces. Two fundamentally different approaches—one inherently digital, the other analog—have been taken. Each is briefly reviewed in the following two subsections.

2.1 Discrete Robot Surfaces

One by now fairly extensively studied approach to creating robot surfaces (Leithinger et al. 2015) is to construct them from (a fairly large number of) discrete elements or physical pixels. Typically, an array of simple linear actuators is constructed. Some such arrays have tactile sensing allowing for humans to reshape them by contact. A fairly typical design for a discrete robot surface (Vijaykumar 2015) is illustrated in Fig. 1.

Discrete robot surfaces have been proposed principally as interactive displays offering novel data visualization. For example, Fig. 1. shows a discrete array being shaped to reflect two-dimensional data being projected on it from above. Digital robot surfaces have also recently been demonstrated performing impressive and novel forms of manipulation (Leithinger et al. 2015).

Digital displays can function as de facto surfaces by connecting the tips of their discrete elements with a deformable material. While not continuous surfaces in the formally correct sense, they can be considered as robot surfaces. However, they remain essentially digital in operation (having a series of "hard points" underlying the surface material), and require significant volume and mass.

2.2 Continuum Robotics and Continuum Robot Surfaces

The last few years has seen the emergence of a new form of robot. Termed continuum robots, these new types of "trunk and tentacle" robots are made up of
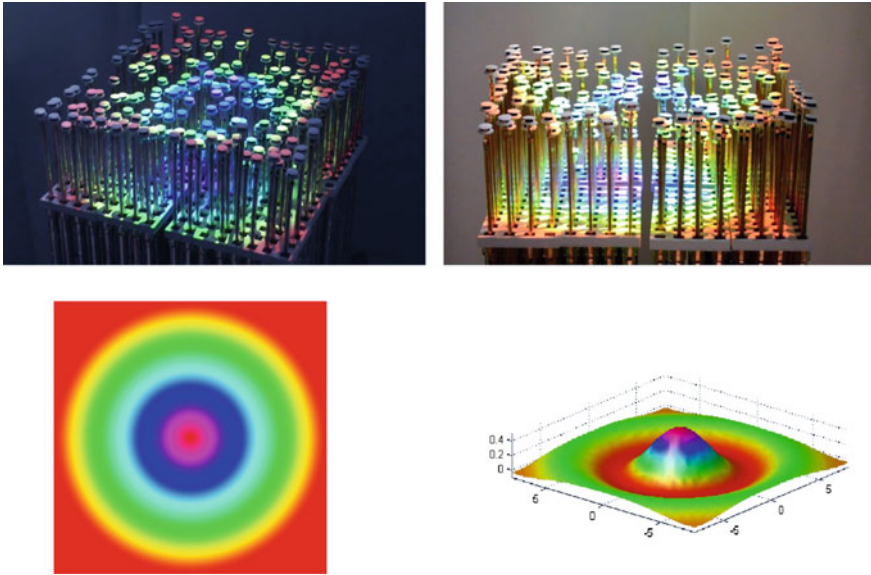
Fig. 1 Typical discrete element robotic surface design (Vijaykumar 2015)

continuous structures, effectively forming spatial curves (Trivedi et al. 2008; Walker 2013; Webster and Jones 2010). In contrast to traditional robots, which are built from a connection of rigid links, continuum robots are more compliant, adaptable, and inherently safe in contact with their environment, most notably humans.

While most research in continuum robots has concentrated on one-dimensional trunk robots, there has been work extending to two-dimensional continuum robot surfaces. Threatt et al. (2014) developed continuum robot surfaces for rehabilitation stoke patients. These pneumatically actuated two-dimensional robot surfaces (Fig. 2) support the limbs of stroke patients as they perform therapist-modulated exercises aimed at recovering lost limb movements. Inherently safe, the robot surfaces support and conform to the contours of the human, and their compliance can be modulated, guided by the therapist, to provide the appropriate resistance during the exercises to accelerate patient recovery (Threatt et al. 2014).

A key advantage of continuum robot surfaces, as compared to their discrete counterparts, is that they are more self-contained and compact, better lending themselves to deployment alongside humans. They are also inherently compliant, in contract to systems built from rigid elements which are inherently stiff. This feature makes them significantly safer to operate in intimate contact with humans. Additionally, being constructed from continuous and smooth materials, they are better suited to conforming to the human body.

Taken together, these qualities suggest the use of continuum robot surfaces for the type of intimate human robot interaction considered herein. While there are other

Fig. 2 Continuum robot stroke patient therapy surface (Threatt et al. 2014)

important applications in healthcare (for example the creation of adaptive surface dressings for treatment of burns), in this paper we focus on the use of robot surfaces for support the human posterior. We discuss these applications in more detail, along with the feasibility of creating robot surfaces for them, in the next section.

3   Application to Adaptive Seating

Use of the posterior for physical support is a universal aspect of human activity. Seating (chairs, benches, etc.) to support sitting has been produced since the dawn of human artifice. Virtually all such designs are of fixed topology, in the sense of retaining their essential shapes, following construction. However, numerous seated human activities involve, or indeed mandate, active posterior shape variation. For example, active horse riding involves significant variation of human posterior shaping depending on the activity, and thus is a particularly challenging case study. In this work, we consider the potential of continuum robot surface technology to realize novel active seating support. As a particularly interesting example, we focus on the creation and potential of robotic saddles.

3.1   Robotic Saddles

Humans and horses have enjoyed a symbiotic relationship for thousands of years (Charlish 1984). For many centuries, (equine) horsepower represented humanities main transportation and industrial engine. While machines have taken over most of these tasks in the past couple of centuries, equestrianism as a leisure and sport activity remains widespread and hugely popular.

Fig. 3 Typical saddles: *left* "English", *right* "Western"

However, the key interface in the human-equine relationship, the saddle, has undergone relatively few changes over the centuries. Saddle designs remain highly specialized, with fixed morphologies customized to different riding styles. Figure 3 shows two typical saddles: one for "English" style riding, and another for "Western" style riding. Note the sleeker form of the saddle used for English style riding, which has a strong focus on jumping. By contrast, Western style riding has a strong focus on comfort (long periods of time in the saddle) and stability (notably during fast, tight turns). Notice in particular the "horn" at the front of the Western saddle.

Figure 4 shows representative types of activities the two types of saddles are used for. The English saddle, being more streamlined, is better suited for angling the body for jumping (Fig. 4, top). However, the horn on the Western saddle is very useful for tight turns (Fig. 4, bottom), and riders often grasp it directly to increase stability (Fig. 5). However, the horn is a fixed feature, and is an impediment to jumping. For example, the presence of a horn in English-style jumping (Fig. 4, top) would either preclude the optimal jumping posture, or result in a rather painful rider/saddle contact.

A saddle which could "morph" between currently disparate types would allow its user to safely and effectively adapt to a much wider range of riding conditions. Currently, riders who know in advance that they will be encountering differing riding conditions (for example, different disciplines in a given event) must provide and switch between multiple saddles—a costly and logistically demanding exercise.

Many of the issues with fixed geometry saddles and potential benefits of adaptive saddles mentioned above also apply to bicycle and motorcycle saddles. Note however in the equestrian situation, benefits of the saddle adaptation are two way—for both the human and the horse. Fitting a saddle well and comfortably to a horse is a notoriously difficult task. An ill-fitting saddle is not only uncomfortable for the animal, but also problematic for the safety of the rider, since an uncomfortable and consequently distracted horse is less inherently controllable than a comfortable one.

Fig. 4 Core functionality of different saddle types: *Top* jumping with English saddle. *Bottom* tight cornering with western saddle

3.2 Adaptive Seating in General

Seating which can adapt to the contours of the human is of significance in applications other than horse riding. For example, a lack of relative movement between human and their substrate (in wheelchairs and beds) leads to bedsores (Takashima et al. 2014), a serious and debilitating condition in elderly and long-term largely immobile medical patients. Active seating (and bedding) pads which can subtly adjust their contours could significantly reduce this widespread, serious, and painful condition.

Fig. 5 Use of the horn on Western saddle to increase rider stability

Adaptively shaping interiors, notably seating, is also likely to become a significant component of self-driving cars as they transform the human driving experience (Hayes 2011). It is anticipated that, freed from the requirement to control their automobiles, people will demand increased comfort inside these compressed environments. Continuous active robotics control of the internal seats is likely to augment and ultimately replace current manually controlled seating adjustments.

3.3   Design and Modeling of Continuum Robot Surfaces

Continuum robot surfaces are inherently well suited for the creation and adaptation of seat and saddle shapes. They are inherently compact and self-contained, being created by embedding artificial muscles within their bodies, in carefully designed arrangements according to the anticipated functionality of the resulting surface. The core surface body element can be selected from materials varying over a wide range of stiffness, which can be biased in different directions to match the design constraints (Merino et al. 2012). The embedded internal muscles shape the surface by contracting and elongating, inducing both stretching and bending in the body core.

This approach could be implemented via a surface with one dimensional shape variability, using a single embedded muscle. In this design the single actuator is routed through the middle of a surface (the muscle would be along the line attaining the greatest height). Lines parallel to the muscle would represent "virtual muscles", and would represent intermediate effects of the coupling between the muscle length and surface body stiffness. Depending on the level of forces and complexity of overall shapes required, some of these virtual muscles could be implemented as physical muscle actuators.

Modeling of these surfaces is relatively straightforward (Merino et al. 2012). Since the overall surface shape is the direct result of the influence of the (known)

shape of the actuators embedded in it, direct interpolation of the actuator shape information can be used to accurately model the surface shape.

For example, for the muscle arrangement in the surface described above, the x and y values of the shape can be modeled by:

$$x(s) = \int\limits^{y_{min}-s} \sin\left(\int\limits^{\sigma} k(v)dv\right) d\sigma$$

$$y(s) = -\left(y_{min} - \int\limits^{y_{min}-s} \cos\left(\int\limits^{\sigma} k(v)dv\right) d\sigma\right)$$

(1)

where the lower limits of all integrals are 0, s is the arc length along the surface muscle, $x(s)$ is orthogonal to the muscle direction (diagonally left to right downwards), and $y(s)$ (vertical in figure) is the "height" of the surface. The variable k is the curvature of the muscle, and the constant y_{min} is the minimum height. The values of the intermediate values, between the muscle and the surface exterior, can be found by suitable interpolation functions for k (Merino et al. 2012).

The resulting overall (continuous) shape of the example above would be generally in the shape of a saddle. The "horn" could be actively reshaped by simple expansion and contraction of the embedded actuator. It can be made more or less pronounced by pre-biasing the stiffness and shape of the core surface material. This simple single actuator design thus can be made to smoothly adjust between the "Western" and "English" saddle shapes.

In practical robot saddle hardware, the longitudinal muscle(s) underlying the above examples can be augmented with interleaved lateral muscles, to allow the saddle to actively adjust its curvature across the back of the horse, to produce a better fit for a given horse, and the ability for the saddle to fit different horses.

We are currently experimenting with physical prototypes exploring this design space. We are exploring the use of small pneumatic (McKibben) artificial muscles embedded within the saddle surface structure, in a pattern inspired by the veins of leaves. The muscles can be supplied via a small pressurized cylinder embedded in the base of the saddle. Control of these muscles can be achieved via simple voltage control of commercial pressure regulators. The results of our ongoing research will be reported in future papers.

4   Conclusions

Robot surfaces which conform to the human body and adapt their shape according to the needs of human activities could revolutionize aspects of leisure and healthcare. The recent emergence of continuum robotics offers a new form of inherently safe and continuously deformable robot surface, offering a new form of intimate

human/robot interaction. In this paper, we have presented a new concept for robotic saddles and seats based on continuum robot surfaces. Possible future applications include active wound dressings and robotic physical therapy aids.

References

Charlish, A. (1984). *A world of horses*. Franklin Watts.

Hagele, M., Nilsson, K., & Norberto Pires, J. (2008). Industrial robotics (Chap. 42). In B. Siciliano & O. Khatib, (Eds.), *Springer handbook of robotics* (pp. 963–986). Heidelberg: Springer.

Hayes, B. (2011). Leave the driving to it. *American Scientist, 99*, 362–366.

Hayes, B., & Scassellati, B. (2013). Challenges in shared-environment human-robot collaboration. *Learning, 8*, 9.

Leithinger, D., Follmer, S., Olwal, A., & Ishii, H. (2015, September/October). Shape displays: Spatial interaction with dynamic physical form. *IEEE Computer Graphics and Applications*, 5–11.

Merino, J., Threatt, A. L., Walker, I. D., & Green, K. E. (2012). Kinematic models for continuum robotic surfaces. In *Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Vilamoura, Portugal (pp. 3453–3460).

Riek, L. D. (2013). The social co-robotics problem space: Six key challenges. In *Proceedings 1st Workshop on Robotics Challenges and Visions* (pp. 13–16), Berlin, Germany.

Takashima, A., Misaki, A., Takasugi, S. I., & Yamamoto, M. (2014). Characteristic analysis of an air cell for active air mattress of prevention for pressure ulcer. *Advanced Robotics, 28*(7), 497–504.

Threatt, A. L., Merino, J., Green, K. E., & Walker, I. D. (2014). An assistive robotic table for older and post-stroke adults: Results from participatory design and evaluation activities with clinical staff. In *Proceedings of CHI 2014: The ACM Conference on Human Factors in Computing Systems* (pp. 673–682), Toronto, Ontario, Canada.

Trevelyan, J. P., Kang, S.-C., & Hamel, W. R. (2008). Robotics in hazardous applications (Chap. 48). In B. Siciliano & O. Khatib (Eds.), *Springer handbook of robotics* (pp. 1101–1125). Heidelbeerg: Springer.

Trivedi, D., Rahn, C. D., Kier, W. M., & Walker, I. D. (2008). Soft robotics: Biological inspiration, state of the art, and future research. *Applied Bionics and Biomechanics, 5*(2), 99–117.

Viikaykumar, A. (2015, August). A scalable and low-cost interactive shape-changing display (M.S. Thesis). Department of Electrical and Computer Engineering, Clemson University.

Walker, I. D. (2013). Continuous backbone "continuum" robot manipulators: A review. *ISRN Robotics, 2013*, 1–19.

Webster, R. J, III., & Jones, B. A. (2010). Design and kinematic modeling of constant curvature continuum robots: A review. *International Journal of Robotics Research, 29*(13), 1661–1683.

Structural Parameter Identification
of a Small Robotic Underwater Vehicle

Martin Langmajer and Lukáš Bláha

Abstract This paper deals with estimating structural parameters of robotic underwater vehicle (submarine) that cannot be easily measured and determined analytically. The robotic submarine was designed for visual inspection and ultrasonic testing of submerged technologies such as tanks with slowly flowing liquid. The identification procedure contains two strategies which are able to find the structural, especially hydrodynamic parameters of a given model. These parameters are then used for stability and motion control design. The validation of gained model parameters was verified in software Simulink/SimMechanics, where the outputs of the model were compared with vehicle prototype responses.

Keywords Parameter identification · ROV/AUV · Hydrodynamic coefficients · Motion control · Model based design

1 Introduction

Research of submerged vehicles has a long history. Especially in recent years the underwater unmanned vehicles have received new attention that comes from advanced technologies and growing industry (Christ and Wernli 2007; Inzartsev 2009). Underwater unmanned vehicles (UUV) can be divided into autonomous underwater vehicles (AUV) and remotely operated vehicles (ROV). These vehicles are mostly small submersible technologies, designed for specific applications. Structural configuration and possibility of utilization are mostly very different from traditional submarines.

An accurate dynamic model is a crucial part of designing a stable, controllable and autonomous robotic system. However, the modeling of submerged system dynamics is full of nonlinearities and uncertainties. Moreover the hydrodynamic

M. Langmajer (✉) · L. Bláha
Department of Cybernetics, Faculty of Applied Sciences,
University of West Bohemia, Pilsen, Czech Republic
e-mail: mlangos3@gmail.com

107
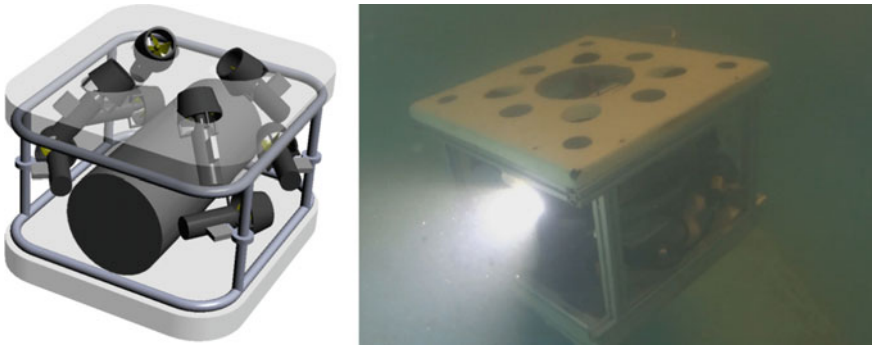
Fig. 1 Schematic representation of an underwater vehicle and corresponding real prototype

effects are often coupled between individual degrees of freedom, see Fossen (1994). Therefore it seems very natural to design a submerged robotic vehicle as symmetric as possible and use an overactuated control strategy. This ensures that matrices of parameters of model can be considered as diagonal matrices and that the system is completely controllable and robust to parametric uncertainties.

This paper deals with estimating structural parameters of a robotic underwater vehicle that cannot be easily measured and determined analytically (Ovalle et al. 2011). The vehicle is designed as a neutrally buoyant rigid body with wheeled platform containing the phased array probe for weld testing, see Fig. 1. The rest of the paper will focus on mathematical modeling and parameter identification of vehicle itself. The key part of modeling is to get a qualitatively good model, respecting all the major aspects of kinematics and dynamics, important to design of stable motion control.

In general, the identification techniques can be split into experimental techniques and into methods based on strip theory or computational fluid dynamics. This paper uses the strategy of experimental techniques.

2 Mathematical Model

The underwater vehicle is highly nonlinear complex system. Nevertheless for a streamlined vehicle, slowly floating in non-flowing liquid, we can make some standard simplifications, as mentioned by Chen (2007) or Wang (2007).

The vehicle is assumed to have almost rounded cubic shape, symmetrical about all principal axes. The velocity of the vehicle is low enough making the vortex shedding not dominant and viscous effects can be modeled as a drag in form of linear or quadratic damping. The added mass and added inertia matrices are hard to identify in whole complexity, but it could be relatively easy to make some experimental measurement to get diagonal terms of these simplified parameters. Moreover, the stabilizing control law will then be less dependent on such poorly

known parameters. All inaccuracies are modeled as external forces, which are subject of identification.

2.1 Equations of Motion

We model the vehicle as a neutrally buoyant rigid body in ideal fluid, where the motion of fluid can be described by single valued potential (Lamb 1932). Moreover the center of gravity is coincident with the center of buoyancy. Under these assumptions, we can use the description of dynamics using Kirchhoff's equations of motion, as mentioned in well known work of Fossen (1994)

$$m(\dot{v} + \omega \times v) = \sum_{j=1}^{8} T_j + F_l,$$
$$J\dot{\omega} + \omega \times J\omega = \sum_{j=1}^{8} \rho_j \times T_j + \tau_l + \tau_t, \tag{1}$$

where v, ω indicates the linear and angular velocity of coordinated C-frame, $m = m_b + m_{ad}$ is the mass of the submarine, given by inherent mass m_b and added mass m_{ad}, $J = J_b + J_{ad}$ is the tensor of inertia given by body inertia matrix J_b and added inertia matrix J_{ad} produced by fluid. T_j is the thruster force produced by jth thruster mounted on the position ρ_j, where $j = 1, 2, \ldots, 8$.

Hydrodynamic drags are modeled as an external forces acting on the vehicle, where F_l, τ_l, τ_t are external damping forces and torques caused by laminar and turbulent frictions

$$F_l = \beta_{tr} v, \quad \tau_j = \beta_{rot} \omega, \quad \tau_t = \gamma_{rot} \omega |\omega|. \tag{2}$$

The structural configuration is depicted in Fig. 2.

3 Parameter Estimation

We used several methods for estimation of unknown parameters from an identification experiment. Besides the least-square method we use nonlinear curve fitting in the least square sense (Coleman and Li 1996) or the identification of linearized system in discrete time domain using Yule-Walker equations, see Brockwell et al. (2005) for details. Here the basic features of these methods are summarized.
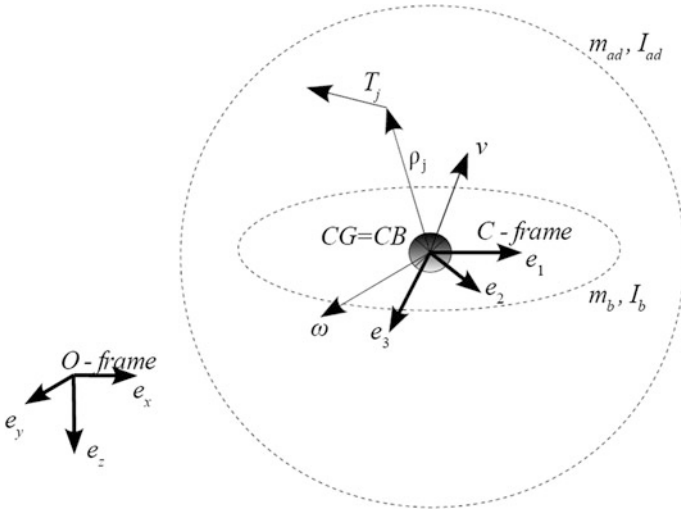
Fig. 2 Structural configuration of symmetric underwater vehicle

3.1 Least-Square Method

This is the standard approach in regression analysis. For identification purposes the least-square method is used for estimation of unknown model parameters according to some experimental measurements.

Given the overdetermined linear system of equation

$$Ax = y, \quad A \in R^{n \times m}, \quad x \in R^m, \quad y \in R^n, \tag{3}$$

where A is the matrix of regressors, which consists of measurements and often even its derivatives. The least-square method then minimizes the sum of squared residuals

$$\min_{x \in R^m} \|Ax - y\|_2. \tag{4}$$

As example, assume that the experiment is described by equation of motion in form

$$J\ddot{\theta} + \beta\dot{\theta} + K\sin(\theta) = 0, \tag{5}$$

where the parameters J, β are subject of estimation. From measurements $\theta(t_i) = \theta_i$, and derivatives $\dot{\theta}_i, \ddot{\theta}_i$ we can construct an overdetermined linear system of equation

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} -\ddot{\theta}_1 & -\dot{\theta}_1 \\ -\ddot{\theta}_2 & -\dot{\theta}_2 \\ \vdots & \vdots \\ -\ddot{\theta}_n & -\dot{\theta}_n \end{bmatrix} \begin{bmatrix} J \\ \beta \end{bmatrix}, \quad y_i = k\sin(\theta_i), \quad i = 1,\dots,n, \tag{6}$$

If A has independent columns, then the solution is unique and has the form

$$\hat{x}_{LS} = (A^T A)^{-1} A^T y. \tag{7}$$

Advantage of this method is that we can use a nonlinear form of differential equations. The drawback is that the residuals has to be linear in all unknowns. Otherwise it is necessary to use nonlinear iterative methods, see more next. Moreover the regressor matrix requires the numerical-derivative, which has to be computed very carefully, especially for noise measurements, see Sect. 3.5.

3.2 Estimation Using Curve Fitting

Standard curve fitting techniques use nonlinear iterative least-square method, as introduced in Dennis (1977). As example, assume that the identification experiment is described by second order linear equation in form

$$J\ddot{\theta} + \beta\dot{\theta} + K\theta = 0 \tag{8}$$

given by linearization from Eq. (5). This could be the case of an underdamped free decay experiment. Then the general solution in closed form can be expressed as

$$\theta(t) = A e^{-\frac{\beta}{2J}t} \quad \sin\left(\sqrt{\frac{K^2}{J^2} - \frac{\beta^2}{4J^2}}t + \theta_0\right) \tag{9}$$

The curve fitting method, for example the lqscurvefit from Matlab, then tries to estimate the unknown parameters in least-square sense. The advantage of this method is that no derivatives of measurements are needed. Drawback is that the convergence of the method very depends on initial guess of parameters, which is often very hard to set. Moreover, the general parametric solution in closed form is almost never known. So this method is rarely used for identification of more complex experiments.

3.3   Estimation in Discrete Time Domain

This estimation method works in discrete time domain. Assume again Eq. (8) describing the motion of damped linear system

$$J\ddot{\theta} + \beta\dot{\theta} + K\theta = 0 \tag{10}$$

Laplace transform gives the transfer function

$$f(s) = \frac{r_0}{(s - r_1)(s - r_2)} \tag{11}$$

which can be converted to discrete time domain

$$y(k) = -a_1 y(k - 1) - a_0 y(k - 2) + b_1 u(k - 1) + b_0(k - 2) \tag{12}$$

where the coefficients has the form

$$a_1 = -e^{r_1 T_s} - e^{r_2 T_s}, \quad a_0 = e^{(r_1 + r_2)T_s},$$
$$b_1 = Aa_1 - B\left(1 + e^{r_2 T_s}\right) - C\left(1 + e^{r_1 T_s}\right), \quad b_0 = Aa_0 + Be^{r_2 T_s} + Ce^{r_1 T_s} \tag{13}$$

and A, B, C are functions of r_0, r_1, r_2.

Then the unknown parameters can be estimated using formula

$$\hat{\theta} = -\left(P^T \phi\right)^{pinv} P^T Y, \tag{14}$$

following from Yule-Walker equations, where

$$\phi = \begin{bmatrix} 0 & 0 \\ y(1) & 0 \\ y(2) & y(1) \\ \vdots & \vdots \\ y(n-1) & y(n-2) \end{bmatrix}, \quad P = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ y(1) & 0 \\ y(2) & y(1) \\ \vdots & \vdots \\ y(n-1) & y(n-2) \end{bmatrix},$$

$$Y = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(n) \end{bmatrix}, \quad \hat{\theta} = \begin{bmatrix} \hat{a}_1 \\ \hat{a}_0 \end{bmatrix}. \tag{15}$$

The reverse conversion to continuous time domain can be done by formulas

$$\hat{r}_2 = \frac{\ln\left(\frac{-\hat{a}_1 \pm \sqrt{\hat{a}_1^2 - 4\hat{a}_0}}{2}\right)}{T_s}, \quad \hat{r}_1 = \frac{\ln(\hat{a}_0)}{T_s} - \hat{r}_2 \tag{16}$$

from them the unknown parameters can be expressed.

The main advantage of this estimation method is that it does not need numerical differentiation. Drawback is a necessity of linear equation to get the transfer function and unfortunately, the conversion of discrete parameters to continuous can be poorly conditioned, as it is in case of Eq. (16). Nevertheless, sometimes it can be useful to continue working in discrete time domain.

3.4 Base Inertial Parameters

The chosen identification procedure together with corresponding equations of motion defines and restricts the number of so-called base inertial parameters to estimate. Khalil and Dombre (2002) speak about the set of independent identifiable parameters as a set of the minimum number of parameters. The result is that the equation of motion describes not only one specific physical system, but a whole set of similar systems, which behavior is governed by equations with the same base inertial parameters. A typical example can be the estimation of inertial parameters of free decay pendulum motion, which can be described by equation

$$J\ddot{\theta} + \beta\dot{\theta} + \gamma\dot{\theta}|\dot{\theta}| + G(\theta) = 0, \quad G(\theta) = mgl\sin(\theta). \tag{17}$$

First, we are not able to identify the mass and length of pendulum, because they have multiplicative relationship. But these parameters can be apriori measured directly.

Second, we have to be careful during identification of the damping parameters β, γ, because they are both the scalers of some element of angular velocity. It means, that it very depends by which velocity the system is moving to excite each element of the equation. For example for swing free-decay pendulum with small initial angle we may not get valid estimate of β, γ because the square of velocity could be too small to excite the given element of equation of motion. The Taylor series expansion around the point $\dot{\theta}_0 = 0$ produces the approximation

$$\beta\dot{\theta} + \gamma\dot{\theta}^2 \approx \beta\dot{\theta} + \gamma\left[\dot{\theta}_0^2 + 2\dot{\theta}_0\left(\dot{\theta} - \dot{\theta}_0\right)\right] = \beta\dot{\theta} + \gamma\left[-\dot{\theta}_0^2 + 2\dot{\theta}_0\dot{\theta}\right] = \beta\dot{\theta} \tag{18}$$

from which it follows, that the γ is unobservable.

3.5 Numerical Differentiation

The identification techniques for dynamic systems usually uses the matrix of regressors, created often using a numeric differentiation of measured signal. Carefully computed numerical derivatives very much affects the results. Especially for system with noisy measurements, or for identification processes, where the input is not any simple periodic function, it is appropriate to use multi-point difference based on any robust differentiation method. Holoborodko (2008) has a very nice overview of many useful techniques, especially the method based on frequency domain criteria. This method works well even for measurement, where the ratio of signal to noise is not so big.

4 Identification Procedure

Because the lack of hydrodynamic tunnel or another flume or towing test lab, the identification experiments were reduced to a free-decay swing test and to oscillatory forced motion test, that are able to identify the structural parameters of model (1). The similar techniques has been used in Eng (2008).

Our identification procedure contains four basic steps in the following order:

- Identification of structural parameters of swing-free mechanism.
- Identification of thrust produced by propeller thruster.
- Identification of rotational parameters of vehicle.
- Identification of translational parameters of vehicle.

4.1 Identification of Mounting Mechanism and Thruster Dynamics

The vehicle has to be mounted on a swing-free mechanism to perform free-decay experiment. Therefore we need to identify the parameters of the pendulum mechanism, that have to be then subtracted to get the parameters of vehicle model alone (1). The appropriate equation of motion of mounting mechanism is in form

$$J_s\ddot{\theta} + \beta_s\dot{\theta} + m_s g l \sin(\theta) = 0 \tag{19}$$

where the actual angle of revolution θ is the measured variable. The mass and effective length were measured explicitly. The parameters J_s, β_s are subject of the identification process. From measurement $\theta(t_i) = \theta_i$ see Fig. 3 (right), and its numeric derivatives $\dot{\theta}_i, \ddot{\theta}_i$ we can construct an overdetermined linear system of equations, as in (6). Using least-square method we get the parameters of the mechanism
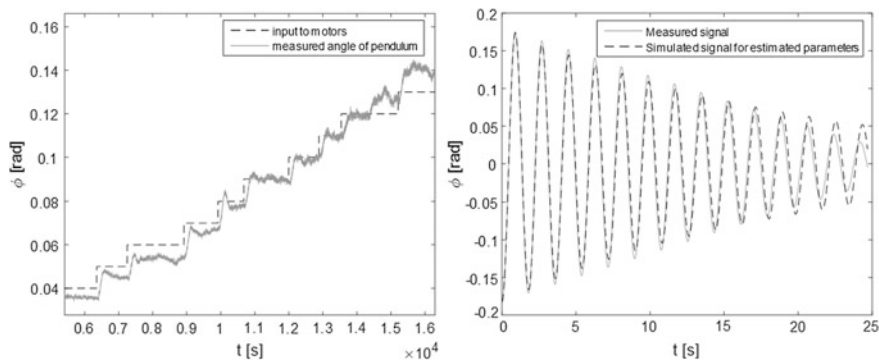
Fig. 3 Angle of inclination generated by the velocity of thruster propellers (*left*). Free-decay output response of swing mechanism and corresponding response of model with estimated parameters (*right*)

$$m_s = 7.88 \,(\text{kg}), \quad r = 0.4975 \,(\text{m}), \quad J_s = 3.13 \,(\text{kg m}^2), \quad \beta_s = 0.31 \,(\text{kg m}^2/\text{s}). \quad (20)$$

For estimating parameters of rotational motion, first we need to identify the thrust produced by propeller thruster to define the input force to the oscillatory experiment. The study of thruster dynamics is presented in Yoerger et al. (1990). The simple model, assuming the slow speed changes of propeller, can be defined as

$$T = C_t n |n|. \quad (21)$$

As before, the vehicle was mounted on a swing mechanism. The thrust was identified from a resulted inclination angle of the mechanism for a given input as a balance between torque generated by thrusters and torque produced by gravity and buoyancy.

The comparative measurement was done using a portable scale to measure the so called bollard pull, see Inzartsev (2009). Both measurements were used for estimation of parameter C_t. Figure 3 (left) shows the angle of mechanism depending on the thruster input. Note, that the behavior of thruster in forward and reverse direction is asymmetric, see Table 2.

4.2 Identification Using Forced Oscillations

For identification of rotational parameters of the model the pure rotational experiment is realized. The system is mounted in such a way, that it is able to perform only rotary motion around chosen body axis e_i. The oscillatory motion is then forced by mounted thrusters, where the thrusters torque was already identified. The equation of motion for the described experiment is

$$J\dot{\omega} + \beta_{rot}\omega + \gamma_{rot}\omega|\omega| = \tau_e, \tag{22}$$

where the meaning of variables is the same as in (1) and $\tau_e = \sum_{j=1}^{k} \rho_j \times T_j$, where k is number of thrusters, which can generate torque around a given axis.

Estimated parameters $\hat{\theta}_1 = J$, $\hat{\theta}_2 = \beta_{rot}$, $\hat{\theta}_3 = \gamma_{rot}$ define the observation matrix in form

$$A = \begin{bmatrix} \dot{\omega}_1 & \omega_1 & \omega_1|\omega_1| \\ \dot{\omega}_2 & \omega_2 & \omega_2|\omega_2| \\ \vdots & \vdots & \vdots \\ \dot{\omega}_n & \omega_n & \omega_n|\omega_n| \end{bmatrix}, \tag{23}$$

and using least-square method the estimation for e_1 axis gives the following parameter values:

$$J = 1.3\,(\text{kg m}^2), \quad \beta_{rot} = 2.2\,(\text{kg m}^2/\text{s}), \quad \gamma = 4.64\,(\text{kg m}^2/\text{rad}).$$

Estimated parameters for other axes are arranged in Table 1. Note that the J is the base inertial parameter which contains the body inertia tensor and added fluid inertia. These parameters are not observable separately by this experiment.

4.3 Free-Decay Identification Procedure

This identification procedure uses the swing-free mounting mechanism described above. The vehicle is attached to the end of the mechanism, submerged in the water and the free decay motion is performed. The structure of experiment is depicted in Fig. 4. This experiment can be described by differential equation

$$\begin{aligned} &\left((m_{ad} + m_b)l^2 + J + J_s\right)\ddot{\varphi} + (l\beta_{tr} + \beta_{rot} + \beta_s)\dot{\varphi} + \gamma_{rot}\dot{\varphi}|\dot{\varphi}| \\ &= ((m_w - m_b)l - m_s r)g\,\sin(\varphi) \end{aligned} \tag{24}$$

where the parameters agree with (1) and (19).

Table 1 Estimated parameters for translational and rotational motion

Axis	Translational motion		Rotational motion		
	m_{ad} (kg)	β_{tr} (kg m/s)	J (kg m^2)	β_{rot} (kg m^2/s)	γ_{rot} (kg m^2/rad)
e_1	70	46	1.3	2.2	4.64
e_2	80	47	2	2.2	4.75
e_3	163	66	6.7	3	5.5

Fig. 4 Free-decay structure
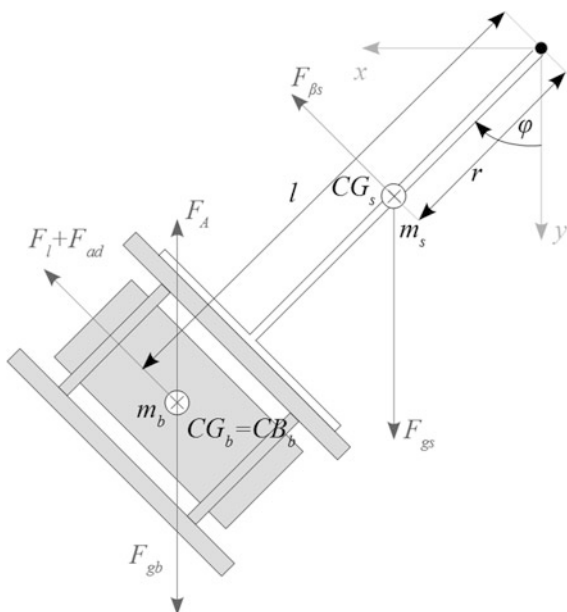
For identification purposes Eq. (24) should be due to low angular velocity (see Sect. 3.4) approximated to

$$((m_{ad} + m_b)l^2 + J + J_s)\ddot{\varphi} + (l\beta_{tr} + \beta_{rot} + \beta_s)\dot{\varphi} = ((m_w - m_b)l - m_s r)g\,\sin(\varphi)$$
(25)

Due to the structure of formula (25) there is no chance to identify more than two base inertial parameters. But the number of remaining estimated parameters is exactly two, namely m_{ad}, β_{tr}. Base Inertial parameters to estimate are

$$\hat{\theta}_1 = ((m_{ad} + m_b)l^2 + J + J_s), \quad \hat{\theta}_2 = (l\beta_{tr} + \beta_{rot} + \beta_s).$$
(26)

Observation matrix is in form

$$A = \begin{bmatrix} \ddot{\varphi}_1 & \dot{\varphi}_1 \\ \ddot{\varphi}_2 & \dot{\varphi}_2 \\ \vdots & \vdots \\ \ddot{\varphi}_n & \dot{\varphi}_n \end{bmatrix},$$
(27)

and the corresponding estimated parameters by least-square method are expressed from base inertial parameters (26). For axis e_1 we get

$$m_{ad} = 70 \,(\text{kg}), \quad \beta_{tr} = 46 \,(\text{kg m/s}). \tag{28}$$

Other estimated parameters are arranged in Table 1.

5 Results of Identification

This section summarizes the results of parameter identification. All the estimated parameters were checked using a simulating model in Matlab/SimMechanics, where the model response is compared to measured data. The comparison of measured and simulated responses is depicted in Figs. 5 and 7. The Fig. 6. depicts the record of thruster input and forced oscillation of vehicle in each axis. The
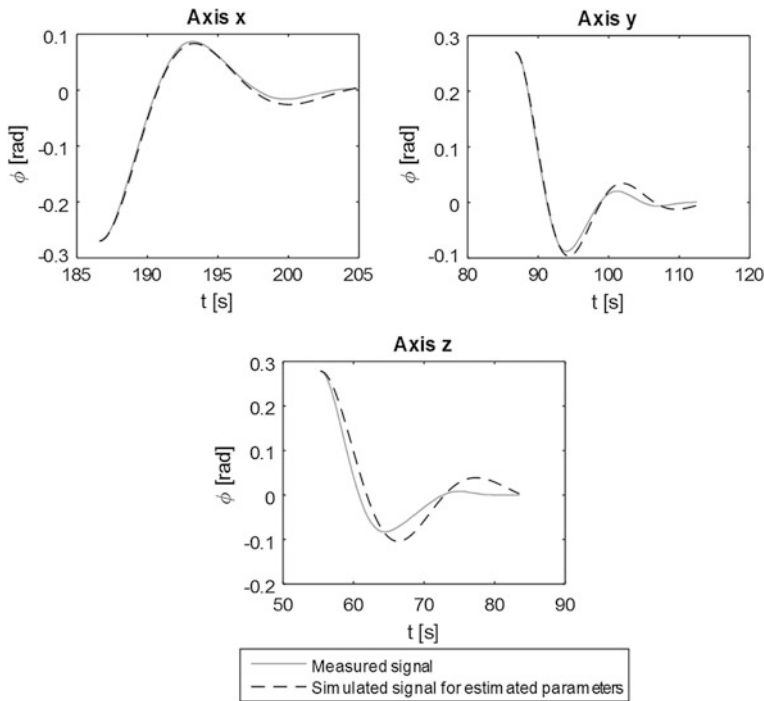
Fig. 5 Response of free-decay structure and corresponding model with estimated parameters
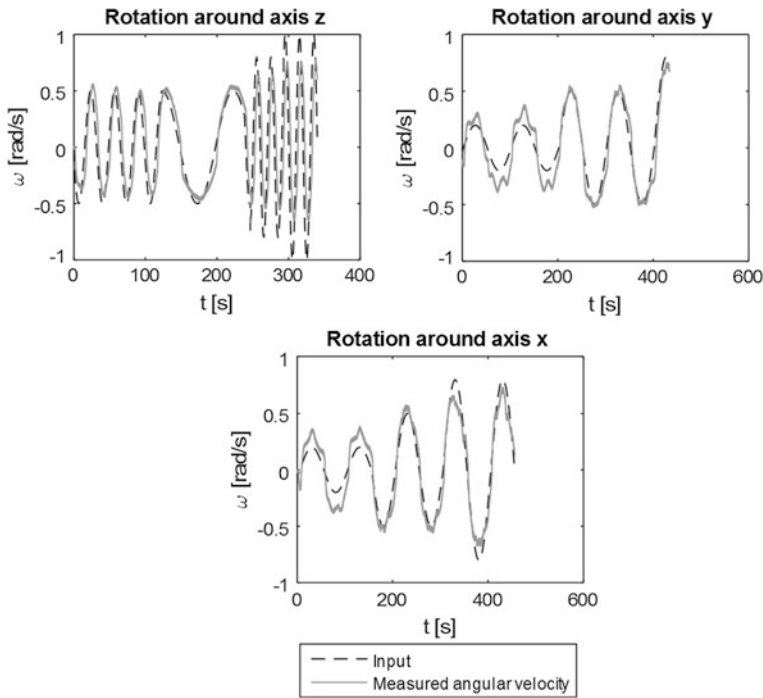
Fig. 6 Sinusoidal input and vehicle response for each axis during oscillatory forced motion experiment

estimated parameters are arranged in Table 1. Table 2 adds a summary of remaining auxiliary parameters.

6 Conclusion

The paper describes an identification procedure for estimating the base inertial parameters of a robotic underwater vehicle. The mathematical model was defined using Kirchhoff's equations with external forces representing the hydrodynamic drags in standard simplification. The meaning of the base inertial parameters is explained. Three methods of identification are described, together with the problem of numerical differentiation. The identification procedure uses two experiments for identification of rotational and translational parameters of given model. The experiments are designed with respect to unknown parameters of the model and
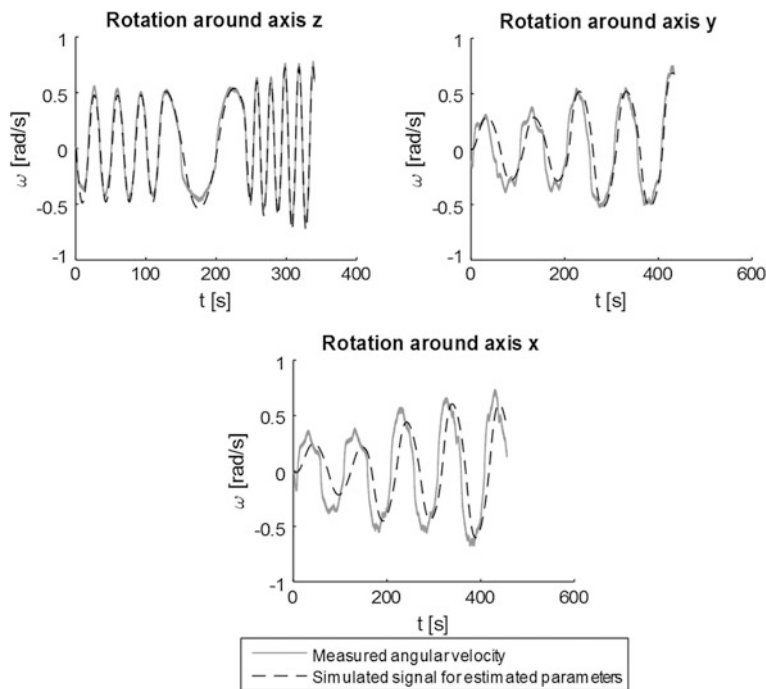
Fig. 7 Vehicle response for each axis (Fig. 6) and corresponding outputs of model with estimated parameters

Table 2 Auxiliary measured parameters

Mounting mechanism					Vehicle body/thrusters			
m_s (kg)	J_s (kg m^2)	β_s (kg m^2/s)	r (m)	l (m)	m_b (kg)	m_w (kg)	C_{tf} (hor/vert)	C_{tb} (hor/vert)
7.88	3.13	0.31	0.5	1.06	40.9	40.9	3.28/6.12	2.22/1.28

according to local conditions. The estimated parameters are validated using simulations, where the measured responses are compared with real measured data. The parametric model was successfully used for design of stable motion controller for a real prototype, which proves that the model is rich enough to be relevant yet simple to be useful for control design.

References

Brockwell, P. J., Dahlhaus, R., & Trindade, A. A. (2005). Modified Burg algorithms for multivariate subset autoregression. *Statistica Sinica, 15*, 197–213.

Chen, Y. (2007). *Modular modeling and control for autonomous underwater vehicle (AUV)* (Thesis of Master of Engineering). Singapore.

Christ, R. D., & Wernli, R. L. Sr. (2007). *The ROV manual: A user guide for observation-class remotely operated vehicles*.

Coleman, T. F., & Li, Y. (1996). An interior, trust region approach for nonlinear minimization subject to Bounds. *SIAM Journal on Optimization, 6*.

Dennis, J. E. Jr. (1977). Nonlinear least-squares. In D. Jacobs (Ed.), *State of the art in numerical analysis*. Academic Press.

Eng, Y. H. (2008). Estimation of the hydrodynamics coefficients of an ROV using free decay pendulum motion. *Engineering Letters, 16*(3), 326–331.

Fossen, T. I. (1994). *Guidance and control of ocean vehicles*. Chichester, England: Wiley.

Holoborodko, P. (2008). *Smooth noise-robust differentiators*. At http://www.holoborodko.com/pavel/numerical-methods/numerical-derivative/smooth-low-noise-differentiators/

Inzartsev, A. V. (2009). *Underwater vehicles*. Austria: In-Tech.

Khalil, W., & Dombre, E. (2002). *Modeling, identification and control of robots* (p. 500). London: Kogan Page Science ed.

Lamb, H. (1932). *Hydrodynamics* (6th ed.). New York: Dover.

Ovalle, D. M., García, J., & Periago, F. (2011). Analysis and numerical simulation of a nonlinear mathematical model for testing the maneuverability capabilities of a submarine. *Nonlinear Analysis: Real World Applications, 12*(3), 1654–1669.

Wang, W. (2007). *Autonomous control of a differential thrust micro rov*.

Yoerger, D. R., Cooke, J. G., & Slotine, J. J. E. (1990). The influence of thruster dynamic on underwater vehicle behavior and their incorporation into control system design. *IEEE Journal of Oceanic Engineering, 15*, 167–168.

Using Online Modelled Spatial Constraints for Pose Estimation in an Industrial Setting

Kenneth Korsgaard Meyer, Adam Wolniakowski, Frederik Hagelskjær, Lilita Kiforenko, Anders Glent Buch, Norbert Krüger, Jimmy Jørgensen and Leon Bodenhagen

Abstract We introduce a vision system that is able to on-line learn spatial constraints to improve pose estimation in terms of correct recognition as well as computational speed. By making use of a simulated industrial robot system performing various pick and place tasks, we show the effect of model building when making use of visual knowledge in terms of visually extracted pose hypotheses as well as action knowledge in terms of pose hypotheses verified by action execution. We show that the use of action knowledge significantly improves the pose estimation process.

Keywords Pose estimation · Online modelling · Pick and place · Stable pose

1 Introduction

Reliable pose estimation is still a challenge, in particularly in unconstrained environments. Finding the correct 6-DOF position of an object towards the camera. The detection method for this will be elaborated in Sect. 3.3. Pose estimation often requires adaptation to different lighting conditions as well as parameter tuning and camera positions. This has multiple reasons: Firstly, even with the most advanced recording systems it is often difficult to generate reliable point clouds in an industrial context where objects are shiny and illumination can only be controlled to a certain degree. Secondly, even if reliable point clouds can be computed, the information provided by one view of a 3D sensor might be insufficient to reliably

K.K. Meyer · A. Wolniakowski · F. Hagelskjær (✉) · L. Kiforenko ·
A.G. Buch · N. Krüger · J. Jørgensen · L. Bodenhagen
Maersk Mc-Kinney Moller Institute, University of Southern Denmark, Odense, Denmark
e-mail: frhag@mmmi.sdu.dk

A.G. Buch
e-mail: anbu@mmmi.sdu.dk

N. Krüger
e-mail: norbert@mmmi.sdu.dk

compute poses due to the lack of shape structure. Humans however perform pose estimation with high reliability even in difficult contexts based on 2D views. One reason for that is that they can make use of context knowledge: humans are able to improve their performance on tasks by learning appropriate context models after very few trials.

In this work, we exploit spatial constraints to improve pose estimation performance, which cover knowledge of poses that can be expected, so called "stable poses", as well as knowledge about the spatial arrangement of items on a conveyor belt or a table. Figure 1 shows a typical situation in which some correct and some wrong pose hypotheses are generated based on the available visual information. By applying knowledge about what poses are actually physically possible and how objects are usually arranged, the wrong hypotheses can be eliminated. In this paper, we show how a model that represents such constraints can be learned on-line in a pose estimation system that is integrated in a complex assembly process.

Once enough data is generated to initialize a suitable model, this model can "kick in" and reduce the search problem by disregarding all pose hypotheses that are found unlikely to be correct according to the model leading to better performance and faster processing time.

We realize two different ways of building models for such constraints: First, we exploit visual knowledge only, i.e., we our model purely from visually extracted pose hypotheses. Here the problem is, however, that many wrong visual hypotheses can lead to wrong model predictions which usually results in drop in performance. As an alternative, we realize modelling based on the already performed actions. Since we can evaluate the success of grasps, we can constrain the model building to pose hypotheses that led to successful action only. We show that by this modelling, we are able to more consistently reduce computational time and increase performance.

The paper is structured as follows; in Sect. 2 an overview of the method compared to the state of the art is given. Section 3 elaborates methods used in the article
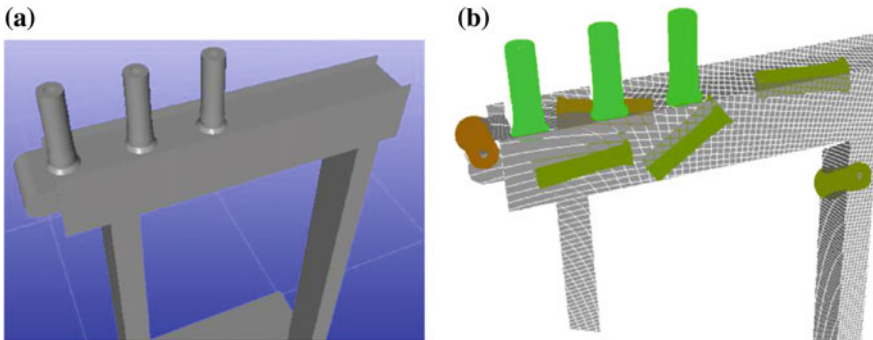
Fig. 1 Result of the pose estimation algorithm. A large number of both correct and incorrect poses are found. **a** Scene image with three rotor caps on the conveyor belt. **b** Found poses in the point cloud, the lighter the color of the rotorcap is the less penalty the estimate has

along with a description of the simulation environment. Section 4 describes the online modelling used in the article. The results are given in Sect. 5. Lastly Sect. 6 contains a conclusion based on the test results of different methods.

2 State of the Art

The use of spatial constraints during 3D pose estimation is relatively rare. Most state of the art methods use more or less complicated hypothesis verification modules for eliminating bad poses. Some noteworthy examples hereof are the recognition systems based on Spin Images (Johnson and Herbert 1999), Tensor Matching (Mian et al. 2006) and more recently the RoPS features (Guo et al. 2013). Common to all these systems, and many other 3D recognition pipelines, is that they rely on individual verifications of pose hypothesis, which are generated from e.g. Hough voting or RANSAC (Fischler and Bolles 1981) pipelines.

Instead, some works have used contextual information during pose hypothesis verification to arrive at a better scene interpretation. In (Papazov and Burschka 2010), a RANSAC scheme is used for gathering pose hypothesis, based on matches from point pair features, which were also successfully used in (Drost et al. 2010). Then, a conflict graph is constructed using the set of pose candidates, and conflicting hypotheses are removed by non-maximum suppression. In Aldoma et al. (2012) a global hypothesis verification algorithm was used. In this work, a set of hypothesis poses is assigned a global cost based on a variety of cues, e.g. occlusions and overlaps. A random subset of hypotheses is then evaluated using this cost function, and a simulated annealing process is used for minimizing the global cost function for selecting the set of hypothesis that best describes the scene in a global manner.

This work builds upon the prior work of Jorgensen et al. (2015), which deals with stable poses of an object. We use such stable poses in our constraint framework in an active manner. Additionally, we compare the use of both vision and action information, and we use an online modeling scheme instead of requiring a sophisticated and error-prone prior modeling scheme.

3 Methods

In this section we will describe the types of constraints that we use for the point cloud segmentation, and methods that were used to construct their models. We first introduce the concept of a *stable pose* in Sect. 3.1, which allows us to classify and filter the detected samples by their orientation. Then, we arrive at spatial constraints constructed from the filtered samples: the *point, line*, and *plane* constraints in Sect. 3.2. All of these spatial constraint models, as well as the stable pose model, are learned online using the RANSAC algorithm. The threshold values ε on the fit

error function E specific for each of the models are used to determine when a detection is part of a model. In Sect. 3.3, we then briefly describe the simulation framework used in our experiments.

3.1 The Concept of a Stable Pose

We use the concept of a stable pose to filter out false detections of objects in the point cloud, by accepting only poses which are physically possible.

The concept of a stable pose of an object, introduced in (Jorgensen et al. 2015), refers to a subset of object orientations in SO(3) corresponding to particular support modes of that object. Stable poses arise due to physical constraints imposed on the object by its context, e.g. gravity and contact forces. For example, for a surface-to-surface contact on a plane perpendicular to the direction of the gravity vector, a free standing object is deprived of 2 degrees of freedom in rotation (see Fig. 2a). In Fig. 2c, a case of a stable pose based on a rolling contact between a cylinder and a plane is presented. In that case, the object is deprived of 1 degree of rotational freedom.

A stable pose model M_S is defined based on observing spatial structures in the distribution of unit vectors subjected to rotational transformations from a candidate subset of SO(3). Figure 3 presents sample distributions of unit vectors.

The stable pose model M_S is defined in terms of planes p_x, p_y, p_z detected in sets of points $\mathbf{x}*$, $\mathbf{y}*$, $\mathbf{z}*$ obtained by transforming respective unit vectors with sampled rotational transformations R_i, e.g. $x_i = R_i$. We define the stable pose model M_S as:

$$M_S = \{n, \ d_x, \ d_y, \ d_z\}, \tag{1}$$

where $\mathbf{n}$ is the normal vector common for all of the detected planes, and d_i are distances from the origin to the ith plane respectively.

The stable pose model requires 3 samples at minimum to be fitted by the RANSAC algorithm. We define the fit error function for the stable pose model as:
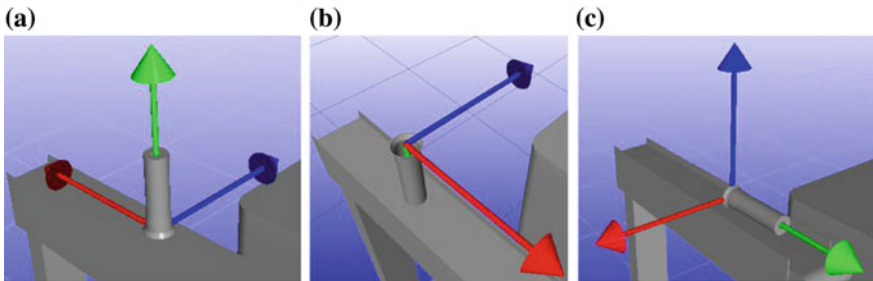
Fig. 2 Stable poses of the rotorcap object. **a** Stable pose 1. **b** Stable pose 2. **c** Stable pose 3
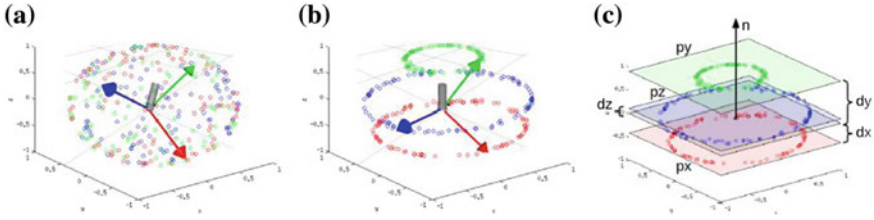
Fig. 3 Distributions of 100 unit vectors subjected to sampled rotational trans-formations R: **a** no stable pose, any xyz position is possible, (R $\in$ *SO*(3)), **b** stable pose, only 1-degree of freedom, (R $\in$ S). Panel **c** shows the planes p_x, p_y, p_z detected for a stable pose model $M_S = \{\mathbf{n}, d_x, d_y, d_z\}$

$$E(\mathbf{R}) = \left\| [R_{11}, R_{22}, R_{33}] - [d_x, d_y, d_z] \right\| \tag{2}$$

3.2 Modeling Spatial Constraints

We construct several types of spatial constraints based on samples obtained from the object pose detection in addition to the stable pose constraint. We also require that the samples match a stable pose model S. The constraints we use are based on definitions introduced in Jorgensen et al. (2015):

Point constraint represents positions arbitrarily close to a point in R^3. This is often a case for objects located in fixtures, or supplied by part feeders. The point constraint is defined as:

$$C_{\text{point}} = \left\{ p \middle| \left\| p_t - M_{\text{point}} \right\| < \varepsilon_{\text{point}} \wedge p_R \in S \right\}, \tag{3}$$

where M_{point} represents the constraint model defined as a point location in R^3, $p = (p_t, p_R)$ is a candidate pose consisting of translation and rotation parts, and $\varepsilon_{\text{point}} = 0.1$ is the constraint threshold.

Line constraint. The line constraint best matches scenarios where the objects are placed in a row, e.g. on a conveyor belt. Hence, we expect the poses to be arbitrarily close to a straight line in R^3. The line constraint is defined as:

$$C_{\text{line}} = \left\{ p \middle| \left\| (p_t - v) - ((p_t - v) \cdot s)s \right\| < \varepsilon_{\text{line}} \wedge p_R \in S \right\}, \tag{4}$$

where the constraint model $M_{\text{line}} = (\mathbf{v}, \mathbf{s})$ consists of a point on the model line $\mathbf{v}$, and the direction vector $\mathbf{s}$. The constraint threshold is $\varepsilon_{\text{line}} = 0.1$.

Plane constraint fits poses of the objects which are arranged on a flat surface. This could be, for example, a table, or—in our case—the fixture for storing the rotorcaps. The plane constraint is defined as:

$$C_{\mathrm{plane}} = \left\{ p \,\big|\, \| (v - p_t) \cdot n) \| < \varepsilon_{\mathrm{plane}} \wedge p_R \in S \right\}, \tag{5}$$

where the table constraint model $M_{\mathrm{plane}} = (\mathbf{v}, \mathbf{n})$ includes a point on the plane $\mathbf{v}$, and the plane normal vector n. For the plane constraint, the threshold is $\varepsilon_{\mathrm{plane}} = 0.15$.

For the spatial constraint model fitting, the error function E is determined based on the distance of the translational part p_t of the sample p to the point, line or plane used as respective constraint model.

3.3 The Simulation Environment

The experiments are based on a case of a pick and place operation. We work in the context of an industrial assembly scenario (see Fig. 4a). In the scenario, a number of rotorcaps need to be picked up by a robotic arm and placed in a fixture (see Fig. 4b). The robot is the modular platform "Little Helper 4", developed at Aalborg university, with a 6-dof UR6-85-5-a robot arm mounted (Madsen et al. 2015). On the robot arm, a universal electric gripper WSG 50 from Weiss Robotics is attached. A camera is mounted on a pan tilt unit. The whole set-up is a simulation of a real set-up (see Fig. 4c).

The objects to be manipulated are rotorcaps, which are a part of an electric motor assembling process. In three different scenarios (see Fig. 5), the rotorcaps are placed on a conveyor belt or on a table. Once the rotorcaps have been grasped by the robotic system, they are placed in a fixture which is a (4×3) grid of male counterparts to the rotorcaps (or: *rotorfixes*—see Fig. 4b). If a grasped rotorcap fails to be placed in the fixture, it is dropped in the nearby crate (red box in Fig. 4a).

A point cloud of the scene is generated by a simulated Kinect camera, and the best possible poses are found by an altered version of the detection algorithm in (Buch et al. 2013). The detection method gives a number of possible poses each with a penalty of the detection. The grasp selection is based on a database of grasps generated for the rotorcaps and the gripper. Using the pose with the smallest penalty
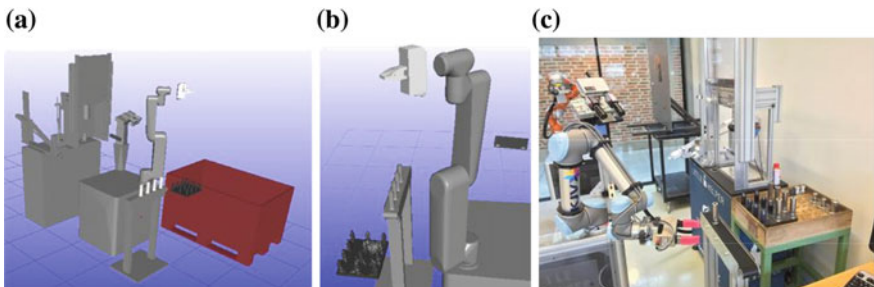
Fig. 4 The setup used in the simulation along with the actual scene. **a** Full scene used in simulation. **b** Robotic arm and fixture. **c** Actual scene simulation is based on
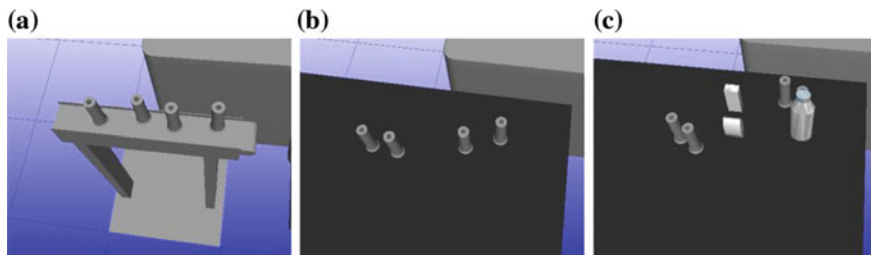
Fig. 5 Simulated rotorcap picking scenarios. In the scenario 1, image (**a**), the rotorcaps are supplied by a conveyor belt. In scenario 2, image (**b**), the rotorcaps are positioned randomly on a table and scenario 3, image (**c**), are scenario 2 with other objects placed randomly on the table
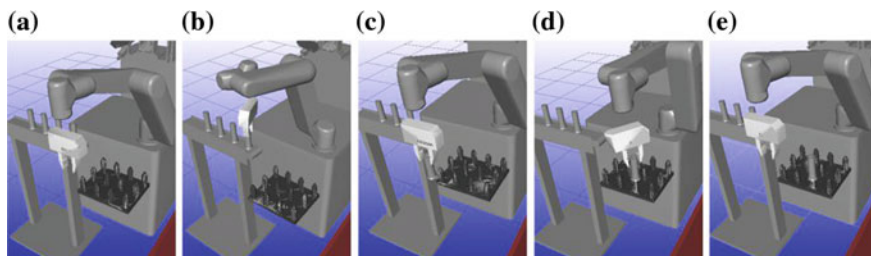
Fig. 6 Different states of the robot-arm during pick and place actions (scenario 1). **a** Default position, **b** picking, **c** default position, **d** placing, **e** default position

from the pose estimation all grasps are tested to find a possible grasp. If this fails the next best pose is tested, etc.

In order to perform a place action, pose estimation is employed again. The camera view is changed to gaze at the fixture, and poses are found of individual rotor fixes. The desired position of the rotorcap is then right above the fixture farthest away from the robot, to ensure free passage for the next rotorcap. Once the position is found, a path is planned to position the rotorcap Fig. 6.

4 On-line Modelling and Application of Constraints

In our experiments, we simulate a simple pick-and-place process. To improve the performance of pose estimation, we employ a set of constraints which we introduced in Sect. 3.2. The constraint models are built online, using the data from the simulation. Once enough data is gathered to ensure the validity of the constraint models, the models are used to segment the point clouds obtained from the cameras, thereby speeding up the process. Since during the process, we perform the pose estimation twice (first to detect the objects to be grasped, and then the fixture to put

them in), in one simulated scenario we build two separate constraint models, one for the rotorcaps and the other for the fixtures.

The constraint models are updated with the information obtained from the pose estimation and the action execution. After a number of data points sufficient to construct models are obtained by iterating the pick-and-place process, we use the constructed constraints to segment the raw point clouds, by only accepting points in the neighborhood of the constraint model:

$$p_{\text{accepted}} = \left\{ p \middle| \|p - M_{\text{constraint}})\|_\infty < \varepsilon_{\text{seg}} \right\}, \tag{6}$$

where p is an element of the point cloud, $M_{\text{constraint}}$ is the constraint model, and ε_{seg} is the distance threshold set to the value of ε of a corresponding model (see Eqs. 3–5). The process of segmentation is shown in Fig. 7.

We perform two versions of the experiments. At first, we build the constraint models based on all pose detections obtained from the vision component of the system (vision based). In the second version, we only use the data points which were confirmed to be associated with successfully executed actions (action based). The knowledge of the status of the action execution is based on the ground truth, i.e. we only consider the actions to be successful, if the objects are placed within an arbitrary threshold of $\varepsilon_{\text{success}} = 0.01$ m to their expected position in the scene:

$$p_{\text{successful}} = \left\{ p \middle| \left\|p^t - p^t_{\text{expected}}\right\|_1 < \varepsilon_{\text{success}} \right\}. \tag{7}$$
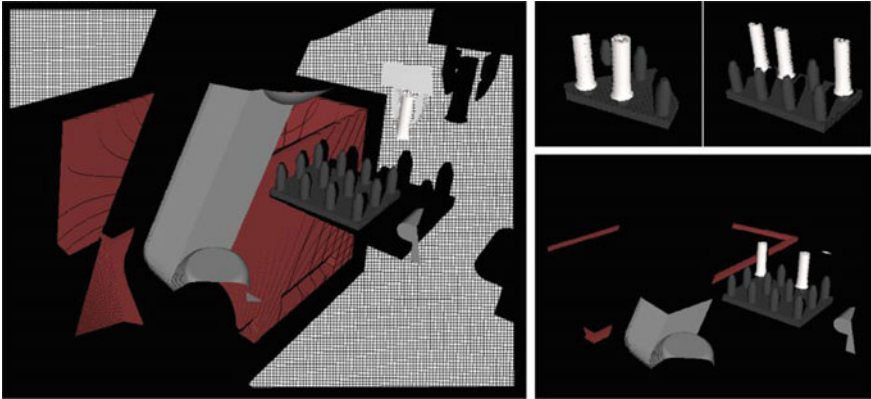
Fig. 7 Result of the different segmentation methods on a scene, *left image*. In the *top right corner* from *left* to *right* is point and line segmentation. And the *bottom right corner* is the plane segmentation

5 Results

Figure 8 shows the typical performance of the system over time. After 11 iterations enough data is generated to apply a suitable model, the search can be constrained and by that performed faster.

Table 1 shows the results for the 3 different scenarios of pick actions and the place action. Performance is given in terms of percentage of successful actions as well as cycle time for the pose estimation. In the "Baseline test" no constraints are applied. As can be seen, performance as well as execution time is quite different for the different scenarios and actions. The three columns under "Vision based" show the results, when we apply the models based on visual pose hypotheses.

Fig. 8 Example of the speed increase that occurs for the pose estimation when enough trials have been done for the segmentation to begin. Fixture location using plane segmentation with visual based modelling
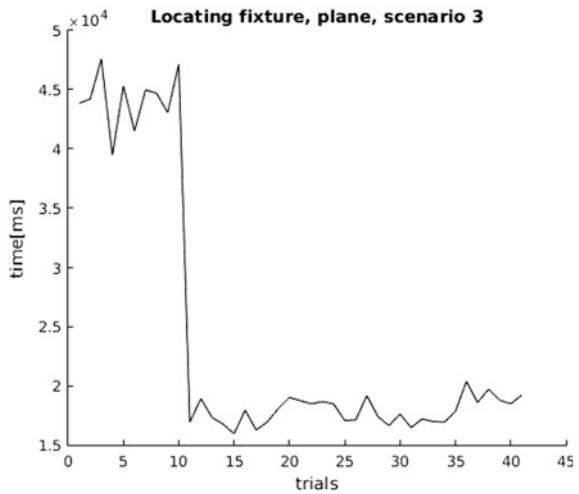
Table 1 Resulting performance of the pick and place action using the different segmentation methods for the vision system

	Scenario 1		Scenario 2		Scenario 3		Place	
	Perf. in %	Time in s	Perf. in %	Time in s	Perf. in %	Time in s	Perf. in %	Time in s
Baseline	**90.2**	**19.7**	**88.2**	**41.0**	**49.1**	**43.9**	**100**	**42.8**
Vision based								
Point	47.5	10.5	49.6	23.7	26.0	15.4	98.3	16.8
Line	**87.8**	**14.0**	17.9	15.0	11.9	14.8	**100**	**13.1**
Plane	82.9	14.4	**91.7**	**41.6**	**49.1**	**40.3**	**100**	17.1
Action based								
Point	88.7	14.4	**90.3**	**35.9**	51.4	35.6	100	17.8
Line	**96.2**	**12.4**	41.3	16.9	51.7	34.9	98.7	15.5
Plane	93.8	13.5	89.7	41.7	**54.8**	**45.5**	**100**	16.5

Best performance in successive grasps for each segmentation method is shown in boldface
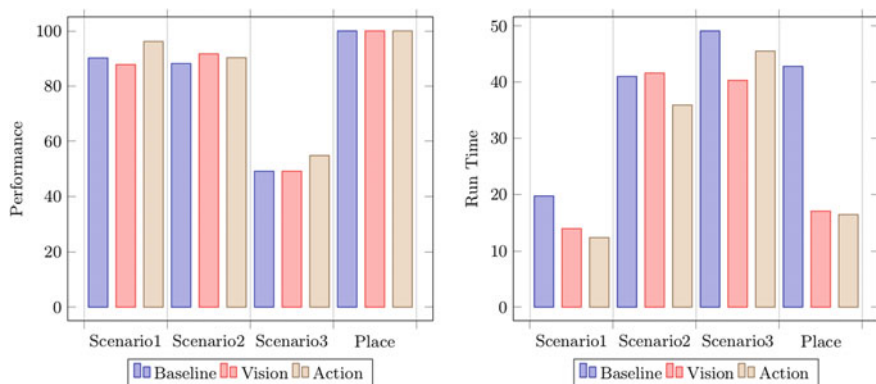
Fig. 9 Results in performance and run time for the best performing methods in both vision and action compared to the baseline

Under "Action based", we only use pose hypotheses when actions have been executed successfully to build our models.

When applying models generated by vision based hypotheses, the line and plane constrained systems can match the baseline system, but only in scenario 1 for the line constrained and only scenario 2 and 3 for the plane constrained. However, significant speed-ups can be achieved with the point and line model.

Looking at the "Action based" results, we achieve better performance for the line constraint in scenario 1 and for the plane constraint in all three scenarios. For the best performing the computational time is improved in scenario 1, slightly improved in scenario 2 and slightly worse in scenario 3. For the place action, computational time is improved by a factor of around 2.5 with same perfect performance. In Fig. 9 the results for each of the best optimizations are shown. Here it is clear that using the model it is possible to speed up computation while maintaining performance.

6   Conclusion

We have described a vision system in which spatial constraints are exploited to improve pose estimation in terms of detection as well as computational speed. The constraints are learned on-line and utilized once enough evidence for the model is accumulated. We showed that using vision knowledge alone, only a rather small improvement could be achieved but that by using pose knowledge that has been verified by action a significant speed up was possible. The simulation framework allowed us to apply such action verified constraints. The next logical step would be to apply the generated models on the real set-up.

References

Aldoma, A., Tombari, F., Di Stefano, L., & Vincze, M. (2012). A global hypotheses verification method for 3D object recognition. In *Computer Vision–ECCV 2012* (pp. 511–524). Heidelberg: Springer.

Buch, A. G., Kraft, D., Kamarainen, J. K., Petersen, H. G., & Kruger, N. (2013, May). Pose estimation using local structure-specific shape and appearance context. In *2013 IEEE International Conference on Robotics and Automation* (*ICRA*), (pp. 2080–2087). IEEE.

Drost, B., Ulrich, M., Navab, N., & Ilic, S. (2010). *Model globally, match locally: Efficient and robust 3D object recognition.*

Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM, 24*(6), 381–395.

Guo, Y., Sohel, F., Bennamoun, M., Lu, M., & Wan, J. (2013). Rotational projection statistics for 3D local surface description and object recognition. *International Journal of Computer Vision, 105*(1), 63–86.

Johnson, A. E., & Hebert, M. (1999). Using spin images for efficient object recognition in cluttered 3D scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 21*(5), 433–449.

Jorgensen, J. A., Rukavishnikova, N., Kruger, N., & Petersen, H. G. (2015, March). Spatial constraint identification of parts in SE3 for action optimization. In *2015 IEEE International Conference on Industrial Technology* (*ICIT*), (pp. 474–480). IEEE.

Madsen, O., Bøgh, S., Schou, C., Andersen, R. S., Damgaard, J. S., Pedersen, M. R., et al. (2015). Integration of mobile manipulators in an industrial production. *Industrial Robot: An International Journal, 42*(1), 11–18.

Mian, A. S., Bennamoun, M., & Owens, R. (2006). Three-dimensional model-based object recognition and segmentation in cluttered scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 28*(10), 1584–1601.

Papazov, C., & Burschka, D. (2010). An efficient RANSAC for 3D object recognition in noisy and occluded scenes. In *Computer Vision–ACCV 2010* (pp. 135–148). Heidelberg: Springer.

Comparison Study of Industrial Robots for High-Speed Machining

Alexandr Klimchik, Alexandre Ambiehl, Sebastien Garnier, Benoit Furet and Anatol Pashkevich

Abstract The paper presents methodology for comparison of industrial robots used for high-speed machining. Particular attention is paid to the robot accuracy in milling operation and evaluation robot capacity to perform the task with desired precision. In contrast to other works, the robot performance is evaluated using an industrial standard that is based on the distortion of the circular shape. The developed approach is applied to four industrial robots of KUKA family, which have been compared with respect to the machining precision.

Keywords Robot-based machining · Circularity index · Industrial robot · Stiffness model · Compliance errors · Robot comparison

1 Introduction

High-speed machining is quite a new application of industrial robots. As follows form related study (Chen and Dong 2013), the machining segment represents less than 5 % of the total market of the industrial robots, but this share is continuously increasing. So, replacement of conventional CNC machines by more competitive industrial robots becomes more and more attractive. The main restraint here is

A. Klimchik (✉)
Innopolis University, Universitetskaya St, 1, Innopolis, Tatarstan 420500, Russia
e-mail: a.klimchik@innopolis.ru

A. Ambiehl · S. Garnier · B. Furet · A. Pashkevich
Institut de Recherches en Communications et en Cybernétique de Nantes,
UMR CNRS 6597, 1 rue de la Noe, 44321 Nantes, France

A. Ambiehl · S. Garnier · B. Furet
Université de Nantes, Chemin de la Censive du Tertre, 44300 Nantes, France

A. Ambiehl
Gébé 2 Productique, Parc Activite Vendée Sud Loire, 85600 Boufféré, France

A. Pashkevich
Ecole des Mines de Nantes, 4 rue Alfred-Kastler, 44307 Nantes, France

rather limited knowledge of robotics by potential customers and lack of competence of the of robotic cell end-users. On the other side, the research labs have already confirmed that CNC machines replacement by robots gives essential benefits, which must be clarified for practicing engineers. For this reason, this paper proposes an industry oriented technique for evaluation of the robot capacities in machining, which can be used as the base for the related comparison study.

In contrast to conventional CNC machines, robots are able to process complex bulky 3D shapes and provide large and easily extendable workspace that can be modified by adding extra axes. Besides, the same workspace can be shared by several robots. However, the robot trajectory generation is much more complex task compared to the Cartesian machines since mapping from the actuator space to the operational space is highly non-linear. Another difficulty arises because of robot redundancy with respect to the technological process. In fact, conventional machining process requires 5 dof only while most of industrial robots have 6 actuators. This redundancy can be used to optimize the tool path, to improve the trajectory smoothness (Zargarbashi et al. 2012a) or to minimize the impact of machining forces (Zargarbashi et al. 2012b; Vosniakos and Matsas 2010).

Another difficulty of robot application in machining is related to non-negligible compliance of robotic manipulators. In some cases the end-effector deflections due to the influence of the cutting forces may overcome 10 mm (Matsuoka et al. 1999). To reduce them, robot manufactures pay particular attention to the manipulator stiffness and compensation of the compliance errors. To improve the manipulator rigidity, designers are obliged either to increase the link cross-sections or to use advanced composite materials. The first solution leads to increasing of moving masses and consequent reduction of dynamic properties. Utilization of composite materials essentially influences on the robot price and decreases its market competitiveness. Nevertheless, both ways improve the link stiffness only, while the major manipulator elasticity is often concentrated in the actuator gears (Dumas et al. 2012) and can be hardly improved in practice. Another method of the compliance error reduction is based on the mechanical gravity compensators. However, this solution does not allow compensating the impact of the machining forces. To overcome the problem of elastic deformations in the actuator gears, robot manufactures tends to use secondary encoders attached to the motor shaft (Devlieg 2010) that allow to modify the actuator input in order to compensate the gear compliance. It is obvious that this approach also increases the robot price. According to our experience, the double encoders enable compensating about 65 % of the compliance errors. The main reason for this is that the robot link deformations are outside of the double encoder observability. It is clear that for the high-speed milling, where the cutting forces are high enough to cause deflection of several millimeters, such level of error compensation is insufficient. In this case, it is reasonable to apply the off-line error compensation technique (Klimchik et al. 2013a; Chen et al. 2013) based on the modifying the reference trajectory used as the controller input. As follows from our previous research, this approach is very efficient. In particular, the off-line technique based on the simple (reduced) manipulator stiffness model allows user to compensate 85–90 % of the end-effector deflections (Klimchik et al. 2012),

while the complete stiffness model ensures the compensation level of about 95 % (Klimchik et al. 2015). However, robot manufactures usually do not provide customers with the manipulator stiffness parameters, so they must carry out dedicated experimental study (Abele et al. 2007; Nubiola and Bonev 2013). In this paper, the above mentioned problem will be also considered.

To advance robot application in machining, end-user should be provided with clear and efficient tool allowing to evaluate the final product quality expressed via the level of the end-effector deflections caused by the manipulator elasticity. It is evident that usual approach based on different performance measures extracted from the Cartesian stiffness matrix (Guo et al. 2015; Nagai and Liu 2008) are not suitable here. For this reason, this paper proposes an industry oriented technique allowing to examine particular robot suitability for a give machining task and to compare several robot-based implementations.

2 Robot-Based Machining

Machining with robots is an intersection of two engineering fields: conventional machining and robotics. Machining sector usually prefers for these operations Cartesian CNC machines that provide end-users with high repeatability (2 μm) and good precision (5 μm). Traditionally, they are used for processing of metal parts from parallelepiped-like crude products with high material removal rate. Contemporary CNC machines possess quite large workspace allowing essentially increase an application area. Besides, their efficiency was also proved for processing of composite materials that are utilized more and more due to perfect mass-to-strength ratio. In addition to milling, finishing and trimming operations can be also performed by CNC machines. Nevertheless, in spite of numerous advantages, the CNC machines remain very expensive and their workspace is limited and cannot be extended, which is crucial for aeronautic and shipbuilding. This motivates users to find an alternative solution.

One of the promising ways to overcome the above mentioned difficulties is replacing the CNC machines by industrial robots, whose cost is competitive and workspace can be easily extended (by adding extra actuated axes). An example for such an application is presented in Fig. 1. Traditionally, the market of industrial robots is shared between handling, pick and place, assembling and welding. The processing (including machining), represents insignificant part of the market, less than 5 %. According to PWC study (McCutcheon and Pethick 2014) these shares will remain the same in the nearest future. Nevertheless, the share of robot-based machining is continuously growing. Large part of this market share corresponds to trimming that was traditionally a high-qualified manual work, but nowadays the robots become competitive here due to increasing of their accuracy. For machining, robots are attractive due to their large and extendable workspace and competitive price that makes them a cost-effective solution for machining of large dimension parts. However, the main obstacle for robots' utilization in machining is their

Fig. 1 Example of
machining process with robot
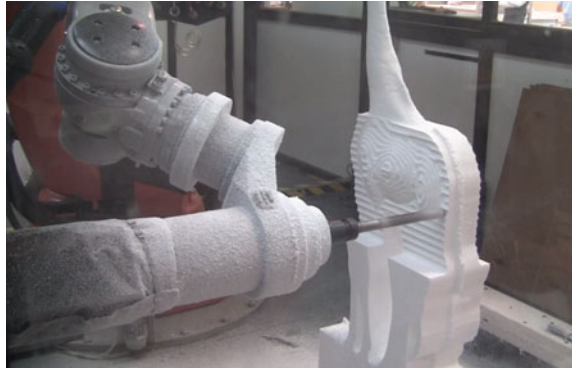
Fig. 2 Cutting force in the
machining process
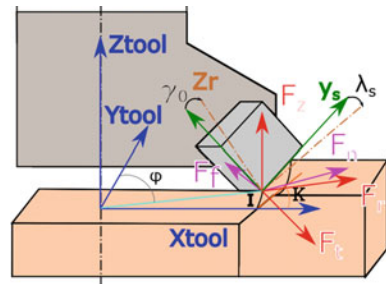
relatively low accuracy (about 0.7 mm) and repeatability (about 0.2 mm) compared
to the CNC machines. Nevertheless, there are a number of efficient solutions to
reduce manipulator positioning errors that were discovered in research labs and
progressively applied in industrial environment. The latter allows robots to compete
with CNC machines in terms of accuracy, while providing essentially larger
workspace.

The problem of the machining process modeling has been known since pioneer
work (Merchant 1945), where Merchant used principle of minimum angle to
develop an analytical 1D cutting force model based on chip thickness and material
removal behavior. Latter, this model has been extended to the cases of 2D and 3D
(Doi and Kato 1955). The further advances in this area lead to mechanistic models
(Koenigsberger and Sabberwal 1961). These models explain the machining process
as the function of the cutting tool geometry and process parameters (such as feed
rate, spindle speed, etc.). Further, the cutting force model was integration in the
commercial software (Weck et al. 1994; Shirase and Altintaş 1996). These models
were able to predict deformation of the tool during machining and adapt cutting
conditions (Wan et al. 2014; Araujo et al. 2015; Ali et al. 2013).

For the face milling application, the carbide insert tool cutting force model is the
most advanced one allowing analyzing the path of each teeth and to predict cor-
responding force (Cheng et al. 1997). The geometry of technological process and
corresponding cutting forces are given in Fig. 2. In the tool coordinate system, the

instantaneous force $\mathbf{F}^i = [F_n^i, 0, F_f^i]^T$ is defined by the normal force F_n acting perpendicular to the tool face and the friction force F_f acting along the face. Here, the superscript "i" indicates the nature of the force, i.e. instantaneous. It should be noted that the friction force F_f^i depends on normal force F_n^i and friction coefficient k_f, i.e. $F_f^i = k_f \cdot F_n^i$. The normal force F_n^i depends on the engagement angle φ, feed rate f_z and the cut depth a_p

$$F_n^i(\varphi) = K_n \cdot f_z \cdot a_p \cdot \sin \varphi \tag{1}$$

where K_n is a specific cutting coefficient, which depends on the material properties and the cutting tool.

The instantaneous force $\mathbf{F}^i$ is usually presented in the cylindrical coordinate system as $\mathbf{F}_0^i = [F_r^i, F_t^i, F_z^i]^T$, where F_r, F_t, F_z are the radial, tangential and axial forces, respectively. The correspondence between this forces are defined by the rotation matrix $\mathbf{R}(\kappa, \lambda_s, \gamma_0)$ that depends on the tool orientation, i.e. the entering angle κ, the helix angle λ_s and the cutting angle γ_0

$$\mathbf{F}_0^i = \mathbf{R}(\kappa, \lambda_s, \gamma_0) \cdot \mathbf{F}^i \tag{2}$$

Using cylindrical coordinate system, it is possible to project the cutting force to the workpiece frame,

$$\mathbf{F}_C^i = \mathbf{R}_z(\varphi) \cdot \mathbf{F}_0^i \tag{3}$$

where $\mathbf{F}_C$ is the cutting force in the Cartesian space, $\mathbf{R}_z$ is the homogeneous rotation matrix and φ is the engagement angle. The cutting force varies with the engagement angle φ, so in practice the force usually is evaluated for the complete tool revolution and can be computed as

$$F_n = \int_{\varphi_S}^{\varphi_E} F_n^i \mathrm{d}\varphi \tag{4}$$

where φ_S and φ_E are the start and end angles for the tool engagement. The engagement angles φ_S, φ_E are defined by the technological task and cutting tool diameter.

Similarly, the cutting force in the Cartesian coordinate system can be computed as

$$\mathbf{F}_C = \int_{\varphi_S}^{\varphi_E} \mathbf{R}_z(\varphi) \cdot \mathbf{R}(\kappa, \lambda_s, \gamma_0) \cdot \mathbf{F}_0^i \mathrm{d}\varphi \tag{5}$$

Taking into account that $\mathbf{F} = F_n \cdot \mathbf{C}$ with $\mathbf{C} = [1, \, 0, \, k_f]^T$, one can rewrite the cutting force expression in the following form

$$\mathbf{F}_C = \int\limits_{\varphi_S}^{\varphi_E} F_n(\varphi) \cdot \mathbf{R}_z(\varphi) \mathrm{d}\varphi \cdot \mathbf{R}(\kappa, \lambda_s, \gamma_0) \cdot \mathbf{C} \tag{6}$$

The latter can be used to compute the cutting force that is applied to the robot end-effector.

Once the cutting force is known, it is possible to estimate the torques generated by the cutting force at the tool reference frame $\mathbf{M} = \mathbf{r} \times \mathbf{F}_0$. Here, $\mathbf{r} = [r\cos\varphi, \, r\sin\varphi, \, 0]^T$, r is the tool radius. In this case, the torque with respect to the tool contact point can be computed as

$$\mathbf{M} = F_n \cdot (\mathbf{r}\times) \cdot \mathbf{R}(\kappa, \lambda_s, \gamma_0) \cdot \mathbf{C} \tag{7}$$

where $(\mathbf{r}\times)$ is the skew-symmetric matrix based on the vector $\mathbf{r}$. Further, this torque can be transferred to the robot end-effector taking into account the tool and spindle geometry.

3 Compliance Errors and Their Estimation

3.1 Manipulator Positioning Errors Under the Loading

In robotics, manipulator stiffness is usually described by the Cartesian stiffness matrix $\mathbf{K}_C$ that allows user to compute the end-effector deflections $\Delta\mathbf{t}$ as

$$\Delta\mathbf{t} = \mathbf{K}_C^{-1} \cdot \mathbf{F} \tag{8}$$

for given external force $\mathbf{F}$. Using the VJM-based technique (Pashkevich et al. 2011), the stiffness model of a typical industrial robot is presented as a serial chain containing rigid links separated by actuators and virtual springs (Kövecses and Angeles 2007) as shown in Fig. 3.

To compute the Cartesian stiffness matrix, it is necessary to consider simultaneously the extended geometric model $\mathbf{t} = \mathbf{g}(\mathbf{q}, \boldsymbol{\theta})$ and static equilibrium equation $\mathbf{J}_\theta^T \cdot \mathbf{F} = \mathbf{K}_\theta \cdot \boldsymbol{\theta}$. The first of them allows computing the end-effector location $\mathbf{t}$ for given actuator coordinates $\mathbf{q}$ and virtual joint deflections $\boldsymbol{\theta}$ caused by the loading $\mathbf{F}$. The static equilibrium equation allows finding relation between the external force and the deflections of the virtual joints. It includes the matrix $\mathbf{K}_\theta$ describing the virtual joint elasticities and the Jacobian of the extended geometric model $\mathbf{J}_\theta = \partial\mathbf{g}(\mathbf{q}, \boldsymbol{\theta})/\partial\boldsymbol{\theta}$. Simultaneous solution of the above mentioned equations leads to the following expression for the desired Cartesian stiffness matrix
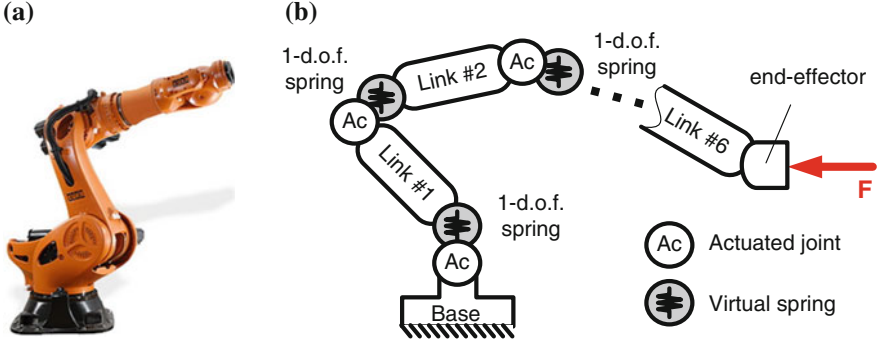
Fig. 3 Typical industrial robot and its VJM-based stiffness model. **a** Typical industrial robot. **b** VJM model of serial robot

$$\mathbf{K}_C = \left(\mathbf{J}_\theta \cdot \mathbf{K}_\theta^{-1} \cdot \mathbf{J}_\theta^T\right)^{-1} \tag{9}$$

In more general case, when the external loading $\mathbf{F}$ is rather high, the Cartesian stiffness matrix also depends on the Hessian $\mathbf{H}_{\theta\theta}$ (Klimchik et al. 2014a). So, the above equation is replaced by.

$$\mathbf{K}_C = \left(\mathbf{J}_\theta \cdot (\mathbf{K}_\theta - \mathbf{H}_{\theta\theta})^{-1} \cdot \mathbf{J}_\theta^T\right)^{-1} \tag{10}$$

where $\mathbf{H}_{\theta\theta} = \partial^2(\mathbf{g}(\mathbf{q}, \boldsymbol{\theta}) \cdot \mathbf{F})/\partial\boldsymbol{\theta}^2$ describes modification of manipulator elasticity due to applied loading.

The above expressions for the Cartesian stiffness matrix were derived assuming that joint stiffness matrix $\mathbf{K}_\theta$ is constant. However, if the manipulator includes the gravity compensators, the equivalent joint stiffness coefficients become configuration dependent and the joint stiffness matrix is presented as a sum

$$\mathbf{K}_\theta = \mathbf{K}_\theta^0 + \mathbf{K}_\theta^{GC}(\mathbf{q}) \tag{11}$$

where the first term $\mathbf{K}_\theta^0$ is constant and corresponds to the manipulator without compensator; and the second term $\mathbf{K}_\theta^{GC}(\mathbf{q})$ depends on the manipulator configuration and describes equivalent elasticity of the gravity compensator with respect to the virtual joints. More details concerning computing of these matrices are given in our previous work (Klimchik et al. 2013b).

3.2 Evaluation of Machining Accuracy Using Circularity

In the industrial practice, there exist different norms to evaluate the final product quality. For example, ISO 12780 and 12781 define the path straightness, the surface

flatness and the path roundness that in some cases is also called as the *circularity*. From our experience, the circularity index is the best one for evaluating the machining process quality, since for linear trajectories the force magnitude/direction are constant and their impact can be easily compensated in the control program (Feng et al. 2015). In contrast, for the circular path the cutting force direction varies and may lead to irregular path distortion that can be hardly compensated. For this reason, the circularity index will be used to evaluate the capacity of industrial robot to perform the machining task.

According to ISO 12181, the circularity evaluation includes two steps: obtaining a reference circle and estimation of the path deviations with respect to this circle. There are four methods to define the reference circle: Minimum Circumscribed Circle (MCC), Maximum Inscribed Circle (MIC), Minimum Zone Circles (MZC) and Least Squares Circle (LSC). They are illustrated by Fig. 4. For all of them, the circularity evaluates the distance between two circles in accordance with the equation

$$\rho = r_{\max} - r_{\min} \tag{12}$$

where $r_{\max}$ and $r_{\min}$ are the radii of the circumscribed and inscribed circles, respectively. The principal difference is related to the circle centers that are computed using different methods. For example, for MIC the center point is computed for the maximum inscribed circle and it is also used for the minimum circumscribed one. In the MCC method, the center is computed for the minimum circumscribed circle and the inscribed circle is build using the same center point. In the case of LSC, the inscribed and circumscribed circles are found for the center point obtained for the least square circle. In contrast, the MZC method uses a center point for

Fig. 4 Circularity evaluation using different industrial norms
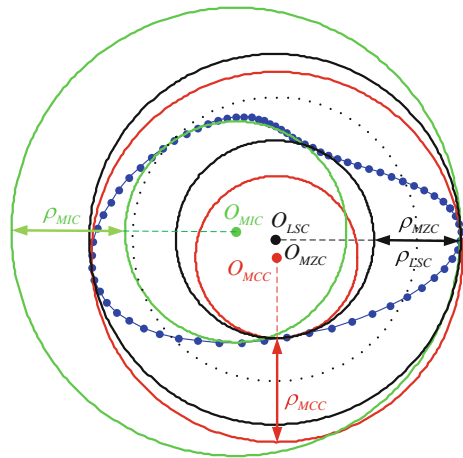
which the distance between the inscribed and circumscribed circles is minimal. In practice, MIC and MCC methods are rarely used if the tool path is essentially distorted and corresponding center points are essentially different (Fig. 4). In the latter case, it is preferable to use either MZC or LSC index. Since there is no considerable difference between MZC or LSC, practicing engineers prefer LSC to evaluate the trajectory circularity.

In the frame of the LSC method, the circle center is obtained as a solution of the following optimization problem

$$\sum_{i=1}^{n} \left(\sqrt{(x_i - x_0)^2 - (y_i - y_0)^2} - r \right)^2 \rightarrow \min_{x_0, y_0, r} \tag{13}$$

where (x_i, y_i) are the measured tool path coordinates, (x_0, y_0) is the circle center and r is its radius. This optimization problem is highly non-linear and cannot be solved analytically. For this reason, a Newton–Raphson method is used.

Using the LSC center point, the circularity is computed in straightforward way in accordance with (12), where

$$r_{\max} = \max(|\mathbf{p}_i - \mathbf{O}_{LS}|, i = \overline{1, n}); \quad r_{\min} = \min(|\mathbf{p}_i - \mathbf{O}_{LS}|, i = \overline{1, n}) \tag{14}$$

Here $\mathbf{p}_i = (x_i, y_i)^T$ is measured tool path points and $\mathbf{O}_{LS} = (x_0, y_0)^T$ is the center of corresponding LS circle.

Assuming that all geometric and elastostatic parameters of the manipulator are known, the resulting tool path for the desired circular trajectory can be computed and evaluated the above presented index. For the set of machining tasks considered in this paper, a benchmark trajectory is a circle of radius 100 mm. For each task and for each path point, there were computed the cutting wrenches $\mathbf{W} = [\mathbf{F}_C^T, \mathbf{M}^T]^T$ which include both the force and the torque components. It is worth mentioning, that the wrench magnitude is constant here, while its direction varies along the path depending on the rotation angle φ. Using this data, the resulting tool path can be computed in the following way

$$\mathbf{p}_i = \mathbf{p}_0 + \mathbf{R}\,\mathbf{r} + \mathbf{J}_\theta^{(p)} \cdot \mathbf{K}_\theta^{-1}\,\mathbf{J}_\theta^{(p)T}\mathbf{R}\,\mathbf{F}_C + \mathbf{J}_\theta^{(p)} \cdot \mathbf{K}_\theta^{-1} \cdot \mathbf{J}_\theta^{(\varphi)T}\mathbf{R}\,\mathbf{M} \tag{15}$$

where $\mathbf{p}_0$ is the center of the reference circle, r is its radius, $\mathbf{r}$ defines the radius vector of the circle for $\varphi = 0$, the matrix $\mathbf{R}(\varphi_i)$ takes into account rotation of the target point by the angle φ_i, the superscripts (p) and (φ) indicate the position and orientation part of the Jacobian matrix presented as $\mathbf{J}_\theta = [\mathbf{J}_\theta^{(p)T}, \mathbf{J}_\theta^{(\varphi)T}]^T$. It should be stressed that dynamics also affects the circularity; however, its influence is much lower comparing to the compliance errors caused by cutting forces.

Based on this approach, it is possible to evaluate robot accuracy with respect to the circularity index for the entire workspace and to determine the region in which machining accuracy is the best (from the circularity point of view).

4 Experimental Result

The technique developed in this paper has been applied to the comparison study of four industrial robots of Kuka family. There were compared with respect to the circular machining task of 100 mm radius that was placed in different workspace points. Some details concerning the examined robots and their principal parameters are given in Table 1. These robots have similar kinematics and provide comparable repeatability/accuracy without loading. However, their payload capacities and workspace size are different. Elastostatic parameters of the examined robots have been identified using dedicated experimental study using methodology developed in our previous works (Klimchik et al. 2013b, 2014b). Corresponding results are presented in Table 2.

In the experimental study, it was used the cutting tool of the radius $R = 5$ mm with three teeth $(z = 3)$. Its machining parameters are the following: $\kappa = 90°$, $\gamma_0 = 7°$, $\lambda_s = 45°$, $f_z = 0.08$ mm/rev, $a_p = 5$ mm, $K_n = 750$ N/mm^2. It is assumed that cutting tool is completely engaged $(\varphi = 180°)$ along the trajectory. This corresponds the cutting force $\mathbf{F}_C = [-440$ N$, -1370$ N$, -635$ N$]$ and cutting torque $\mathbf{M} = [0$ Nm$, 3$ Nm$, 10.5$ Nm$]$. These vectors are rotating while the tool is moving along the circular trajectory.

In our study it was assumed that the reference circular trajectory was located in the plane of joint q_2 and q_3 movements (XOZ if $q_1 = 0$), which is the most critical one for all articulated robots in machining application. For these conditions, the circularity maps have been computed for all examined manipulators. Relevant

Table 1 Principal characteristics of examined robots

Robot	Repeatability (mm)	Workspace volume (m^3)	Working radius (m)	Maximum payload (kg)
KR 100 HA	0.05	46	2.6	100
KR 270	0.06	55	2.7	270
KR 360	0.08	118	3.3	360
KR 500	0.08	68	2.8	500

Table 2 Stiffness parameters of examined robots

Robot	Equivalent joint compliances, μm/N					
	k_1	k_2	k_3	k_4	k_5	k_6
KR 100 HA	1.92	0.34	0.56	3.31	3.83	5.42
KR 270	0.54	0.29	0.42	2.79	3.48	2.07
KR 360	0.86	0.17	0.25	2.17	1.47	2.96
KR 500	0.47	0.14	0.19	0.72	0.95	1.44

results are presented in Figs. 5, 6, 7 and 8. In addition, these figures contain the optimal regions for locating the machining tasks of size 100 × 100, 200 × 200 and 500 × 500 mm. Summary of circularity indices for different machining task is given in Table 3. It should be stressed that in the case of machining task different from the considered one, the results should be scaled according to cutting force magnitude.

As follows from the presented results, robot Kuka KR 500 ensures the best performance for the considered technological task. This advantage is achieved due to less complaint actuators, which obviously affect the robot price. In general, all examined robots ensure circularity level about 3 mm within entire workspace (without compensation). So, if this accuracy is sufficient, the robot can be chosen taking into account the workspace and payload properties. It should be stressed that robot KR 100 cannot be used for machining of hard materials because of the

Fig. 5 Circularity maps for robot KR 100 HA
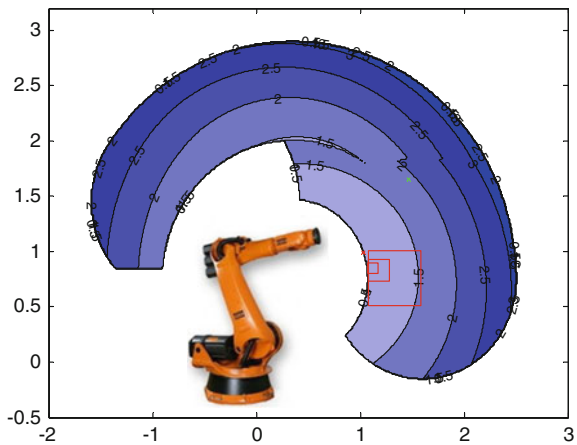
Fig. 6 Circularity maps for robot KR 270
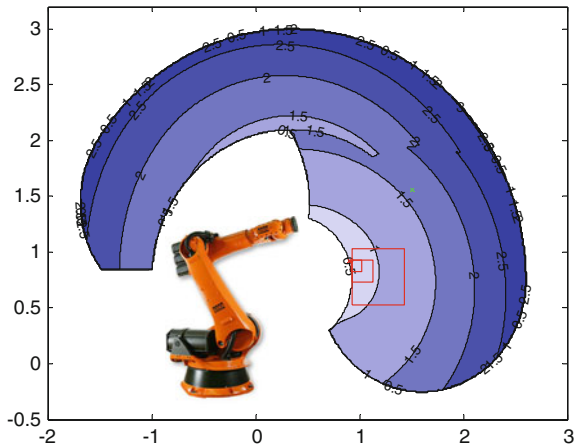
Fig. 7 Circularity maps for robot KR 360
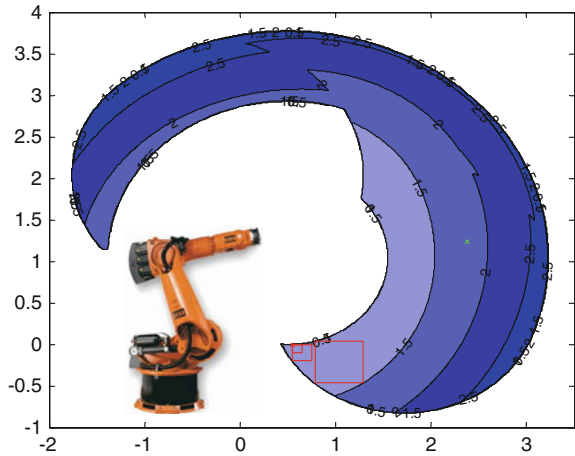
Fig. 8 Circularity maps for robot KR 500
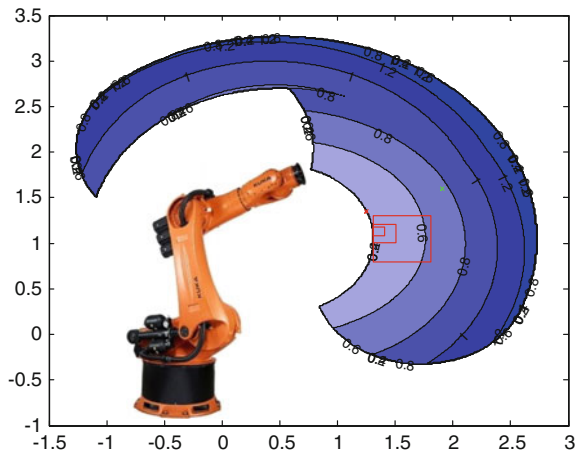
payload limitation. However, because of its lower price comparing to other examined robots, it is competitive for machining if the cutting force magnitude is less than 1 kN and the desired circularity is about 1 mm. On the other hand, robot KR 360 is competitive for large-dimensional tasks only and for milling with forces higher than 2.5 kN. Otherwise, KR 270 is preferable that may ensure better performance within its workspace. The obtained results are also summarized in Table 4 that presents robots suitable for machining with desired accuracy (in terms of the circularity) for different force magnitudes. The latter allows practicing engineers to justify robot selection for given machining task.

Table 3 Circularity indices for examined robots, mm

Robot	Min	Max	Middle
KR 100 HA	1.03	3.36	1.91
KR 270	0.84	3.13	1.64
KR 360	1.02	2.81	1.80
KR 500	0.41	1.42	0.76

Table 4 Suitability of examined robots for different machining tasks (without compensation)

Circularity (mm)	Force magnitude			
	500 N	1000 N	2000 N	3000 N
0.2	KR 500			
0.5	KR 100	KR 270	KR 500	
	KR 270	KR 500		
	KR 360			
	KR 500			
1.0	KR 100	KR 100	KR 270	KR 500
	KR 270	KR 270	KR 360	
	KR 360	KR 360	KR 500	
	KR 500	KR 500		
1.5	KR 100	KR 100	KR 270	KR 500
	KR 270	KR 270	KR 360	
	KR 360	KR 360	KR 500	
	KR 500	KR 500		
2.0	KR 100	KR 100	KR 270	KR 360
	KR 270	KR 270	KR 360	KR 500
	KR 360	KR 360	KR 500	
	KR 500	KR 500		

5   Conclusions

The paper presents an industry oriented technique for evaluation of robot capacity to perform machining operations. It proposes methodology for evaluation of the circularity widely used in industrial practice. Obtained theoretical results have been applied to the experimental study that provides comparison analysis of four industrial robots with respect to their accuracy in machining. In future, this methodology will be applied to a wider set of industrial robots, technological tools and materials.

References

Abele, E., Weigold, M., & Rothenbücher, S. (2007). Modeling and Identification of an industrial robot for machining applications. *CIRP Annals—Manufacturing Technology, 56*(1), 387–390.

Ali, M. H., Khidhir, B. A., Ansari, M. N. M., & Mohamed, B. (2013). FEM to predict the effect of feed rate on surface roughness with cutting force during face milling of titanium alloy. *HBRC Journal, 9*(3), 263–269.

Araujo, A. C., Mello, G. M., & Cardoso, F. G. (2015). Thread milling as a manufacturing process for API threaded connection: Geometrical and cutting force analysis. *Journal of Manufacturing Processes, 18*, 75–83.

Chen, Y., & Dong, F. (2013). Robot machining: Recent development and future research issues. *The International Journal of Advanced Manufacturing Technology, 66*(9–12), 1489–1497.

Chen, Y., Gao, J., Deng, H., Zheng, D., Chen, X., & Kelly, R. (2013). Spatial statistical analysis and compensation of machining errors for complex surfaces. *Precision Engineering, 37*(1), 203–212.

Cheng, P. J., Tsay, J. T., & Lin, S. C. (1997). A study on instantaneous cutting force coefficients in face milling. *International Journal of Machine Tools and Manufacture, 37*(10), 1393–1408.

Devlieg, R. (2010). Expanding the use of robotics in airframe assembly via accurate robot technology. SAE Technical Paper.

Doi, S., & Kato, S. (1955). *Chatter vibration of lathe tools*. ASME.

Dumas, C., Caro, S., Cherif, M., Garnier, S., & Furet, B. (2012). Joint stiffness identification of industrial serial robots. *Robotica, 30*(04), 649–659.

Feng, W. L., Yao, X. D., Azamat, A., & Yang, J. G. (2015). Straightness error compensation for large CNC gantry type milling centers based on B-spline curves modeling. *International Journal of Machine Tools and Manufacture, 88*, 165–174.

Guo, Y., Dong, H., & Ke, Y. (2015). Stiffness-oriented posture optimization in robotic machining applications. *Robotics and Computer-Integrated Manufacturing, 35*, 69–76.

Klimchik, A., Chablat, D., & Pashkevich, A. (2014a). Stiffness modeling for perfect and non-perfect parallel manipulators under internal and external loadings. *Mechanism and Machine Theory, 79*, 1–28.

Klimchik, A., Furet, B., Caro, S., & Pashkevich, A. (2015). Identification of the manipulator stiffness model parameters in industrial environment. *Mechanism and Machine Theory, 90*, 1–22.

Klimchik, A., Pashkevich, A., Chablat, D., & Hovland, G. (2013a). Compliance error compensation technique for parallel robots composed of non-perfect serial chains. *Robotics and Computer-Integrated Manufacturing, 29*(2), 385–393.

Klimchik, A., Pashkevich, A., Wu, Y., Caro, S., & Furet, B. (2012). Design of calibration experiments for identification of manipulator elastostatic parameters. *Applied Mechanics and Materials, 162*, 161–170.

Klimchik, A., Wu, Y., Caro, S., Furet, B., & Pashkevich, A. (2014b). Geometric and elastostatic calibration of robotic manipulator using partial pose measurements. *Advanced Robotics, 28*(21), 1419–1429.

Klimchik, A., Wu, Y., Dumas, C., Caro, S., Furet, B., & Pashkevich, A. (2013b). Identification of geometrical and elastostatic parameters of heavy industrial robots. In *IEEE International Conference on Robotics and Automation (ICRA)*.

Koenigsberger, F., & Sabberwal, A. J. P. (1961). An investigation into the cutting force pulsations during milling operations. *International Journal of Machine Tool Design and Research, 1*(1–2), 15–33.

Kövecses, J., & Angeles, J. (2007). The stiffness matrix in elastically articulated rigid-body systems. *Multibody System Dynamics, 18*(2), 169–184.

Matsuoka, S.-I., Shimizu, K., Yamazaki, N., & Oki, Y. (1999). High-speed end milling of an articulated robot and its characteristics. *Journal of Materials Processing Technology, 95*(1–3), 83–89.

McCutcheon, R., & Pethick, R. (2014). *The new hire: How a new generation of robots is transforming manufacturing*. Price waterhouse Coopers.

Merchant, M. E. (1945). Mechanics of the metal cutting process. I. Orthogonal cutting and a type 2 chip. *Journal of Applied Physics, 16*(5), 267–275.

Nagai, K., & Liu, Z. (2008). A systematic approach to stiffness analysis of parallel mechanisms. In *IEEE International Conference on Robotics and Automation, 2008 (ICRA 2008)*. IEEE.

Nubiola, A., & Bonev, I. A. (2013). Absolute calibration of an ABB IRB 1600 robot using a laser tracker. *Robotics and Computer-Integrated Manufacturing, 29*(1), 236–245.

Pashkevich, A., Klimchik, A., & Chablat, D. (2011). Enhanced stiffness modeling of manipulators with passive joints. *Mechanism and Machine Theory, 46*(5), 662–679.

Shirase, K., & Altintaş, Y. (1996). Cutting force and dimensional surface error generation in peripheral milling with variable pitch helical end mills. *International Journal of Machine Tools and Manufacture, 36*(5), 567–584.

Vosniakos, G.-C., & Matsas, E. (2010). Improving feasibility of robotic milling through robot placement optimisation. *Robotics and Computer-Integrated Manufacturing, 26*(5), 517–525.

Wan, M., Pan, W.-J., Zhang, W.-H., Ma, Y.-C., & Yang, Y. (2014). A unified instantaneous cutting force model for flat end mills with variable geometries. *Journal of Materials Processing Technology, 214*(3), 641–650.

Weck, M., Altintas, Y., & Beer, C. (1994). CAD assisted chatter-free NC tool path generation in milling. *International Journal of Machine Tools and Manufacture, 34*(6), 879–891.

Zargarbashi, S. H. H., Khan, W., & Angeles, J. (2012a). The Jacobian condition number as a dexterity index in 6R machining robots. *Robotics and Computer-Integrated Manufacturing, 28*(6), 694–699.

Zargarbashi, S. H. H., Khan, W., & Angeles, J. (2012b). Posture optimization in robot-assisted machining operations. *Mechanism and Machine Theory, 51*, 74–86.

Adaptive Robust Control and Fuzzy-Based Optimization for Flexible Serial Robot

Fangfang Dong, Jiang Han and Lian Xia

Abstract An adaptive robust control design for flexible joint serial robot is considered. The system contains uncertainty which is assumed to lie in fuzzy set. Since the overall system does not meet the matching condition, a virtual control is implanted and the system is transformed into fuzzy dynamical system. Then an adaptive robust controller can be designed to guarantee the uniform boundedness and uniform ultimate boundedness of the transformed system. No other knowledge of uncertainty is required other than the existence of its bound. The control is deterministic and is not IF-THEN heuristic rules-based. The optimization of control parameters is also considered by solving two quartic equations. The extreme minimum value of the equation is proven to be the optimal solution.

Keywords Fuzzy control · Adaptive robust control · Uncertain system · Flexible joint robot · Uniform ultimate boundedness

1 Introduction

Flexible joint serial robot is widely used in industry applications. It provides less mass, faster motion and stronger payload capacity. All these advantages make it become incomparable. However, a flexible joint serial robot is a highly coupled, nonlinear, complicated system, and often surfers from matched or mismatched uncertainty. Therefore, scholars devoted themselves to find an effective control strategy for achieving high performance in the past decades. Although significant developments have been made, the simplicity, efficiency, and reliability of controls for flexible robots are still unsolved.

Once the exact information of uncertainty is known a priori, some deterministic control schemes can be applied, such as PD control (Piltan and Sulaiman 2011), H_∞

F. Dong · J. Han · L. Xia (✉)
School of Mechanical Engineering, Hefei University of Technology,
193 Tunxi Road, Hefei 230009, Anhui, China
e-mail: xia_nian@126.com

control (Shen and Tamura 1995), sliding mode control (Huang and Chen 2004) and so forth. To deal with the uncertainty, Alonge et al. (2004) and Yeon et al. (2008) adopted robust control and the performance would be guaranteed as long as uncertainty lies in some (typically compact) set. By using adaptive methodology, the unknown parameter of controller can approach to the real value for achieving better tracking trajectory (Chien and Huang 2007; Pradhan and Subudhi 2012). The fuzzy-neural control (Sun et al. 2003) and other feasible strategies are well summarized in, e.g., Kiang et al. (2015) and its bibliographies.

If the uncertainty cannot be clearly identified, which means the known portion cannot be separated from the unknown portion, a fuzzy perspective via the degree of occurrence of certain applications comes into account as an alternative. Fuzzy theory was originally used to describe the object, which is in lack of a sharp boundary (Zadeh 1965), in a human linguistic manner. Thus, fuzzy theory is mostly applied to fuzzy reasoning, estimation, decision-making, etc. The road of fuzzy theory incorporate with system theory (i.e., fuzzy dynamical system) is insufficiently investigated. Past efforts on this domain can be found in Hanss (2005) and Klir and Yuan (1995). Our work endeavors to adopt a new fuzzy dynamical system approach to complement the previous efforts and to broaden the scope.

The main contributions of this work are threefold. First, the uncertainty bound is described by using fuzzy set. This renders the crisp bound description a special case and provides an alternative to probability description. Second, by creatively implanting a fictitious control, the system is transformed into a new representation. An adaptive robust control, which is deterministic and not if-then rules-based, is designed to guarantee the uniform boundedness and uniform ultimate boundedness of the transformed system. Third, a fuzzy-based system performance index based on the fuzzy-based uncertainty bound description is proposed. The optimal choice of the control parameter is formulated as a (semi-infinite) constrained optimization problem. The extreme minimum value to this optimization problem can be obtained by solving two scalar quartic algebraic equations.

2 Fuzzy Dynamical Model of the System

Consider a flexible joint serial robot described as follows

$$\begin{cases} D(q_1, \sigma_1)\,\ddot{q}_1 + C(q_1, \dot{q}_1, \sigma_1)\,\dot{q}_1 + G(q_1, \sigma_1) + K(\sigma_1)(q_1 - q_2) = 0 \\ J(\sigma_2)\,\ddot{q}_2 + F(q_2, \dot{q}_2, \sigma_2)\,\dot{q}_2 - K(\sigma_2)(q_1 - q_2) = u, \end{cases} \tag{1}$$

where $q_1 = [q^2 q^4 \ldots q^{2n}]^T$ is link position vector and $q_2 = [q^1 q^3 \ldots q^{2n-1}]^T$ is joint position vector, here, the superscripts $1, 2, \ldots, 2n$ denote the nth coordinate of the robot. $\sigma_1 \in \Sigma_1 \subset \mathbf{R}^{n_1}$ and $\sigma_2 \in \Sigma_2 \subset \mathbf{R}^{n_2}$ are unfictitious control implanted into the certain parameter vectors with Σ_1 and Σ_2 prescribed and compact. Let $q = \begin{bmatrix} q_1^T & q_2^T \end{bmatrix}^T$ be a $2n$-dimension vector representing the generalized coordinate for

the system. The joint flexibility is denoted by a linear torsional spring whose elasticity coefficient is represented by a diagonal positive matrix $K(\sigma_{1,2})$. $D(q_1, \sigma_1)$ is the link inertia matrix and $J(\sigma_2)$ is a diagonal matrix representing the inertia of actuator; $C(q_1, \dot{q}_1, \sigma_1)\, \dot{q}_1$ and $F(q_2, \dot{q}_2, \sigma_2)\, \dot{q}_2$ represent the Coriolis and centrifugal forces of links and actuators, respectively; $G(q_1, \sigma_1)$ is the gravitation force, and u denotes the input force from the actuators. We note that certain entries of the uncertain parameter vectors σ_1 and σ_2 may represent the same physical parameters because these two subsystems are connected.

Assumption 1 The inertia matrix $D(q_1, \sigma_1)$ is uniformly positive definite (Chen et al. 1998). That is, there exists a constant $\underline{\lambda} > 0$, $\bar{\lambda}_1 > 0$, $\bar{\lambda}_2 \geq 0$, $\bar{\lambda}_3 \geq 0$, such that

$$\bar{\lambda}_1 + \bar{\lambda}_2 \|q_1\| + \bar{\lambda}_3 \|q_1\|^2 \geq \| D(q_1, \sigma_1) \| \geq \underline{\lambda}, \quad \forall q_1 \in \mathbf{R}^n \tag{2}$$

We state this as an assumption other than a fact.

Let us rewrite the first part of (1) as

$$D(q_1, \sigma_1)\,\ddot{q}_1 + C(q_1, \dot{q}_1, \sigma_1)\,\dot{q}_1 + G(q_1, \sigma_1) + K(\sigma_1)q_1 - K(\sigma_1)(q_2 - u_1)$$
$$= K(\sigma_1)u_1 \tag{3}$$

where u_1 is a fictitious control implanted into the system. It is only used to formulate the real control u without changing the dynamics of original system. With u_1 introduced, the system could be divided into link position subsystem and joint position subsystem, which are controlled by u_1 and u, respectively.

Let $x_1 = q_1, x_2 = \dot{q}_1, x_3 = q_2 - u_1, x_4 = \dot{q}_2 - \dot{u}_1$, then $X_1 = [x_1 \quad x_2]^T$, $X_2 = [x_3 \quad x_4]^T$, $X = [X_1 \quad X_2]^T$. Thus, by multiplying $K^{-1}(\sigma_1)$ on both sides of (3) the dynamics of flexible joint robot can be expressed as follows by using new state variable:

$$\hat{D}(x_1, \sigma_1)\,\dot{x}_2 = -\hat{C}(x_1, \dot{x}_1, \sigma_1)\,\dot{x}_1 - \hat{G}(x_1, \sigma_1) - x_1 + x_3 + u_1, \tag{4}$$

$$J(\sigma_2)\,\dot{x}_4 = -J(\sigma_2)\,\ddot{u}_1 - F(x_1, \dot{x}_1, x_3, \dot{x}_3, \sigma_2)\,\dot{x}_3 - F(x_1, \dot{x}_1, x_3, \dot{x}_3, \sigma_2)\,\dot{u}_1$$
$$- K(\sigma_2)x_3 + K(\sigma_2)x_1 - K(\sigma_2)u_1 + u, \tag{5}$$

where $\hat{D} = K^{-1}D, \hat{C} = K^{-1}C, \hat{G} = K^{-1}G$.

Assumption 2

1. Suppose the initial state is $X(t_0) = X_0$, where t_0 is the initial time. For each entry of X_0, namely $X_{0i}, \; i = 1, 2, \ldots, 4n$, there exists a fuzzy set M_{0i} in a universe of discourse $\Theta_i \subset \mathbf{R}$, characterized by a membership function $\mu_{\Theta_i} \colon \Theta_i \to [0, 1]$ (Chen 2011). That is,

$$M_{0i} = \{(X_{0i}, \mu_{\Theta_i}(X_{0i})) | X_{0i} \in \Theta_i\}. \tag{6}$$

2. For each entry of σ_1, σ_2, namely σ_{1i}, σ_{2i}, $i = 1, 2, \ldots, n$, the function $\sigma_{1i}(\cdot)$, $\sigma_{2i}(\cdot)$ are Lebesgue measurable.
3. For each σ_{1i}, σ_{2i}, there exist fuzzy sets N_{1i}, N_{2i} in universe of discourses $\Sigma_{1i} \subset \mathbf{R}$, $\Sigma_{2i} \subset \mathbf{R}$, characterized by membership functions $\mu_{1i} : \Sigma_{1i} \to [0, 1]$, $\mu_{2i} : \Sigma_{2i} \to [0, 1]$. That is,

$$\begin{aligned} N_{1i} &= \{(\sigma_{1i}, \mu_{1i}(\sigma_{1i})) | \sigma_{1i} \in \Sigma_{1i}\}, \\ N_{2i} &= \{(\sigma_{2i}, \mu_{2i}(\sigma_{2i})) | \sigma_{2i} \in \Sigma_{2i}\}. \end{aligned} \tag{7}$$

Remark 1 The Assumption 2 employs fuzzy description on the uncertainties in the overall system. This approach to the system representation earns much more advantage than the probability avenue which often requires a large number of repetitions to acquire the observed data (always limited by nature).

Assumption 3

1. There exist an unknown j-dimension constant vector $\psi_1 \in (0, \infty)^j$ and a known function $\Pi_1 : \mathbf{R}^n \times \mathbf{R}^n \times (0, \infty)^j \to \mathbf{R}_+$ such that for all $x_1 \in \mathbf{R}^n$, $\dot{x}_1 \in \mathbf{R}^n$,

$$\Pi_1(x_1, \dot{x}_1, \psi_1) \geq \|\Phi_1(x_1, \dot{x}_1, \sigma_1)\| \tag{8}$$

$$\begin{aligned} \Phi_1(x_1, \dot{x}_1, \sigma_1) = \frac{1}{2}\dot{\hat{D}}(x_1, \dot{x}_1, \sigma_1)(\dot{x}_1 + S_1 x_1) - \hat{C}(x_1, \dot{x}_1, \sigma_1)\dot{x}_1 \\ - \hat{G}(x_1, \sigma_1) - x_1 + \hat{D}(x_1, \sigma_1)S_1\dot{x}_1, \end{aligned} \tag{9}$$

$$S_1 = diag[S_{1i}]_{n \times n}, \quad S_{1i} > 0, \quad i = 1, 2, \ldots, n. \tag{10}$$

2. For each entry of ψ_1, namely $\psi_{1i}, i = 1, 2, \ldots, j$, there exists a fuzzy set Q_{1i} in a universe of discourse $\Xi_{1i} \subset \mathbf{R}$, characterized by a membership function $\mu_{\Xi_{1i}} : \Xi_{1i} \to [0, 1]$. That is,

$$Q_{1i} = \{(\psi_{1i}, \mu_{\Xi_{1i}}(\psi_{1i})) | \psi_{1i} \in \Xi_{1i}\}. \tag{11}$$

3. The function $\Pi_1(x_1, \dot{x}_1, \cdot) : (0, \infty)^j \to \mathbf{R}_+$ is C^2 (i.e., 2-times continuously differentiable) and concave (i.e., $-\Pi_1(x_1, \dot{x}_1, \cdot)$ is convex), and it is non-decreasing with respect to each ψ_1.
4. The function $\Pi_1(x_1, \dot{x}_1, \cdot)$ and $\frac{\partial \Pi_1}{\partial \psi_1}(\cdot)$ are both continuous.

Assumption 4

1. There exist an unknown s-dimension constant vector $\psi_2 \in (0, \infty)^s$ and a known function $\Pi_2 \colon \mathbf{R}^{2n} \times \mathbf{R}^{2n} \times (0, \infty)^s \to \mathbf{R}_+$ such that for all $X_1 \in \mathbf{R}^{2n}, X_2 \in \mathbf{R}^{2n}$,

$$\Pi_2(X_1, X_2, \psi_2) \geq \|\Phi_2(X, \sigma_1(t), \sigma_2)\|, \tag{12}$$

where $S_2 = diag[S_{2i}]_{n \times n}$, $S_{2i} > 0$, $i = 1, 2, \ldots n$,

$$\Phi_2(X, \sigma_1, \sigma_2) = -J(\sigma_2)\,\ddot{u}_1(X, \sigma_1, \sigma_2) - F(X, \sigma_2)(\dot{x}_3 + \dot{u}_1(X_1, \sigma_1)) - K(\sigma_2)x_3$$

$$+ K(\sigma_2)x_1 - K(\sigma_2)u_1(X_1) + J(\sigma_2)S_2\,\dot{x}_3 + \frac{1}{2}\dot{J}(\sigma_2)(x_4 + S_2 x_3). \tag{13}$$

2. For each entry of ψ_2, namely $\psi_{2i}, i = 1, 2, \ldots, s$, there exists a fuzzy set Q_{2i} in a universe of discourse $\Xi_{2i} \subset \mathbf{R}$, characterized by a membership function $\mu_{\Xi_{2i}} \colon \Xi_{2i} \to [0, 1]$. That is,

$$Q_{2i} = \{(\psi_{2i}, \mu_{\Xi_{2i}}(\psi_{2i})) | \psi_{2i} \in \Xi_{2i}\}. \tag{14}$$

3. The function $\Pi_2(X_1, X_2, \cdot) : (0, \infty)^s \to \mathbf{R}_+$ is C^2 (i.e., 2-times continuously differentiable) and concave (i.e., $-\Pi_2(X_1, X_2, \cdot)$ is convex), and it is nondecreasing with respect to each ψ_2.
4. The function $\Pi_2(X_1, X_2, \cdot)$ and $\frac{\partial \Pi_2}{\partial \psi_2}(\cdot)$ are both continuous.

3 Deterministic Adaptive Robust Control Design

After dividing the system into the cascades of (4) and (5), both of them have "inputs". Therefore, the matching condition is satisfied (Chen and Leitmann 1987). For given constant positive scalars $\gamma_1, \varepsilon_1, k_1$, we construct the fictitious control u_1 for the link position subsystem as follows:

$$u_1(t) = -\gamma_1(\dot{x}_1(t) + S_1 x_1(t))\,\Pi_1{}^2(x_1(t), \dot{x}_1(t), \hat{\psi}_1) - \varepsilon_1(\dot{x}_1(t) + S_1 x_1(t)), \tag{15}$$

where $\hat{\psi}_1$ is the adaptation law with the following form:

$$\dot{\hat{\psi}}_1(t) = k_1^{-1} \frac{\partial \Pi_1^T}{\partial \psi_1}(x_1(t), \dot{x}_1(t), \hat{\psi}_1(t)) \|\dot{x}_1(t) + S_1 x_1(t)\| - \hat{\psi}_1(t). \tag{16}$$

Then, the input torque u can be constructed with given constant positive scalars $\gamma_2, \varepsilon_2, k_2$ as

$$u(t) = -\gamma_2(\dot{x}_3(t) + S_2 x_3(t)) \, \Pi_2^{2}(X(t), \hat{\psi}_2(t)) - \varepsilon_2(\dot{x}_3(t) + S_2 x_3(t)), \qquad (17)$$

where $\hat{\psi}_2$ is the adaptation law with the following form:

$$\dot{\hat{\psi}}_2(t) = k_2^{-1} \frac{\partial \Pi_2^T}{\partial \psi_2}(X(t), \hat{\psi}_2(t)) \|\dot{x}_3(t) + S_2 x_3(t)\| - \hat{\psi}_2(t). \qquad (18)$$

Here, k_1, k_2 determine the "rate" of adaptation, the adaptive parameter will increase quickly so as to effectively compensate the uncertainty as k_1 is large.

Theorem 1 *Let* $\tilde{\psi}_1 = \hat{\psi}_1 - \psi_1, \tilde{\psi}_2 = \hat{\psi}_2 - \psi_2, Z = [X_1^T \quad \tilde{\psi}_1^T \quad X_2^T \quad \tilde{\psi}_2^T]^T$, *subject to Assumptions* 1–4*, the control* (17) *renders the combined mechanical/adaptive systems* (4)*,* (5)*,* (16) *and* (18) *the following performances:*

1. **Uniform boundedness**: Given any constant $r > 0$ with $\|Z_0\| \leq r$, there exists $d(r) > 0$ such that $\|Z(t)\| \leq d(r)$ for all $t \in [t_0, \infty)$.
2. **Uniform ultimate boundedness**: Given any $\bar{d} > \underline{d}$ and any $r > 0$, there exists a finite time $T(\bar{d}, r)$ such that $\|Z_0\| \leq r$ implies $\|Z(t)\| \leq \bar{d}$ for all $t \geq t_0 + T_z(\bar{d}, r)$.

Proof Let $\psi = [\tilde{\psi}_1^T \quad \tilde{\psi}_2^T]^T$, choose the Lyapunov function candidates for the system as following:

$$V(X, \psi) = V_1(X_1, \tilde{\psi}_1) + V_2(X_2, \tilde{\psi}_2), \qquad (19)$$

where

$$V_1(X_1, \tilde{\psi}_1) = V_{X_1} + V_{\psi_1} = \frac{1}{2}(x_2 + S_1 x_1)^T \hat{D}(x_2 + S_1 x_1) + \frac{1}{2}\tilde{\psi}_1^T k_1 \tilde{\psi}_1,$$

$$V_2(X_2, \tilde{\psi}_2) = V_{X_2} + V_{\psi_2} = \frac{1}{2}(x_4 + S_2 x_3)^T J(x_4 + S_2 x_3) + \frac{1}{2}\tilde{\psi}_2^T k_2 \tilde{\psi}_2. \qquad (20)$$

To show $V(X)$ is a legitimate Lyapunov function candidate for any flexible joint robot, we need to prove that $V(X)$ is positive definite and decrescent. Let $\Psi_1 = \begin{bmatrix} S_1^2 & S_1 \\ S_1 & I \end{bmatrix}$, $\bar{S}_1 = \lambda_{\max}(\Psi_1)$, $\underline{S}_1 = \lambda_{\min}(\Psi_1)$, based on Assumption 1,

$$\begin{aligned} V_1 &\geq \frac{1}{2}\underline{\lambda}_0 [x_1 \quad \dot{x}_1] \begin{bmatrix} S_1^2 & S_1 \\ S_1 & I \end{bmatrix} \begin{bmatrix} x_1 \\ \dot{x}_1 \end{bmatrix} + \frac{1}{2}k_1 \|\tilde{\psi}_1\|^2 \\ &= \frac{1}{2}\underline{\lambda}_0 \bar{S}_1 \|X_1\|^2 + \frac{1}{2}k_1 \|\tilde{\psi}_1\|^2 \\ &\geq \lambda_0^1 \|Z_1\|^2. \end{aligned} \qquad (21)$$

where $\lambda_0^1 = \min\{\frac{1}{2}\underline{\lambda}_0 \bar{S}_1, \frac{1}{2}k_1\}$, $Z_1 = \begin{bmatrix} X_1^T & \tilde{\psi}_1^T \end{bmatrix}^T$.

For the upper bound condition (2) of inertia matrix,

$$V_1 \le (\bar{\lambda}_1^1 + \bar{\lambda}_2^1\|x_1\| + \bar{\lambda}_3^1\|x_1\|^2)\|\dot{x}_1 + S_1 x_1\|^2 + \frac{1}{2}k_1\|\tilde{\psi}_1\|^2$$

$$\le (\bar{\lambda}_1^1 + \bar{\lambda}_2^1\|X_1\| + \bar{\lambda}_3^1\|X_1\|^2)\bar{S}_1\|X_1\|^2 + \frac{1}{2}k_1\|\tilde{\psi}_1\|^2$$

$$\le \lambda_1^1\|Z_1\|^2 + \lambda_2^1\|Z_1\|^3 + \lambda_3^1\|Z_1\|^4, \tag{22}$$

where $\lambda_1^1 = \max\{\bar{\lambda}_1^1 \bar{S}_1, \frac{1}{2}k_1\}, \lambda_2^1 = \bar{\lambda}_2^1 \bar{S}_1, \lambda_3^1 = \bar{\lambda}_3^1 \bar{S}_1$. Similarly to V_1, we can prove that V_2 is also positive definite and decrescent, it follows

$$V_2 \ge \lambda_0^2(\|X_2\|^2 + \|\tilde{\psi}_2\|^2) =: \lambda_0^2\|Z_2\|^2, \tag{23}$$

where $\lambda_0^2 = \min\{\frac{1}{2}\underline{\theta}\underline{S}_2, \frac{1}{2}k_2\}, \underline{\theta} = \lambda_{\min}(J), Z_2 = [X_2^T \quad \tilde{\psi}_2^T]^T$.

$$V_2 \le \lambda_1^2(\|X_2\|^2 + \|\tilde{\psi}_2\|^2) =: \lambda_1^2\|Z_2\|^2, \tag{24}$$

where $\lambda_1^2 = \max\{\frac{1}{2}\bar{\theta}\bar{S}_2, \frac{1}{2}k_2\}, \bar{\theta} = \lambda_{\max}(J)$. Therefore, we have

$$\xi_1(\|Z\|) = \lambda_0\|Z\|^2 \le V \le \lambda_1\|Z\|^2 + \lambda_2\|Z\|^3 + \lambda_3\|Z\|^4 = \xi_2(\|Z\|), \tag{25}$$

where $\quad \lambda_0 = \min\{\lambda_0^1, \lambda_0^2\}, \lambda_1 = \max\{\lambda_1^1, \lambda_1^2\}, \lambda_2 = \lambda_2^1, \lambda_3 = \lambda_3^1, Z = [Z_1^T \quad Z_2^T]^T$. That is, V is positive definite and decrescent for all $Z \in \mathbf{R}^{2n+j+s}$.

We analyze derivative of V_1 along the trajectory of the controlled system separately.

$$\dot{V}_{X_1} = (\dot{x}_1 + S_1 x_1)^T \hat{D}(\ddot{x}_1 + S_1 \dot{x}_1) + \frac{1}{2}(\dot{x}_1 + S_1 x_1)^T \dot{\hat{D}}(\dot{x}_1 + S_1 x_1)$$

$$= (\dot{x}_1 + S_1 x_1)^T(\frac{1}{2}\dot{\hat{D}}\dot{x}_1 + \frac{1}{2}\hat{D}S_1 x_1 - \hat{C}\dot{x}_1 - \hat{G} - x_1 + \hat{D}S_1\dot{x} + u_1 + x_3). \tag{26}$$

From (9), it can be seen that

$$\dot{V}_{X_1} = (\dot{x}_1 + S_1 x_1)^T\Phi_1 + (\dot{x}_1 + S_1 x_1)^T u_1 + (\dot{x}_1 + S_1 x_1)^T x_3. \tag{27}$$

According to Assumption 3 and (12), we substitute the control (15), then

$$
\begin{aligned}
\dot{V}_{X_1} &= (\dot{x}_1 + S_1 x_1)^T \Phi_1 + (\dot{x}_1 + S_1 x_1)^T [-\gamma_1 (\dot{x}_1 + S_1 x_1) \Pi_1^2 (X_1, \hat{\psi}_1) - \varepsilon_1 (\dot{x}_1 + S_1 x_1)] \\
&\quad + (\dot{x}_1 + S_1 x_1)^T x_3 \\
&\leq \|\dot{x}_1 + S_1 x_1\| \Pi_1 (X_1, \psi_1) + (\dot{x}_1 + S_1 x_1)^T [-\gamma_1 (\dot{x}_1 + S_1 x_1) \Pi_1 (X_1, \hat{\psi}_1)^2 \\
&\quad - \varepsilon_1 (\dot{x}_1 + S_1 x_1)] + (\dot{x}_1 + S_1 x_1)^T x_3 \\
&= \|\dot{x}_1 + S_1 x_1\| \Pi_1 (X_1, \psi_1) - \gamma_1 \|\dot{x}_1 + S_1 x_1\|^2 \Pi_1^2 (X_1, \hat{\psi}_1) \\
&\quad - \varepsilon_1 \|\dot{x}_1 + S_1 x_1\|^2 + (\dot{x}_1 + S_1 x_1)^T x_3.
\end{aligned}
\tag{28}
$$

Based on (16), the derivative of V_{ψ_1} follows

$$
\begin{aligned}
\dot{V}_{\psi_1} &= (\hat{\psi}_1 - \psi_1)^T k_1 \dot{\hat{\psi}}_1 \\
&= (\hat{\psi}_1 - \psi_1)^T \frac{\partial \Pi_1^T}{\partial \psi_1} (X_1, \hat{\psi}_1) \|\dot{x}_1 + S_1 x_1\| - (\hat{\psi}_1 - \psi_1)^T k_1 \dot{\hat{\psi}}_1.
\end{aligned}
\tag{29}
$$

From Assumption 3, we know that $-\Pi_1(X_1, \cdot)$ is convex for all $X_1 \in \mathbf{R}^{2n}$, this leads to

$$
\frac{\partial \Pi_1}{\partial \psi_1} (X_1, \hat{\psi}_1)(\hat{\psi}_1 - \psi_1) \leq \Pi_1(X_1, \hat{\psi}_1) - \Pi_1(X_1, \psi_1),
\tag{30}
$$

therefore, we have

$$
\dot{V}_{\psi_1} \leq (\Pi_1(X_1, \hat{\psi}_1) - \Pi_1(X_1, \psi_1)) \|\dot{x}_1 + S_1 x_1\| - (\hat{\psi}_1 - \psi_1) k_1 \dot{\hat{\psi}}_1
\tag{31}
$$

By combining (28) and (31), we have

$$
\begin{aligned}
\dot{V}_1 &= \|\dot{x}_1 + S_1 x_1\| \Pi_1(X_1, \psi_1) - \gamma_1 \|\dot{x}_1 + S_1 x_1\|^2 \Pi_1^2(X_1, \hat{\psi}_1) - \varepsilon_1 \|\dot{x}_1 + S_1 x_1\|^2 \\
&\quad + (\dot{x}_1 + S_1 x_1)^T x_3 + (\Pi_1(X_1, \hat{\psi}_1) - \Pi_1(X_1, \psi_1)) \|\dot{x}_1 + S_1 x_1\| - (\hat{\psi}_1 - \psi_1) k_1 \dot{\hat{\psi}}_1 \\
&= \|\dot{x}_1 + S_1 x_1\| \Pi_1(X_1, \hat{\psi}_1) - \gamma_1 \|\dot{x}_1 + S_1 x_1\|^2 \Pi_1^2(X_1, \hat{\psi}_1) - \varepsilon_1 \|\dot{x}_1 + S_1 x_1\|^2 \\
&\quad + (\dot{x}_1 + S_1 x_1)^T x_3 - (\hat{\psi}_1 - \psi_1) k_1 \dot{\hat{\psi}}_1 \\
&\leq \frac{1}{4\gamma_1} - \varepsilon_1 \|\dot{x}_1 + S_1 x_1\|^2 + (\dot{x}_1 + S_1 x_1)^T x_3 - (\hat{\psi}_1 - \psi_1) k_1 \dot{\hat{\psi}}_1.
\end{aligned}
\tag{32}
$$

According to the inequality $ab \leq (a^2 + b^2)/2$, $a, b \in \mathbf{R}$, we have

$$
(\dot{x}_1 + S_1 x_1)^T x_3 \leq \frac{1}{2} \omega_1 \|\dot{x}_1 + S_1 x_1\|^2 + \frac{1}{2} \omega_1^{-1} \|x_3\|^2,
\tag{33}
$$

where $\omega_1 > 0$ is a constant. Therefore,

$$\dot{V}_1 \le \frac{1}{4\gamma_1} - \varepsilon_1 \|\dot{x}_1 + S_1 x_1\|^2 - (\hat{\psi}_1 - \psi_1) k_1 \hat{\psi}_1 + \left(\frac{1}{2} \omega_1 \|\dot{x}_1 + S_1 x_1\|^2 + \frac{1}{2} \omega_1^{-1} \|x_3\|^2 \right)$$

$$\le \frac{1}{4\gamma_1} - (\varepsilon_1 - \frac{1}{2}\omega_1) \|\dot{x}_1 + S_1 x_1\|^2 + \frac{1}{2}\omega_1^{-1} \|x_3\|^2 - (\hat{\psi}_1 - \psi_1) k_1 (\hat{\psi}_1 - \psi_1 + \psi_1)$$

$$\le -\underline{\lambda}_1 \|Z_1\|^2 + \frac{1}{4\gamma_1} + \frac{1}{2} k_1 \|\psi_1\|^2 + \frac{1}{2}\omega_1^{-1}\|x_3\|^2, \tag{34}$$

where $\underline{\lambda}_1 = \min\{(\varepsilon_1 - \frac{1}{2}\omega_1) \underline{S}_1, \frac{1}{2} k_1\}$. Next, the derivative of V_2 is given by

$$\dot{V}_2 = \dot{V}_{X_2} + \dot{V}_{\psi_2}. \tag{35}$$

According to (5) and (13),

$$\dot{V}_{X_2} = (\dot{x}_3 + S_2 x_3)^T J (\ddot{x}_3 + S_2 \dot{x}_3) + \frac{1}{2}(\dot{x}_3 + S_2 x_3)^T \dot{J}(\dot{x}_3 + S_2 x_3)$$

$$= (\dot{x}_3 + S_2 x_3)^T (-J\ddot{u}_1 - Kx_3 - F\dot{x}_3 - F\dot{u}_1 + Kx_1 - Ku_1$$

$$\quad + JS_2 \dot{x}_3 + \frac{1}{2}\dot{J}\dot{x}_3 + \frac{1}{2}\dot{J}S_2 x_3 + u)$$

$$= (\dot{x}_3 + S_2 x_3)^T \Phi_2 + (\dot{x}_3 + S_2 x_3)^T u. \tag{36}$$

Concerning $\dot{V}_{\psi_2}$, it follows from (18)

$$\dot{V}_{\psi_2} = (\hat{\psi}_2 - \psi_2)^T k_2 \dot{\hat{\psi}}_2$$

$$= (\hat{\psi}_2 - \psi_2)^T \frac{\partial \Pi_2^T}{\partial \psi_2} (X, \hat{\psi}_2) \|\dot{x}_3 + S_1 x_3\| - (\hat{\psi}_2 - \psi_2)^T k_2 \hat{\psi}_2. \tag{37}$$

Similarly to the process of $\dot{V}_1$, we have

$$\dot{V}_2 \le \frac{1}{4\gamma_2} - \varepsilon_2 \|\dot{x}_3 + S_2 x_3\| - (\hat{\psi}_2 - \psi_2)^T k_2 (\hat{\psi}_2 - \psi_2 + \psi_2)$$

$$\le \frac{1}{4\gamma_2} - \varepsilon_2 \lambda_{\min}(\Psi_2) \|X_2\|^2 - k_2 \|\tilde{\psi}_2\|^2 + \frac{1}{2} k_2 \left(\|\tilde{\psi}_2\|^2 + \|\psi_2\|^2 \right)$$

$$= \frac{1}{4\gamma_2} - \varepsilon_2 \lambda_{\min}(\Psi_2) \|X_2\|^2 - \frac{1}{2} k_2 \|\tilde{\psi}_2\|^2 + \frac{1}{2} k_2 \|\psi_2\|^2. \tag{38}$$

Therefore, with $\|x_3\|^2 \le \|X_2\|^2$, the total derivative of V is given by

$$\dot{V} = -\underline{\lambda}_1 \|Z_1\|^2 + \frac{1}{4\gamma_1} + \frac{1}{2}k_1\|\psi_1\|^2 + \frac{1}{2}\omega_1^{-1}\|x_3\|^2 + \frac{1}{4\gamma_2}$$

$$- (\varepsilon_2\underline{S}_2 - \frac{1}{2}\omega_1^{-1})\|X_2\|^2 - \frac{1}{2}k_2\|\tilde{\psi}_2\|^2 + \frac{1}{2}k_2\|\psi_2\|^2$$

$$\le - \underline{\lambda}_1\|Z_1\|^2 - \underline{\lambda}_2\|Z_2\|^2 + \frac{1}{4\gamma_1} + \frac{1}{4\gamma_2} + \frac{1}{2}k_1\|\psi_1\|^2 + \frac{1}{2}k_2\|\psi_2\|^2$$

$$=: -\underline{\lambda}\|Z\|^2 + h, \tag{39}$$

here $\quad h = \frac{1}{4\gamma_1} + \frac{1}{4\gamma_2} + \frac{1}{2}k_1\|\psi_1\|^2 + \frac{1}{2}k_2\|\psi_2\|^2 \quad \underline{\lambda}_2 = \min\{\varepsilon_2\underline{S}_2 - \frac{1}{2}\omega_1^{-1}, \frac{1}{2}k_2\}$, $\underline{\lambda} = \min\{\underline{\lambda}_1, \underline{\lambda}_2\}$. If we choose suitable $\varepsilon_1, \varepsilon_2$ such that $\varepsilon_1 - \omega_1/2 > 0$, and $\varepsilon_2\underline{S}_2 - \omega_1^{-1}/2 > 0$, then $\dot{V} < 0$, for all $\|Z\| \ge \sqrt{h/\underline{\lambda}}$. The uniform boundedness performance follows Corless and Leitmann (1981), that is given any $r > 0$ with $\|Z(t_0)\| \le r$, where t_0 is the initial time, there exists a $d(r)$ given by

$$d(r) = \begin{cases} r\sqrt{\dfrac{\lambda_1 + \lambda_2 r + \lambda_3 r^2}{\lambda_0}} & \text{if } r > R, \\[3mm] R\sqrt{\dfrac{\lambda_1 + \lambda_2 R + \lambda_3 R^2}{\lambda_0}} & \text{if } r \le R, \end{cases}$$
$$R = \sqrt{h/\underline{\lambda}}, \tag{40}$$

such that $\|Z(t)\| \le d(r)$ for all $t \ge t_0$. Uniform ultimate boundedness also follows. That is, given any $\bar{d}$ with

$$\bar{d} > R\sqrt{\frac{\lambda_1 + \lambda_2 R + \lambda_3 R^2}{\lambda_0}}, \tag{41}$$

we have $\|Z(t)\| \le \bar{d}$ for all $t \ge t_0 + T(\bar{d}, r)$, with

$$T(\bar{d}, r) = \begin{cases} 0 & \text{if } r \le \bar{R}, \\[2mm] \dfrac{\lambda_1 r^2 + \lambda_2 r^3 + \lambda_3 r^4 - \lambda_0 \bar{R}^2}{\lambda_{\min}(A)\,\bar{R}^2 - \alpha} & \text{if } r > \bar{R}, \end{cases}$$
$$\bar{R} = \xi_2^{-1}(\lambda_0\,\bar{d}^2) \tag{42}$$

<div align="right">Q.E.D.</div>

Remark 2 The resulting performance of the controlled mechanical system is deterministic. Given the control parameters and the adaptive scheme, both uniform boundedness and uniform ultimate boundedness are guaranteed. This performance assures the bottom line. In addition, since there is fuzzy information in the uncertain parameter, it allows us to further address performance considerations. This will be explored in the later section.

4 Optimal Gain Design

Section 3 shows that the overall system performance can be guaranteed by a deterministic control scheme. By the analysis, the size of the uniform ultimate boundedness region decreases as $\gamma_{1,2}$ increase. This rather strong performance is accomplished by a (possibly) large control effort, which is reflected by $\gamma_{1,2}$. From the practical design point of view, the designer may be interested in seeking an optimal choice of $\gamma_{1,2}$ for a compromise among various conflicting criteria. This is associated with the minimization of a performance index.

Lemma 1 *For any* $Z(t) \in \mathbf{R}^{4n+j+s}, t \geq t_0$, *there exists a known continuous function* $g(\cdot)$ *such that*

$$g(\|Z\|) = \rho_1 \|Z\|^4 + \rho_2 \|Z\|^2 \geq \xi_2(\|Z\|) \geq V. \tag{43}$$

Based on the uniform boundedness and uniform ultimate boundedness performance proven above, we know that $(\xi_1^{-1} \circ g)(Z_0)$ serves as the upper bound of the transient state performance and $(\xi_1^{-1} \circ g \circ \xi_3^{-1})(h)$ serves as the upper bound of the steady state performance. Let

$$\eta^2(t_0) := (\xi_1^{-1} \circ g)(Z_0) = \frac{1}{\lambda_0}\left(\rho_1 \|Z_0\|^4 + \rho_2 \|Z_0\|^2\right), \tag{44}$$

$$\begin{aligned}
\eta_\infty^2(\gamma_1, \gamma_2, \psi_1, \psi_2) &= (\xi_1^{-1} \circ g \circ \xi_3^{-1})(h) \\
&= \frac{1}{\gamma_1}\frac{\rho_2}{4\lambda_0\underline{\lambda}} + \frac{1}{\gamma_1}\frac{k_1}{\lambda_0}\|\psi_1\|^2 + \frac{1}{\gamma_1}\frac{k_2}{\lambda_0}\|\psi_2\|^2 + \frac{1}{\gamma_2}\frac{\rho_2}{4\lambda_0\underline{\lambda}} \\
&\quad + \frac{1}{\gamma_2}\frac{k_1}{\lambda_0}\|\psi_1\|^2 + \frac{1}{\gamma_2}\frac{k_2}{\lambda_0}\|\psi_2\|^2 + \frac{1}{\gamma_1^2}\frac{\rho_1}{16\underline{\lambda}^2\lambda_0} \\
&\quad + \frac{1}{\gamma_2^2}\frac{\rho_1}{16\underline{\lambda}^2\lambda_0} + \frac{1}{\gamma_1\gamma_2}\frac{\rho_1}{8\underline{\lambda}^2\lambda_0} + \frac{\rho_1}{4\underline{\lambda}^2 da_0}k_1^2\|\psi_1\|^4 \\
&\quad + \frac{\rho_2}{2\underline{\lambda}\lambda_0}k_1\|\psi_1\|^2 + \frac{\rho_1}{4\underline{\lambda}^2\lambda_0}k_2^2\|\psi_2\|^4 + \frac{\rho_2}{2\underline{\lambda}\lambda_0}k_2\|\psi_2\|^2 \\
&\quad + \frac{k_1 k_2}{2\underline{\lambda}^2\lambda_0}\|\psi_1\|^2\|\psi_2\|^2. \tag{45}
\end{aligned}$$

Since there is no exact knowledge of the uncertainty, it is only realistic to refer to $\eta(t_0)$ and $\eta_\infty(\gamma_1, \gamma_2, \psi_1, \psi_2)$ while analyzing the system performance.

Applying *D*-operation to $\eta(t_0)$ and $\eta_\infty(\gamma_1, \gamma_2, \psi_1, \psi_2)$, we now propose the following performance index: For any Z_0, let

$$J(\gamma_1, \gamma_2, t_0) = J_1(\gamma_1, \gamma_2, t_0) + \alpha J_2(\gamma_1, \gamma_2, \psi_1, \psi_2) + \beta J_3(\gamma_1, \gamma_2) \tag{46}$$

where

$$J_1(\gamma_1, \gamma_2, t_0) = D[\eta^2(t_0)], \tag{47}$$

$$J_2(\psi_1, \psi_2, \gamma_1, \gamma_2) = D\big[\eta_\infty^2(\gamma_1, \gamma_2, \psi_1, \psi_2)\big], \tag{48}$$

$$J_3(\gamma_1, \gamma_2) = \gamma_1^2 + \gamma_2^2, \tag{49}$$

Because $\eta(t_0)$ does not depend on the design parameters $\gamma_{1,2}$, thus J_1 is a constant

$$J_1(\gamma_1, \gamma_2, t_0) = D[\eta^2(t_0)] =: \kappa_0. \tag{50}$$

Next, applying D-operation to J_2 we obtain

$$J_2(\psi_1, \psi_2, \gamma_1, \gamma_2) = \kappa_1 + \frac{\kappa_2}{\gamma_1} + \frac{\kappa_2}{\gamma_2} + \frac{\kappa_3}{\gamma_1^2} + \frac{\kappa_3}{\gamma_2^2} + \frac{\kappa_4}{\gamma_1\gamma_2}, \tag{51}$$

here,

$$
\begin{aligned}
\kappa_1 &= D\left[\frac{\rho_1}{4\,\underline{\lambda}^2\,\lambda_0}k_1^2\|\psi_1\|^4 + \frac{\rho_2}{2\underline{\lambda}\lambda_0}k_1\|\psi_1\|^2\right] \\
&\quad + D\left[\frac{\rho_1}{4\,\underline{\lambda}^2\,\lambda_0}k_2^2\|\psi_2\|^4 + \frac{\rho_2}{2\underline{\lambda}\lambda_0}k_2\|\psi_2\|^2\right] + D\left[\frac{k_1k_2}{2\,\underline{\lambda}^2\,\lambda_0}\|\psi_1\|^2\|\psi_2\|^2\right], \\
\kappa_2 &= \frac{\rho_2}{4\lambda_0\underline{\lambda}} + \frac{k_1}{\lambda_0}D\big[\|\psi_1\|^2\big] + \frac{k_2}{\lambda_0}D\big[\|\psi_2\|^2\big], \\
\kappa_3 &= \frac{\rho_1}{16\,\underline{\lambda}^2\,\lambda_0}, \kappa_4 = \frac{\rho_1}{8\,\underline{\lambda}^2\,\lambda_0},
\end{aligned}
\tag{52}
$$

with (49), (50) and (51) into (46), we have

$$J(\gamma_1, \gamma_2, t_0) = \kappa_0 + \alpha\left(\kappa_1 + \frac{\kappa_2}{\gamma_1} + \frac{\kappa_2}{\gamma_2} + \frac{\kappa_3}{\gamma_1^2} + \frac{\kappa_3}{\gamma_2^2} + \frac{\kappa_4}{\gamma_1\gamma_2}\right) + \beta(\gamma_1^2 + \gamma_2^2). \tag{53}$$

The optimal design problem is then equivalent to the following constrained optimization problem: For any t_0

$$\min_{\gamma_1, \gamma_2} J(\gamma_1, \gamma_2, t_0) \text{ subject to } \gamma_1, \gamma_2 > 0. \tag{54}$$

For any t_0, taking the first derivative of J with respect to γ_1, γ_2,

$$\frac{\partial J}{\partial \gamma_1} = -\frac{\alpha\kappa_2}{\gamma_1^2} - \frac{\alpha\kappa_3}{\gamma_1^3} - \frac{\alpha\kappa_4}{\gamma_1^2\gamma_2} + 2\beta\gamma_1,$$

$$\frac{\partial J}{\partial \gamma_2} = -\frac{\alpha\kappa_2}{\gamma_2^2} - \frac{\alpha\kappa_3}{\gamma_2^3} - \frac{\alpha\kappa_4}{\gamma_1\gamma_2^2} + 2\beta\gamma_2. \tag{55}$$

That

$$\frac{\partial J}{\partial \gamma_1} = 0, \quad \frac{\partial J}{\partial \gamma_2} = 0 \tag{56}$$

leads to

$$\begin{cases} 2\beta\gamma_1^4 = \alpha\kappa_3 + \alpha(\kappa_2 + \frac{\kappa_4}{\gamma_2})\gamma_1 \\ 2\beta\gamma_2^4 = \alpha\kappa_3 + \alpha(\kappa_2 + \frac{\kappa_4}{\gamma_1})\gamma_2 \end{cases} \tag{57}$$

with $\gamma_1, \gamma_2 > 0$.

The solution can be obtained numerically. Suppose the solution are γ_1^*, γ_2^*, taking the second derivative of J, one gets

$$A = \frac{\partial^2 J}{\partial \gamma_1^2}\Big|_{\gamma_1^*,\gamma_2^*} = \frac{2\alpha\kappa_2}{\gamma_1^{*3}} + \frac{3\alpha\kappa_3}{\gamma_1^{*4}} + \frac{2\alpha\kappa_4}{\gamma_1^{*3}\gamma_2^*} + 2\beta =: h_1(\gamma_1^*) + \frac{2\alpha\kappa_4}{\gamma_1^{*3}\gamma_2^*},$$

$$B = \frac{\partial}{\partial \gamma_2}\left(\frac{\partial J}{\partial \gamma_1}\right)\Big|_{\gamma_1^*,\gamma_2^*} = \frac{\partial}{\partial \gamma_1}\left(\frac{\partial J}{\partial \gamma_2}\right)\Big|_{\gamma_1^*,\gamma_2^*} = \frac{\alpha\kappa_4}{\gamma_1^{*2}\gamma_2^{*2}},$$

$$C = \frac{\partial^2 J}{\partial \gamma_2^2}\Big|_{\gamma_1^*,\gamma_2^*} = \frac{2\alpha\kappa_2}{\gamma_2^{*3}} + \frac{3\alpha\kappa_3}{\gamma_2^{*4}} + \frac{2\alpha\kappa_4}{\gamma_1^*\gamma_2^{*3}} + 2\beta =: h_2(\gamma_2^*) + \frac{2\alpha\kappa_4}{\gamma_1^*\gamma_2^{*3}}, \tag{58}$$

here, $h_1(\gamma_1^*) = \frac{2\alpha\kappa_2}{\gamma_1^{*3}} + \frac{3\alpha\kappa_3}{\gamma_1^{*4}} + 2\beta > 0$, $h_2(\gamma_2^*) = \frac{2\alpha\kappa_2}{\gamma_2^{*3}} + \frac{3\alpha\kappa_3}{\gamma_2^{*4}} + 2\beta > 0$.

$$B^2 - AC = \left(\frac{\alpha\kappa_4}{\gamma_1^{*2}\gamma_2^{*2}}\right)^2 - h_1(\gamma_1^*)\frac{2\alpha\kappa_4}{\gamma_1^*\gamma_2^{*3}} - h_2(\gamma_2^*)\frac{2\alpha\kappa_4}{\gamma_1^{*3}\gamma_2^*} - h_1(\gamma_1^*)h_2(\gamma_2^*) - \frac{4\alpha^2\kappa_4^2}{ma_1^{*4}\gamma_2^{*4}}$$

$$=: -\frac{3\alpha^2\kappa_4^2}{\gamma_1^{*4}\gamma_2^{*4}} - h_3(\gamma_1^*,\gamma_2^*), \tag{59}$$

where $h_3(\gamma_1^*,\gamma_2^*) = h_1(\gamma_1^*)\frac{2\alpha\kappa_4}{\gamma_1^*\gamma_2^{*3}} + h_2(\gamma_2^*)\frac{2\alpha\kappa_4}{\gamma_1^{*3}\gamma_2^*} + h_1(\gamma_1^*)h_2(\gamma_2^*) > 0$.

Therefore, $A > 0$ and $B^2 - AC < 0$, thus, γ_1^*, γ_2^* solve the constrained minimization problem.

5 Conclusion

With uncertainty and states of the system being fuzzily described, an adaptive robust control for the multi-link flexible serial robot is proposed. The control design procedure is only based on the existence of bound of the uncertainty, although the bound is unknown. The parameters, which are related to the uncertainty, in the adaptive laws, are assumed to lie in fuzzy sets. The system performance is assured by the given control which is deterministic and not if-then rules-based. The minimization of the control efforts is considered. The optimal control gains can be obtained by solving a constrained performance index. This performance index is a combination of both deterministic performance (including uniform boundedness and uniformly ultimate boundedness) and fuzzy performance (considering the average control effort).

References

Alonge, F., D'Ippolito, F., & Raimondi, F. M. (2004). Globally convergent adaptive and robust control of robotic manipulators for trajectory tracking. *Control Engineering Practice, 12*(9), 1091–1100.

Chen, Y. H. (2011). A new approach to the control design of fuzzy dynamical systems. *Journal of Dynamic Systems, Measurement, and Control, 133*(6), 061019.

Chen, Y. H., & Leitmann, G. (1987). Robustness of uncertain systems in the absence of matching assumptions. *International Journal of Control, 45*(5), 1527–1542.

Chen, Y. H., Leitmann, G., & Chen, J. S. (1998). Robust control for rigid serial manipulators: A general setting. In *American Control Conference* (Vol. 2, pp. 912–916). IEEE.

Chien, M. C., & Huang, A. C. (2007). Adaptive control for flexible-joint electrically driven robot with time-varying uncertainties. *IEEE Transactions on Industrial Electronics, 54*(2), 1032–1038.

Corless, M. J., & Leitmann, G. (1981). Continuous state feedback guaranteeing uniform ultimate boundedness for uncertain dynamic systems. *IEEE Transactions on Automatic Control, 26*(5), 1139.

Hanss, M. (2005). *Applied fuzzy arithmetic: An introduction with engineering applications*. Berlin Heidelberg: Springer.

Huang, A. C., & Chen, Y. C. (2004). Adaptive sliding control for single-link flexible joint robot with mismatched uncertainties. *Control Systems Technology, IEEE Transactions on, 12*(5), 770–775.

Kiang, C. T., Spowage, A., & Yoong, C. K. (2015). Review of control and sensor system of flexible manipulator. *Journal of Intelligent and Robotic Systems, 77*(1), 187–213.

Klir, G., & Yuan, B. (1995). *Fuzzy sets and fuzzy logic* (Vol. 4). New Jersey: Prentice Hall.

Piltan, F., Sulaiman, N., et al. (2011). Position control of robot manipulator: Design a Novel SISO adaptive sliding mode fuzzy PD fuzzy sliding mode control. *International Journal of Artificial intelligence and Expert System, 2*(5), 208–228.

Pradhan, S. K., & Subudhi, B. (2012). Real-time adaptive control of a flexible manipulator using reinforcement learning. *IEEE Transactions on Automation Science and Engineering, 9*(2), 237–249.

Sun, F.-C., Sun, Z.-Q., Zhang, L.-B., & Xu, F.-J. (2003). Dynamic neuro-fuzzy adaptive control for flexible-link manipulators. *International Journal of Fuzzy Systems, 5*(2), 89–97.

Shen, T., & Tamura, K. (1995). Robust H_∞ control of uncertain nonlinear system via state feedback. *IEEE Transactions on Automatic Control, 40*(4), 766–768.

Yeon, J. S., & Park, J. H. (2008). Practical robust control for flexible joint robot manipulators. In *IEEE International Conference on Robotics and Automation (ICRA 2008)*, pp. 3377–3382.

Zadeh, L. A. (1965). Fuzzy sets. *Information and control, 8*(3), 338–353.

Wired Autonomous Vacuum Cleaner

Emin Faruk Kececi and Fatih Kendir

Abstract In this study, with the help of sensors and encoders, the modification of an existing vacuum cleaner into an autonomous robot is explained. The autonomous motion method proposed in this study is tested with a simulation program in MATLAB as well as on an existing vacuum cleaner. In an environment with unknown obstacle numbers, size, position and orientation, the distance to the robot is measured with the sensors and motion direction and the distance traveled is measured with the encoders. Inside the wire dispensing mechanism, another encoder is used to measure the maximum allowed distance to travel. The robot keeps the motion information on a digital map and completes the cleaning job autonomously.

Keywords Cleaning robot · Wired mobile · Random navigation

1 Introduction

Wireless vacuum cleaners are widely used commercial products. However; the biggest disadvantage of these products is that they carry the battery which power them. The wireless vacuum cleaners can move freely in a room and plan their paths. During the travel inside a room these robots use different mapping algorithms, such as random walk or wall following (Gantz 2012; Milford 2008).

Even though they can move freely, they also have the drawback of power. In some designs the robot has the self-charging capacity, when the battery finishes the robot goes back to its charging station. However the vacuum process requires significant power which cannot be supplied with a battery for a long time.

In order to solve the power problem we have designed and built a vacuum cleaner robot with a power cable which is long enough to cover 40 m^2 of surface. This design has the advantage of power, but has the disadvantage of path planning

E.F. Kececi (✉) · F. Kendir
Department of Mechanical Engineering, Istanbul Technical University, Istanbul, Turkey
e-mail: kececi@itu.edu.tr

167

and traveling inside a room to be cleaned. After the vacuum cleaner is plugged on the wall the robot has a cable mechanism to keep track of the cable length so that the cable does not get tangled and the robot also has a memory algorithm to keep track of the path it has followed to come to the current position with respect to the plug so that after it cleans the surface it can go back to its original position and start to clean other parts of the room.

Since every user will be using the robot in a different environment, where obstacle size, orientation and position is unknown, using a landmark system (Darwish 2012) becomes too difficult and costly to install. In order to overcome this difficulty the robot is designed to have a sensory system to detect the obstacles and a memory system to keep track of the cleaned area.

Robot position and heading are measured and calculated with the help of encoders on the wheels of the robot (Fernandes and Claraco 2013). The robot power cable length is also calculated so that no tangling occurs. The simulations of the autonomous motion is achieved by using MATLAB (MATLAB Control OpenGL Rendering Documentation), where the limitation of cable length and coverage area can be proven.

2 Motion Planning

In order to keep track of the motion of the robot and the cable dispensing mechanism for the mapping algorithm, the motion of the vacuum robot is modeled in the Cartesian coordinate system where the center of the robot is considered as the target point and the end of the wire is considered as the center of the coordinate system. Figure 1 shows in a time interval how the position of the robot changes and how it can be represented in Eqs. 1 through 3.

Fig. 1 Position of the robot in Cartesian coordinates
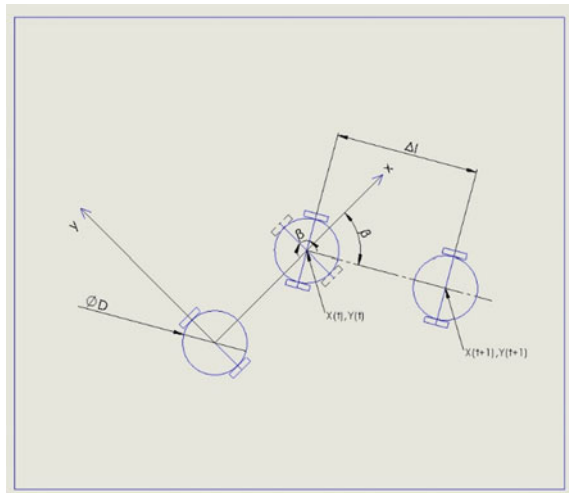
$$x(t + 1) = x(t) + \Delta l \cos(\beta). \tag{1}$$

$$y(t + 1) = y(t) + \Delta l \sin(\beta). \tag{2}$$

$$\beta(t+1) = \beta(t) + \Delta\beta. \tag{3}$$

where Δl is the distance traveled by the robot in unit time and β is the rotation angle about the robot's vertical axes.

The vacuum robot is driven with 2 wheels where the rotational speed of these right and left wheels are respectively $\omega1$ and $\omega2$. By using Fig. 1, β rotation angle can be found as

$$\frac{D}{2}\beta = \left(\frac{\omega_2 - \omega_1}{2}\right)\frac{d}{2}\Delta t. \tag{4}$$

$$v = \left(\frac{\omega_2 + \omega_1}{2}\right)\frac{d}{2}. \tag{5}$$

$$\frac{\Delta x}{\Delta t} = v \cos(\beta). \tag{6}$$

$$\frac{\Delta y}{\Delta t} = v \sin(\beta) \tag{7}$$

where D is the diameter of the robot body, d is the wheel diameter, $\omega1$ and $\omega2$ are right and left wheel rotational speed and it is assumed that apart from the rotational motion these two parameters are same during linear motion and Δt is the time where the motor rotation command signal is applied.

The vacuum cleaner robot is designed to move at 70–80 mm/s speed to be able to pick up the possible dirt efficiently. It is assumed that the robot moves at linear intervals with Δt time intervals. Since the speed of the robot is very slow, the equation of motion is represented with a kinematic model and calculated as

$$\begin{bmatrix} x \\ y \\ \beta \end{bmatrix} = \begin{bmatrix} x(t) + \Delta l \cos\beta \\ y(t) + \Delta l \sin\beta \\ \beta(t) + \Delta\beta \end{bmatrix}. \tag{8}$$

$$\Delta l = \left(\frac{\omega_2 + \omega_1}{2}\right)\frac{d}{2}\Delta t. \tag{9}$$

3 Robot Hardware

The robot has 2 driving wheels which are powered by two 12 VDC motors, 1 DC motor to control the length of the cable and encoders attached to all 3 motors. With a transformer the city supply drops down to 12 VDC level. Motor drivers are also

Fig. 2 Positions of the
sensors and electronic
hardware in the robot
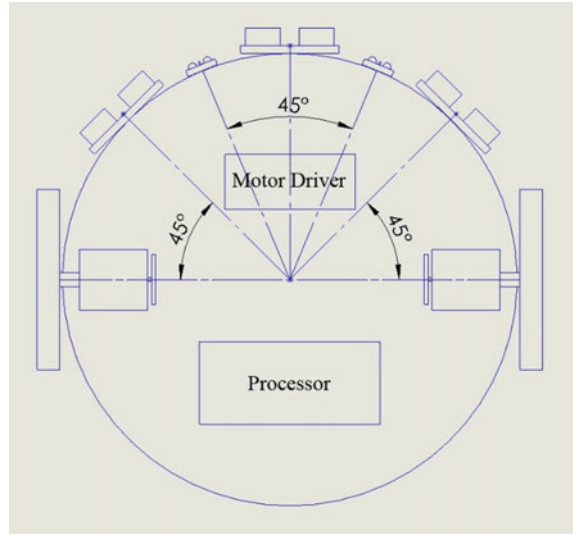
powered by the same voltage regulator and controlled by the microprocessor. The IR and sonar obstacle detection sensors are placed on the front part of the robot body. In this design the robot maps and finds its path on the fly, no prior knowledge of the cleaning area is given. Figure 2 shows the plate on the robot where the sensors are placed.

When the drive motors are selected, the torque calculations are carried out such that the forward speed of the robot should be 70–80 mm/s, depending on the surface friction, and the robot total weight is 6 kg where the dust bag is completely full. Depending on the wheel size the robot rotation speed is chosen as 10 rpm.

In order to drive the motors MC 33926 motor drivers are used. Especially, motor current feedback is used to understand if the robot cannot move. In this case, the robot is rotated and obstacle problem is solved.

To process the sensory data and drive the motors a microprocessor namely, an Arduino Mega, is used. Especially the 6 interrupts become necessary to read the encoder data from the 3 motors. By using the encoders which have 48 counts resolution, the amount of wheel rotation and cable position is measured. Later the speed of motors is also calculated.

Sonar sensors are used to detect the obstacles around the room and they use the concept of time of flight to measure the distance. On the robot, 3 HC-SR04 type sonar sensor is used and they are placed at $0°$, $45°$ and $-45°$ to look in front of the robot. With these sensors, 20–4000 mm distance can be measured with the accuracy of 3 mm. In order to create a safer system apart from sonar sensors, infra-red sensors, TCRT5000, are also placed in front or the robot at $22.5°$ and $-22.5°$ to create redundant system.

4   Navigation Algorithm

The existing vacuum cleaning robot has the advantage of not having a cable so that their motion is not restricted. In the design of the wired vacuum cleaner navigation algorithm the biggest constrain is the cable. It is assumed that the cable is long enough that the robot can cover the surface area and in order to stop the tangling of the cable the robot lays down the cable when it is moving and the cable stays on the ground until the robot returns.

Figure 3 shows how the navigation algorithm works. The robot moves in a random direction where it is plugged in at the distance of Lmax – L(1). Lmax shows the cable length and L(1) is the first step cleaning length. After the robot comes to the cleaning start point, the vertical distance Y(1) is considered as constant. By using Eq. 10, the distance to travel on the x axis is calculated while the robot does not pull the cable.

$$X(1) = \frac{Y(1)}{\tan \varphi} \tag{10}$$

Later, the robot draws a rectangle by moving and turning counterclockwise and comes back to the cable end point. At this stage since more cable is retracted, the distance that the robot can move on the x axis increase considering the y axis distance is constant. The increased cable length to cover the surface is formulated as

$$\Delta X = \sqrt{2^2 L^2(1) - Y^2(2)} - X(1). \tag{11}$$

Fig. 3 Geometrical representation of the navigation algorithm
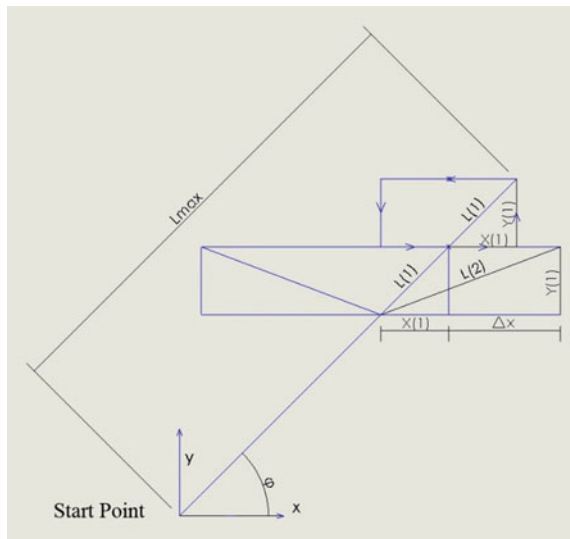
Where ΔX is the extra cable length which the robot can use to carry and it can be generalized as

$$\Delta X(i) = \sqrt{i^2 L^2(1) - Y^2(1)} - X(1). \tag{12}$$

If the robot hits an obstacle on the cleaning surface, it continues its own routine while drawing a smaller rectangle at that step. After the first cleaning is completed, the robot cleans the mirror side of the room. Later another angle is assigned by looking at the unvisited areas and the robot continues to clean as started.

To keep track of the obstacle in the room and the borderlines of the room a matrix is used as the memory of the robot. In order to represent unvisited areas, obstacles/borders and covered areas 0, 1 and 2 are used respectively. After the first cleaning cycle, the robot looks at the cleaning matrix to decide where to clean.

5 Simulation Results

In order to understand the effectiveness of the navigation algorithm a simulation study is carried out in Matlab. When the simulation is designed especially simple mathematical operations were used considering that the algorithm will be running on a microprocessor on the robot. The obstacles are, shown in black and are placed randomly inside a room at a random size, where the total number of obstacle are given. In the real application, the size and position of the obstacle will vary. Figure 4 shows how the robot moves without any obstacle.

Point 1 refers to the start point. Point S refers to the vacuum cleaning robot and the blue zones show the cleaned area.

Figures 5 and 6 show the simulation results for a case where there are 3 obstacles in the room. The obstacles are random sized are randomly placed in the

Fig. 4 Navigation algorithm with no obstacle
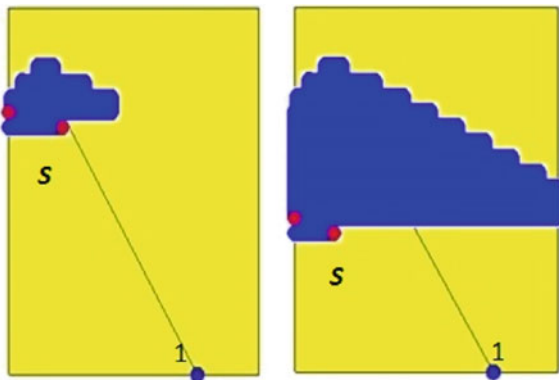
Fig. 5 Simulation start with obstacles
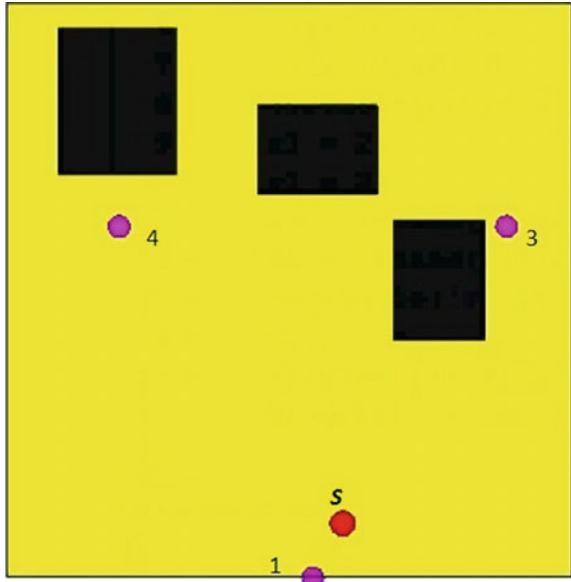
Fig. 6 Robot motion around the obstacle
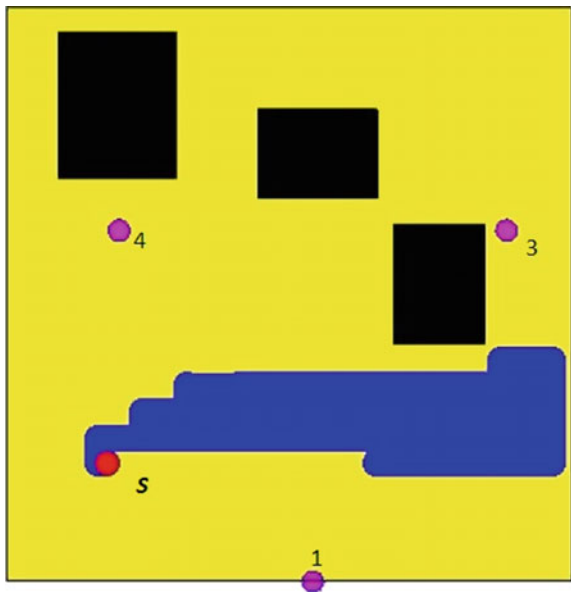
room. Point 1 shows the starting point and point S shows the vacuum cleaner. Point 3 is the random target point and point 4 is the mirror point of point 3. When the robot is moving toward the point 3, because of the obstacle it cannot pass so it works around the obstacle, Fig. 6.

Fig. 7 Second step of the cleaning process
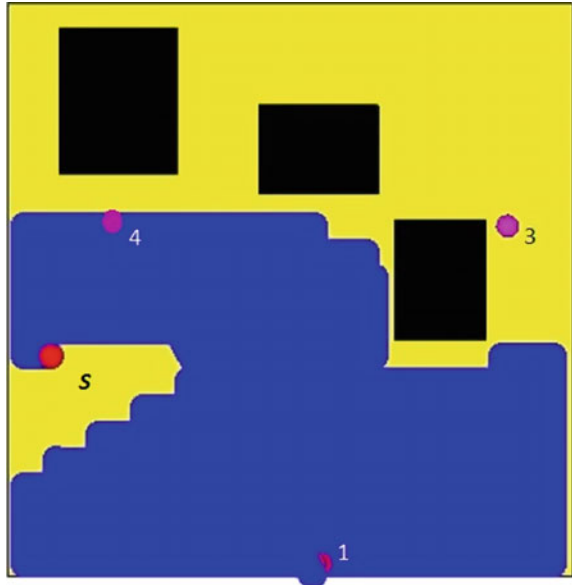
Fig. 8 Memory matrix for the cleaning robot

1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
0	0	2	1	1
0	0	2	1	1
0	0	2	1	1
0	0	2	1	1
0	0	2	1	1

The first part of the cleaning stops when the robot reaches to the start point 1. In the second part of the cleaning, the robot moves towards the point 4 and continues to clean the area (Fig. 7).

While the robot is moving it keeps track of the unvisited area, obstacles/borders and cleaned area are shown in Fig. 8.

6 Conclusions and Future Work

This paper explain the necessary software, hardware and algorithm to convert a classic vacuum cleaner into an autonomous cabled robot vacuum cleaner. The effectiveness of the navigation algorithm is tested with a simulation program and it

is observed that the algorithm can navigate the robot around the obstacles and cover the area as much as the cable length allows. In the future work of the project the navigation algorithm will be designed to run on a microprocessor. The computing power and allowed memory poses a challenge.

References

Darwish, A. (2012). *Cost based navigation for autonomous vacuum cleaners. Robot Intelligence Technology and Applications* (pp. 415–423).
Fernandes, J., & Claraco J. (2013). *Robotic bases. Simultaneous localization and mapping for mobile robots* (pp. 30–38). Hershey: Springer.
Gantz, C. (2012). *1990 to the present. The vacuum cleaner: A history* (pp. 201–205). North Carolina: McFarland.
Milford M. (2008). *Mobile robots. Robot navigation from nature* (pp. 1–3). Berlin, Germany: Springer.
MATLAB Control OpenGL Rendering Documentation, The Math Works Inc.

Human Safety Index Based on Impact Severity and Human Behavior Estimation

**Gustavo Alfonso Garcia Ricardez, Akihiko Yamaguchi,
Jun Takamatsu and Tsukasa Ogasawara**

Abstract With the increasing physical proximity of Human–Robot Interaction (HRI), ensuring that robots do not harm surrounding humans has become crucial. We propose to quantitatively evaluate human safety by modeling the human behavior so that it maximizes the potential injuries in a given situation. The potential injuries are rated using the impact severity of a collision. Therefore, we estimate the human motion that maximizes the impact severity, which is calculated considering a collision between the estimated state of the human and a future state of the robot given a control input. Through simulation experiments using two test cases and three HRI scenarios, we demonstrate that our method keeps human safety while achieving a competitive performance.

Keywords Safety index · Human injury · Human behavior estimation · Human safety · Human robot interaction · Humanoid robot

1 Introduction

Robots are increasingly finding transcendent applications in daily-life environments due to the potential economical and social benefits of the symbiosis between humans and robots, as stated by Ogorodnikova (2008). The harmonic coexistence

G.A. Garcia Ricardez (✉) · A. Yamaguchi · J. Takamatsu · T. Ogasawara
Graduate School of Information Science, Nara Institute of Science
and Technology, 8916-5 Takayama, Ikoma, Nara, Japan
e-mail: garcia-g@is.naist.jp

A. Yamaguchi
e-mail: akihiko-y@is.naist.jp

J. Takamatsu
e-mail: j-taka@is.naist.jp

T. Ogasawara
e-mail: ogasawar@is.naist.jp

between humans and robots depends on the ability of the robots to keep human safety, as explained by De Santis et al. (2008). Therefore, the measurement of how safe or dangerous the interaction between human and robots is has become crucial.

The challenge of human safety is to control the robot in a safe and efficient way. Concretely, the problem can be described as assessing human safety given a state and a robot control input while satisfying a performance requirement. The assessment of human safety is given by a *safety index*, which is a scalar value that represents the safety of humans when interacting with a robot. This index provides a quantitative evaluation of human safety which can be used for robot control.

This chapter proposes a safety index to assess the human safety by considering the consequences of the most dangerous situation. The danger of a situation is derived from the potential injuries, i.e., the situation which could cause the most severe injuries is considered the most dangerous situation. Once we estimate the most dangerous situation, we can control the robot to avoid potential injuries which exceed acceptable limits.

The proposed safety index is based on a physical model to calculate the danger, and a human model to predict the human motion. We rate the potential injuries according to the severity of a collision between the human and the robot. The severity of a collision derives from an analysis of physical quantities before and at the time of a collision. As an actual collision is undesired, we consider a collision between an estimated human state and a future robot state. To do this, we predict the human behavior from the observed human state and the human model, and calculate the future state of the robot from the control input.

In the literature, we find that Ikuta et al. (2003) proposed a safety index based on impact force and impact stress where safety is evaluated as the ratio between a producible impact force and a predefined *critical impact force*. The *critical impact force* is the minimum impact force required to injure humans. In our proposed method, the danger is associated to the injuries a robot can inflict to a human if a collision occurs. The harshness of the injuries derives from the impact severity of such collision. Haddadin et al. (2012) studied the relation between inertia, velocity, and geometry at possible impact locations and the resulting injury of a collision. Then, the obtained knowledge about injuries is used for control. Echávarri et al. (2013) proposed a safety index based on head injury indexes from the automobile industry. In this work, we propose to approximate the potential human injuries by calculating the impact severity of a collision. The work of Ding et al. (2011) estimates the human arm motion to generate a forbidden region that the planner should avoid. Our proposed method estimates the human behavior by considering the position of multiple points on the human to evaluate human safety.

To test the proposed safety index, we perform simulation experiments where the human and robot do simple tasks in shared workspaces. The experimental results validate the proposed approach to guarantee human safety while keeping a competitive efficiency.

The rest of the chapter is organized as follows. Section 2 details the proposed safety index. Section 3 describes the test bed. Section 4 presents the simulation results. Finally, Sect. 5 concludes this chapter.

2 Safety Index

In this section, a novel human safety metric based on the severity of a collision and human behavior estimation is proposed.

The proposed human safety metric, or *safety index*, is a solution to the challenge of human safety assessment. Using distance and the direction of robot's motion as components of a safety metric has been proved to be effective to keep human safety, as described by Garcia Ricardez et al. (2015). Nevertheless, this approach is not explicitly related to potential human injuries and is purely reactive. The proposed method associates the danger to potential human injuries by calculating the impact severity of a collision and anticipates the human motions to restrict the robot speed earlier, which decreases the risk of a collision.

The inputs of the safety index algorithm are the current human and robot states, which are defined by position and linear velocity, and the robot control input defined by the target velocity. The output is a numerical value associated to the impact severity of the worst human action, which represents the danger of the situation (Fig. 1).

Algorithm 1 Safety Index

Input: $\mathbf{x}_{h0}$, $\mathbf{v}_{h0}$, $\mathbf{x}_{r0}$, $\mathbf{v}_r^*$
Output: *SI*

1. $\hat{\mathbf{a}}_h \leftarrow WHA(\mathbf{x}_{h0}, \mathbf{v}_{h0}, \mathbf{x}_{r0}, \mathbf{v}_r^*)$
2. $\mathbf{x}_h' \leftarrow \mathbf{x}_{h0} + \mathbf{v}_{h0}\Delta t + \frac{1}{2}\hat{\mathbf{a}}_h\Delta t^2$
3. $\mathbf{v}_h' \leftarrow \mathbf{v}_{h0} + \hat{\mathbf{a}}_h\Delta t$
4. $\mathbf{x}_r' \leftarrow \mathbf{x}_{r0} + \mathbf{v}_r^*\Delta t$
5. $SI \leftarrow -DS(\mathbf{x}_h', \mathbf{v}_h', \mathbf{x}_r', \mathbf{v}_r^*)$

Fig. 1 Human and robots states, both current (*t*) and at the time horizon (*t*+Δ*t*). The worst human action is $\hat{\mathbf{a}}_h$
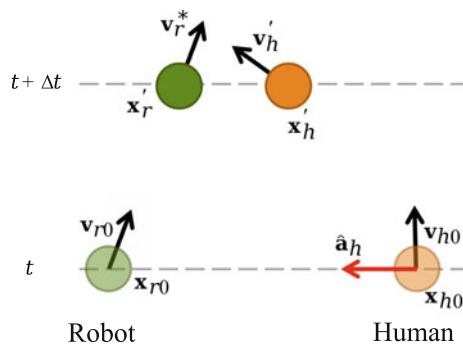
2.1 Algorithm

The purpose of this algorithm is to provide an instant measurement of the human safety given a situation and a robot control input.

The safety index algorithm takes the current human and robot states as input, and is executed at each time step. The algorithm outputs the safety index *SI* which is a numerical value that represents how dangerous the input states are. The values of *SI* are real numbers where lower values indicate more danger. Therefore, the lowest value represents the most dangerous situation, i.e., a collision, while the highest value implies a safe situation. The outline of the algorithm is as follows:

(a) Estimate the most dangerous human motion given the human and robot states and the robot control input.
(b) Calculate the human and robot states at the time horizon $t + \Delta t$ using the estimated human motion and the robot control input.
(c) Calculate a danger score *DS* assuming a collision between the calculated states.
(d) Output the safety index *SI* as the additive inverse of the danger score *DS*.

The complete computation at each time step is shown in Algorithm 1, where the input $\mathbf{x}_{h0}$, $\mathbf{v}_{h0}$, and $\mathbf{x}_{r0}$ are the current human and robot positions and velocities, respectively, and $\mathbf{v}_r^*$ is the robot control input. The worst human action is denoted by $\hat{\mathbf{a}}_h$. The human and robot states at the time horizon $t + \Delta t$ are denoted by $\mathbf{x}_h'$, $\mathbf{v}_h'$, and $\mathbf{x}_r'$, respectively. Note that the human states at the time horizon depend on the worst human action while the robot state depends on the control input. Finally, the safety index is denoted by *SI* and the danger score by *DS*.

2.2 Danger Score

The danger is assessed using the impact severity of a collision between the human and the robot. Actually, this is a collision that would occur if the human takes the worst action, i.e., the most dangerous situation. Therefore, the collision takes place between an estimated human state and a future robot state, where the danger is rated using a danger score *DS*.

More specifically, the danger score *DS* is a function of the human and the robot states, defined as follows:

$$DS(\mathbf{x}_h, \mathbf{v}_h, \mathbf{x}_r, \mathbf{v}_r) = \begin{cases} ds_{\max} & \|\mathbf{d}\| = 0, \\ f_S(\|\mathbf{J}\| f_G(\|\mathbf{d}\|)) & \|\mathbf{d}\| > 0, \frac{\pi}{2} < \phi \leq \pi, \\ f_S((1-\rho)\|\mathbf{J}\| f_G(\|\mathbf{d}\|)) & \|\mathbf{d}\| > 0, 0 \leq \phi \leq \frac{\pi}{2}, \end{cases} \quad (1)$$

where $\mathbf{x}_h$, $\mathbf{v}_h$, $\mathbf{x}_r$, and $\mathbf{v}_r$ are the human and robot states, respectively.[1] The impact severity is associated to the impulse $\mathbf{J}$. The displacement vector between human and robot is denoted by $\mathbf{d}$. The functions f_S and f_G are Sigmoid and Gaussian functions, respectively. The angle ϕ is the angle between $\mathbf{d}$ and the vector of the relative velocity between the human and the robot. The reliability of the human estimation is denoted by ρ. The maximum value of the danger score is denoted by ds_{max}. In this work, the danger score DS takes values of $\mathbb{R} \to \{2, [0, 1]\}$, i.e., $ds_{\mathrm{max}} = 2$. The details of the danger formulation are described below.

2.2.1 Impact Severity

The impact severity is obtained by calculating the impulse generated during a collision. Impulse $\mathbf{J}$ is defined as the change in the momentum $\mathbf{p}$,

$$\mathbf{J} = \Delta\mathbf{p} = m\mathbf{v}_f - m\mathbf{v}_0, \tag{2}$$

where m is the combined mass of the human and the robot, $\mathbf{v}_f$ is the final velocity, and $\mathbf{v}_0$ is the initial velocity.

Under the assumption that the collision is perfectly inelastic, the final velocities are discarded and we obtain:

$$\mathbf{J} = -m(\mathbf{v}_{h0} - \mathbf{v}_{r0}), \tag{3}$$

where $\mathbf{v}_{h0}$ is the human initial velocity, and $\mathbf{v}_{r0}$ is the robot's initial velocity.

The numerical value of the impact severity corresponds to the impulse norm $\|\mathbf{J}\|$. We normalize this value using a Sigmoid function f_S, as numerical values of DS are easier to interpret if they fall within a known range, namely $\mathbb{R} \to [0, 1]$.

2.2.2 Collisions

To be able to calculate the DS we assume that the human and robot collide. This collision actually does not happen but we use it to calculate the impact severity.

We consider three cases:

  i. A collision will occur in the worst case.
 ii. No collision occurs but the future situation seems to be still dangerous.
iii. No collision occurs and the future situation seems to be safe.

The case (i) reflects situations where the distance between the human and the robot is zero before the time horizon, which constitutes a very dangerous situation for the human. The discrimination of cases (ii) and (iii) is based on the angle

---

[1]The *SI* algorithm depends on the human estimation but the danger score does not. Therefore, the input of DS is any human and robot states, whether estimated or current.

between the relative velocity vector $\mathbf{v}_{\text{rel}}$ and the displacement vector $\mathbf{d}$, derived from their dot product:

$$\phi = \cos^{-1}\left(\frac{\mathbf{d} \cdot \mathbf{v}_{\text{rel}}}{\|\mathbf{d}\| \|\mathbf{v}_{\text{rel}}\|}\right), \tag{4}$$

where $\mathbf{v}_{\text{rel}} = \mathbf{v}_h - \mathbf{v}_r$ and $\mathbf{d} = \mathbf{x}_h - \mathbf{x}_r$. In case (ii), the angle between the relative velocity and displacement vectors is $\frac{\pi}{2} < \phi \leq \pi$, while in case (iii) this angle is $0 \leq \phi \leq \frac{\pi}{2}$.

A human estimation reliability factor ρ is introduced in the case (iii) to account for the uncertainty of the human motion. This is because we consider less dangerous a situation where a collision is less likely to happen. Nevertheless, the human could erratically change his motion, leading to a danger underestimation. If the reliability is null ($\rho = 0$), the danger around the boundary between cases (ii) and (iii) is similar. In the ideal (but unrealistic) case that the reliability is maximum ($\rho = 1$), the danger in case (iii) can be neglected. In this research, a typical value of the reliability is 0.2.

The distance between the human and the robot at the moment of the collision is used as an attenuating factor of the danger. This is because a collision far away from the human is less likely to become an actual collision and, therefore, the danger that it represents for the human decreases. On the contrary, a collision when the human and the robot are very close should be considered as more dangerous since it is more likely to become an actual collision. The proposed method models the attenuation caused by distance using a Gaussian function f_G.

2.2.3 Asymmetricity

We approach the trade-off between human safety and robot's efficiency using an asymmetric speed restriction.

The key to make the safety index asymmetric resides in assigning higher scores to cases where the human and the robot are approaching each other (case (ii)) and lower danger scores to the cases where the human and the robot are moving away from each other (case (iii)). To do this, the parameters of the Gaussian and Sigmoid functions are set differently for cases (ii) and (iii). In practical terms, the case (ii) has a wide Gaussian bell and a slowly-increasing Sigmoid function, while the case (iii) has a narrow Gaussian bell and a fast-increasing Sigmoid function.

Figure 2 shows an example of the danger relation to speed and distance for case (ii). The small circles highlight the following features of this example. The danger has a value of 0.1 when the distance is 0.6 [m] and the speed is 0.25 [m/s], while the danger has a value of 0.5 at the same speed when the distance is 0.25 [m].
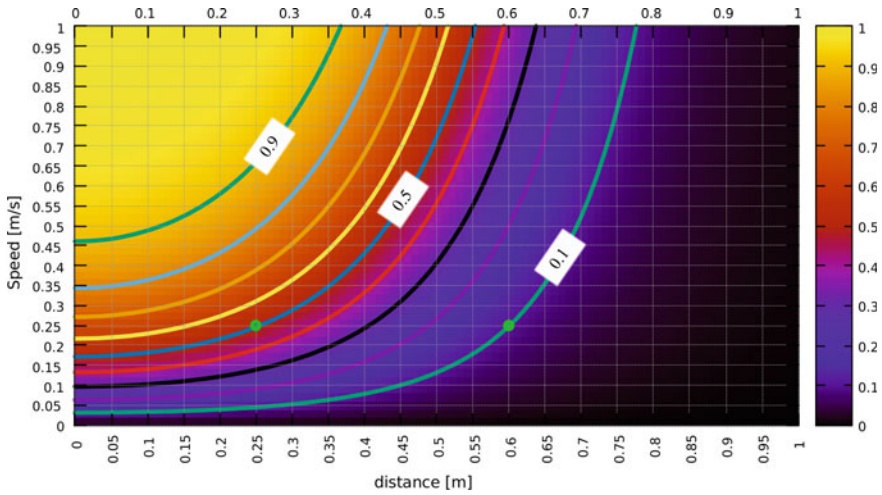
Fig. 2 Relation of the danger score to the speed and distance for case (ii). The danger is 0.1 and 0.5 at 0.6 [m] and 0.25 [m], respectively, at 0.25 [m/s]

2.3 Human Behavior Modeling

The purpose of modeling the human behavior is to find the human action that results into the most dangerous situation for the human, i.e., the worst human action (WHA). This idea is based on the assumption that if the most dangerous situation is accounted for, the human safety assessment of all other (less dangerous) situations is, consequently, included.

In the proposed scheme, a human action is represented by an acceleration vector. Since this acceleration modifies the human state (position and velocity), it should not produce motions that exceed the physical limits of the human body. We introduce a constant A_h that represents the norm of the maximum humanly feasible acceleration.

The estimation of the WHA can be formulated as an optimization problem:

$$\begin{aligned}
\underset{\mathbf{a}_h, t}{\text{maximize}} \quad & DS\big(\mathbf{x}_h(\mathbf{a}_h, t), \mathbf{v}_h(\mathbf{a}_h, t), \mathbf{x}_r(t), \mathbf{v}_r^*\big), \\
\text{subject to} \quad & \|\mathbf{a}_h\| \leq A_h, \quad 0 < t < \Delta t,
\end{aligned} \tag{5}$$

where $\mathbf{x}_h$ and $\mathbf{v}_h$ are the human states depending on the worst human action and time, and $\mathbf{x}_r$ is the robot position depending on the control input $\mathbf{v}_r^*$ and time. A_h is the maximum human acceleration and the objective function is the danger score DS. As the proposed safety index uses a short-term human behavior estimation, we set a finite period of time Δt (e.g., 400 [ms]) within which the effects of human actions are predicted. In other words, $\mathbf{a}_h$ is a feasible human action with the highest danger score within the time horizon.

Fig. 3 The WHA algorithm finds the human action $\hat{\mathbf{a}}_h$ that maximizes the danger
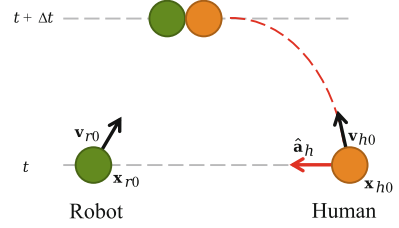
The optimization problem is solved using a Gradient Descent method. To reduce the complexity of this optimization problem, the objective function can be substituted by a function of the distance between the human and robot, namely:

$$F(\mathbf{x}_{h0}, \mathbf{x}_{r0}) = \frac{1}{2}\|\mathbf{x}_{h0} - \mathbf{x}_{r0}\|^2. \tag{6}$$

Algorithm 2: Worst Human Action

Input: $\mathbf{x}_{h0}$, $\mathbf{v}_{h0}$, $\mathbf{x}_{r0}$, $\mathbf{v}_r^*$
Output: $\hat{\mathbf{a}}_h$
1: **for** each $t_i \in [t, t + \Delta t]$ **do**
2:    $\mathbf{x}'_{r_i} \leftarrow \mathbf{x}_{r0} + \mathbf{v}_r^* t_i$
3:    Find $\mathbf{a}'_{h_i}$ that minimizes F for $\mathbf{x}'_{r_i}$
4: **end for**
5: $\hat{\mathbf{a}}_h \leftarrow \mathbf{a}'_{h_i}$ with minimum F value

---

With this simplification, the optimization problem can be reformulated as: (Fig. 3)

$$\begin{aligned} \underset{\mathbf{a}_h, t}{\text{minimize}} \quad & F(\mathbf{x}_h(\mathbf{a}_h, t), \mathbf{x}_r(t)), \\ \text{subject to} \quad & \|\mathbf{a}_h\| \le A_h, 0 < t < \Delta t. \end{aligned} \tag{7}$$

Algorithm 2 shows the entire computation of the proposed WHA algorithm at each time step. The input is the current human and robot states $\mathbf{x}_{h0}$, $\mathbf{v}_{h0}$, and $\mathbf{x}_{r0}$, respectively. The robot control input is denoted by $\mathbf{v}_r^*$. The output is the worst human action represented by the acceleration vector $\hat{\mathbf{a}}_h$. The time Δt is discretized to make the algorithm practical and is denoted as t_i.

The outline of the algorithm is as follows:

(a) Calculate the robot positions $\mathbf{x}'_{r_i}$ for each time sample t_i (step 2).
(b) Estimate a worst human action candidate $\mathbf{a}'_{h_i}$ for each robot positions $\mathbf{x}'_{r_i}$ using a Gradient Descent method (step 3).
(c) Output the candidate $\mathbf{a}'_{h_i}$ with the minimum value of F as the worst human action (step 5).

3 Experimental Setup

This section describes the testbed, the proposed test cases and HRI scenarios used for the simulation experiments.

We use a virtual environment where the human and robot are modeled using cylinders and spheres. The robot model corresponds to the humanoid robot HRP-4 described by Kaneko et al. (2011). The human model is obtained from a motion capture system consisting of sensors attached to the human body.

3.1 Compared Methods

We compare the effect of these methods on human safety and the robot's efficiency:

Conventional. This method separates the human and the robot during the interaction by stopping the robot every time the human enters the robot's workspace, which is delimited by a fixed distance threshold.
AVM. Asymmetric Velocity Moderation (AVM) is a method that calculates speed restrictions for every combination of robot joints and human links. This method restricts the robot speed considering the distance and the direction of the motion.
SI (proposed). Safety index that calculates the impact severity and estimates the most dangerous situation.
No safety restriction. This is a reference to compare the above listed methods to the case where no safety method is applied and no human is present, i.e., the robot moves freely.

By comparing SI to the conventional method, we demonstrate how the proposed approach benefits the robot's efficiency without sacrificing human safety. By comparing SI to AVM, we validate the advantage of using human behavior estimation.

3.2 Test Cases

We use two test cases to verify the proposed safety index. We consider that these test cases are dangerous situations that humans may face in daily-life scenarios.

TC-1. The human and the robot are moving their hands toward a direct collision, as shown in Fig. 4. In this dangerous situation, the velocity vectors are parallel.
TC-2. The human and the robot are moving their hands toward a common region. In this dangerous situation, the velocity vectors are perpendicular.
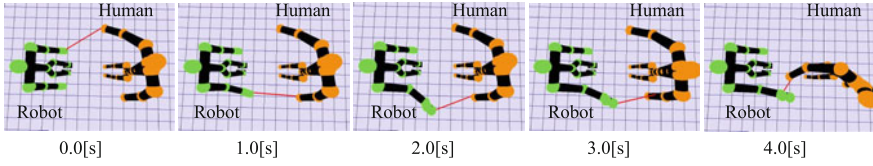
Fig. 4 Screenshots of the test case TC-1. The human and the robot move their hands toward each other

3.3 Simplified HRI Scenarios

To verify the performance of the proposed method, we use three experimental setups which depict HRI scenarios where a humanoid robot could endanger a human, shown in Fig. 5. We selected these scenarios considering what we envision could be common situations in daily-life scenarios where a robot performs a task in human proximity.

The human and the robot perform independent tasks around the table. The task of the human consists of moving balls between three boxes on a table. The task of the robot consists of placing the end-effector 30 times in different positions on the table, chosen randomly out of five predefined positions. The circles in Fig. 5 indicate these five different positions, which are on the same plane about 20 cm above the table and are reachable by the robot.
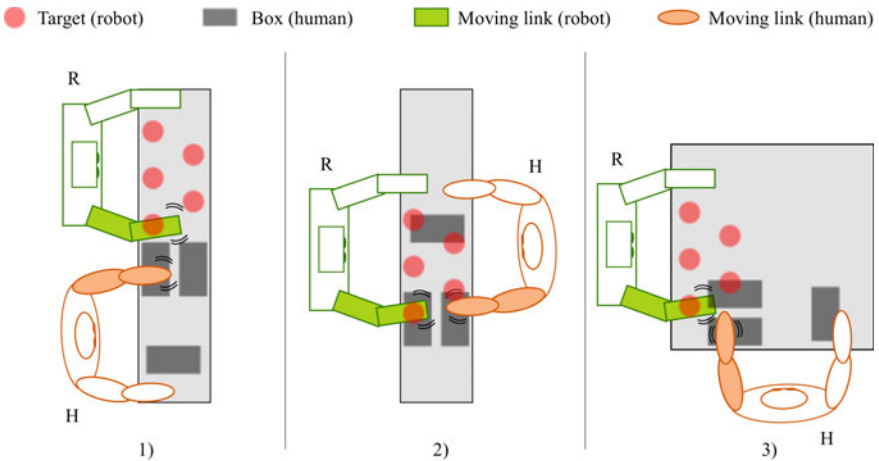
Fig. 5 Simplified HRI scenarios with a humanoid (R) and a human subject (H) performing independent tasks. The *circles* indicate the five different positions above the table where the robot can place the end-effector. The *rectangles* are the boxes among which the human subject moves objects. The moving links are indicated in *solid color*

4 Results

In this section, we compare the proposed safety index to AVM in terms of (a) human safety by simulating dangerous situations with two test cases, and (b) efficiency using three simplified HRI scenarios. We use a fair comparison criteria consisting of setting the parameters of AVM and the safety index to make them start the robot speed restriction simultaneously when the human is standing still.

4.1 Human Safety Comparison

In these experiments, we compare the robot's response to the incoming collision with the human hand using the proposed safety index and AVM. We compare the speed profiles of these methods using the same human motion and the same robot's trajectory.

In TC-1, the robot starts to restrict the robot speed earlier than AVM, as shown in Fig. 6(a). This is because the safety index is estimating the WHA which consists of increasing the velocity of the human end-effector toward the robot's end-effector. The impact severity of such estimated velocity contributes to increase the danger score. Moreover, the estimated human position is closer than the actual human which also increases the danger score.

Similarly in TC-2, the robot speed restriction starts earlier than AVM, as shown in Fig. 6(b). In this case, the WHA consists of increasing the velocity toward the point where the trajectories of the human and robot end-effectors intersect. The increase in the velocity contributes to a higher danger score. The estimated human position in the vicinity of the trajectories' intersection also contributes to increase the danger score.

Figure 7 shows the speed profiles of the human and robot's end-effectors for TC-1. Before the human moves, the robot speed with both methods is the same. With the proposed safety index, the robot speed is restricted earlier than with AVM, approximately 200 [ms] before. This translates into an increase of distance for reaction to the human motions, e.g., if the relative velocity is 0.8 [m/s], the restriction of SI would occur 0.16 [m] before AVM (Fig. 8).

In summary, the safety index is more cautious when approaching the human in anticipation of a collision.

4.2 Efficiency Evaluation

Simulation experiments are carried out to evaluate the efficiency of the proposed safety index. We tested the proposed method in three simplified HRI scenarios,
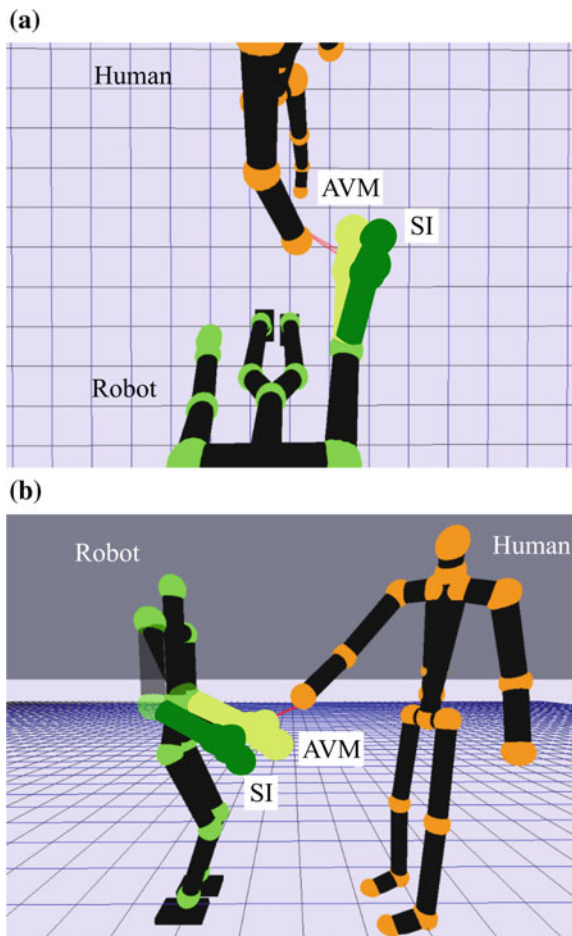
Fig. 6 Simulation experiments with the test cases for the human safety comparison. The superposition of the robot states shows the configuration where the robot stops with each method. The *clear color* corresponds to AVM while the *dark color* corresponds to SI
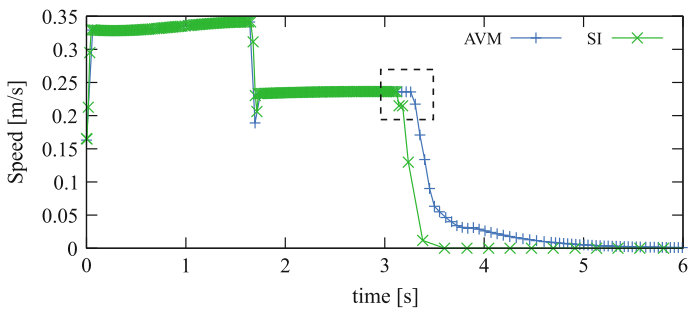
Fig. 7 Speed profiles of SI and AVM for the test case TC-1. The speed profiles show an earlier stop of SI which demonstrates the advantage of the human behavior estimation
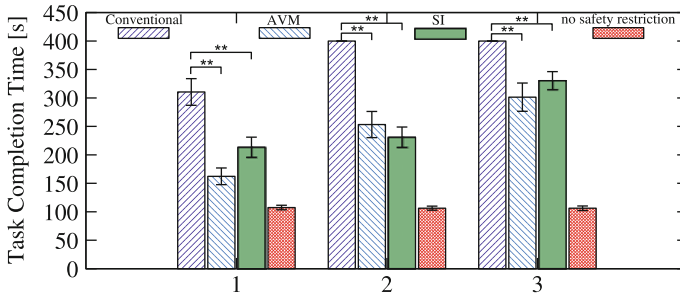
Fig. 8 Task completion time of the compared methods in the three HRI scenarios. Lower time indicates better robot's performance. Setups 1 and 3 show the expected delay of SI caused by the earlier restriction. In setup 2, SI shows a better performance but not significantly

where we measure the task completion time as a notion of the the robot's efficiency. The experiments consist of executing each method 10 times for each setup.

As the proposed safety index anticipates the human motion, the speed restriction occurs earlier than AVM. In setups 1 and 3, the delay caused by this earlier restriction is patent. In setup 2, the proposed safety index shows a better performance but not significantly.

In all setups, SI shows a better performance than the conventional method. While SI is outperformed by AVM in terms of the task completion time, SI is still competitive.

5 Conclusions

We proposed a safety index as a human safety metric which considers the potential human injuries and estimates the human behavior to anticipate dangerous situations. Simulation experiments involving a human and a humanoid robot were carried out. Additionally, we evaluated the robot's efficiency in simplified HRI scenarios in terms of the task completion time.

The proposed safety index estimates the worst possible human action to determine the safety of the interaction. By considering the worst human action, the proposed method provides a numerical evaluation of the most dangerous situation. Human safety assessment is related to the injuries that the robot could inflict to the human if a collision occurred when the human takes the worst possible action.

The experiments results show an earlier restriction and a cautious robot behavior as it effectively estimates the human behavior anticipating the most dangerous situation. Furthermore, the experiments results show that the proposed safety index benefits the efficiency of the robot.

References

De Santis, A., Siciliano, B., De Luca, A., & Bicchi, A. (2008). An atlas of physical Human-Robot Interaction. *Mechanism and Machine Theory, 43*(3), 253–270.

Ding, H., Reißig, G., Wijaya, K., Bortot, D., Bengler, K., and Stursberg, O. (2011). Human arm motion modeling and long-term prediction for safe and efficient Human-Robot Interaction. In *Proceedings of the 2011 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 5875–5880).

Echávarri, J., Ceccarelli, M., Carbone, G., Alén, C., Muñoz, J. L., Díaz, A., & Munoz-Guijosa, J. M. (2013). Towards a safety index for assessing head injury potential in service robotics. *Advanced Robotics, 27*(11), 831–844.

Garcia Ricardez, G. A., Yamaguchi, A., Takamatsu, J., & Ogasawara, T. (2015). Asymmetric Velocity Moderation for human-safe robot control. *Advanced Robotics, 29*(17), 1111–1125.

Haddadin, S., Haddadin, S., Khoury, A., Rokahr, T., Parusel, S., Burgkart, R., et al. (2012). On making robots understand safety: Embedding injury knowledge into control. *The International Journal of Robotics Research, 31*(13), 1578–1602.

Ikuta, K., Ishii, H., & Nokata, M. (2003). Safety evaluation method of design and control for human-care robots. *International Journal of Robotics Research, 22*(5), 281–297.

Kaneko, K., Kanehiro, F., Morisawa, M., Akachi, K., Miyamori, G., Hayashi, A., et al. (2011). Humanoid robot HRP-4—humanoid robotics platform with lightweight and slim body. In *Proceedings of the 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 4400–4407).

Ogorodnikova, O. (2008). Methodology of safety for a human robot interaction designing stage. In *Proceedings of the 2008 Human System Interactions Conference* (pp. 452–457).

Swarm Robots' Communication and Cooperation in Motion Planning

Khiem N. Doan, An T. Le, Than D. Le and Nauth Peter

Abstract This paper focuses on swarm communication and how swarm robot individuals communicate with each other. By imitating swarms in nature, we can develop many features for swarm robots depending on the working environment, size of the robot, the number of individuals and the available budget. We researched on the development of our low budget swarm robot project's communication solution using HM-10 Bluetooth modules. In addition, we also studied about several basic tasks in swarm robotics: synchronize information, calculate individuals' position in the workspace, path planning and cooperate to solve many complicated tasks. To demonstrate swarm robot's main tasks-space, we decided to use a group of two RP6 robots. Although they are completed built robots and cannot be minimized in the size or simplified in the structure, they still have various free ports and connections so that we can attach extensions or sensors to solve some basic swarm robotics problems.

Keywords Generated path-planner · Mobile robot · Multi-robots communication

1 Introduction

The two main reasons for us to pay attention on swarm robot are the simplicity of each individual and the complex behavior of the swarm. Thanks to the simplification in each individual's design and the high population of swarm robot, they show advantages over normal robots in: solving tasks which cover large area, in

K.N. Doan · A.T. Le · T.D. Le (✉) · N. Peter
Vietnamese-German University, Binh Duong, Vietnam
e-mail: than.ld@vgu.edu.vn

N. Peter
e-mail: pnauth@fb2-uas.de

K.N. Doan · A.T. Le · T.D. Le · N. Peter
Frankfurt University of Applied Science, Frankfurt am Main, Germany

Fig. 1 RP6 robot
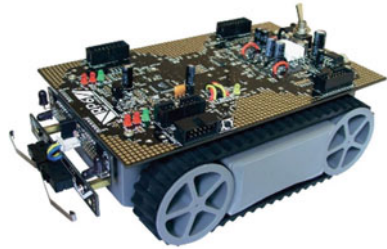
tasks dangerous to normal robots and in tasks which require a scaling population (Ying and Zhong-Yang 2013).

As mentioned above, RP6 robot kits (Fig. 1) are used as swarm individuals. The reason is that we can get deeper into how the robot work, and freely develop the robot without concerning about the lack of hardware (if the design is too simple) or the complexity of it (if the design is too complicated).

After researching on communication solutions and swarm robot behaviors, Bluetooth HM-10 Module has been chosen as robot's communication solution. In the project's range, this module is the best suitable choice thanks to its complete build, ease of use, range of connection and low cost. However, this is just for the prototype because this solution provides only one active connection in the same time. Hence, robots can connect together only in pairs. For real swarm robots, we need to find another solution for communication that support multi-connection.

Path planning is another important part that must be paid attention on in swarm robot. RP6 robot's kinematic model and obstacle detecting algorithms are researched, which help us in robot position monitoring and path planning programming.

About the structure, there are two main parts in this paper. Section 2 (Kinematic models and obstacle avoiding algorithms) describes theories that is used in the project, and Sect. 3 (Experiments and results) depictures how the system is realized and tested.

2 Kinematic Models and Obstacle Avoiding Algorithms

In this section, the kinematics for the swarm mobile robots as well as obstacle avoidance strategies are explained.

2.1 Differential Drive Kinematics

By varying the rotating speed of each wheel Canudas de Wit et al. (1996), the robot can take the defined trajectory.
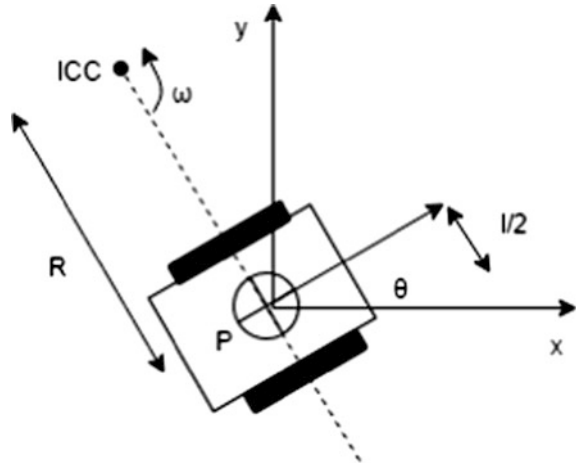
During rotation, the robot will turn around a point belong to its wheel axis, which is called ICC—Instantaneous Center of Curvature (Fig. 2). While ω is the angular rate of rotation about ICC of robot, we get:

$$\omega(R + \frac{l}{2}) = V_R \tag{1}$$

$$\omega(R - \frac{l}{2}) = V_L \tag{2}$$

with
R   Distance between ICC and robot's center.
l   Distance between two wheels.
V_R, V_L velocities of right and left wheel.
   From (1) and (2) we get:

$$\frac{V_R}{V_L} = \frac{2R + l}{2R - l} \tag{3}$$

Therefore:

$$R = \frac{l}{2}\frac{V_R + V_L}{V_R - V_L} \tag{4}$$

And

$$\omega = \frac{V_R - V_L}{l} \tag{5}$$

By varying velocities of right and left wheel, there are three special cases for Eq. (4) can be met:

$V_L = V_R$, then R is infinite, robot goes straightforward.
$V_L = -V_R$, then R is zero, robot rotates around its center (the mid-point between two wheels).
If $V_L = 0$ or $V_R = 0$, then R is $\pm\frac{l}{2}$, robot rotates around one wheel (Dudek and Jenkin 2010).

2.2 Forward Kinematics for Differential Drive Robots

According to (4), knowing wheels' velocities V_R and V_L gives the ability to calculate R. From R, robot position (x, y) and angle θ between robot's instantaneous direction and X axis, called ICC—Instantaneous Center of Curvature can be calculated:

$$ICC = [x - R\,\sin(\theta), y + R\,\cos(\theta)] \tag{6}$$

At the moment $t + \delta t$, parameters about robot's position will be:

$$\begin{bmatrix} x' \\ y' \\ \theta' \end{bmatrix} = \begin{bmatrix} \cos(\omega\delta t) & -\sin(\omega\delta t) & 0 \\ \sin(\omega\delta t) & \cos(\omega\delta t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x - ICC_x \\ y - ICC_y \\ \theta \end{bmatrix} + \begin{bmatrix} ICC_x \\ ICC_y \\ \omega\delta t \end{bmatrix} \tag{7}$$

2.3 Inverse Kinematics of a Mobile Robot

According to Dudek and Jenkin, the position of the robot heading in the direction θ with a speed V can be calculated as follows:

$$x(t) = \int_0^t V(t) \cos[\theta(t)] dt \tag{8}$$

$$y(t) = \int_0^t V(t) \sin[\theta(t)] dt \tag{9}$$

However, these equations are used for *holonomic robots*, which have position that can be calculated using integration. Since differential drive robots are *non-holonomic robots*, their motion can be simplified by dividing it into two different steps: moving straight forward and rotating in place. By this way, the rotating angle, speed and time travel, which is known as the inverse kinematics, can be calculated easily.

2.4   Obstacles and the Configuration Space

As normally, robots are assumed to operate in a planar which can be called workspace *W*. As WO_i stands for the obstacles, we can define the *free workspace* as:

$$W_{free} = W \backslash WO_i \tag{10}$$

However, path planning occurs in the set of all robot configurations *Q*, the configuration space. Apparently, an obstacle can be defined as configurations of robot that intersect the obstacle in the workspace *W*:

$$QO_i = \{q | R(q) \cap WO_i \neq \varnothing\} \tag{11}$$

where $R(q)$ is the set of points occupied by the robot.

Hence, the *free configuration space* is:

$$Q_{free} = Q \backslash QO_i \tag{12}$$

2.5   Bug Algorithms

Bug algorithms are the earliest and simplest sensor-based planners for mobile robots. According to Howie Choset (2005), these methods can be divided into 2 main kinds:

– *Bug1* and *Bug2* for robots with isotropic contact sensor (or object detecting sensor).
– *Tangent Bug* for robots with 360° rotating contact sensor (or object detecting sensor).

The Bug 1 algorithm contains three steps:

– Step 1: Head toward goal.
– Step 2: If an obstacle is encountered, circumnavigate it and remember how close the robot get to the goal.
– Step 3: Return to the closest point (by wall following) and continue.

Assuming that the robot is a point with perfect positioning, we can do some analysis about upper/lower bounds on the path length:

$$D, \quad \text{for lower bound} \tag{12.a}$$

$$D + 1,5 \sum_i P_i, \quad \text{for upper bound} \tag{12.b}$$

where D denotes the straight-line distance from start to goal and P_i denotes the perimeter of the ith obstacle.

The Bug 2 algorithm contains three steps:

- Step 1: Head toward goal on the m-line—straight line from starting point to goal.
- Step 2: If an obstacle is encountered, follow its border until the robot encounter the m-line again closer to the goal.
- Step 3: Leave the obstacle and go toward the goal on the m-line.

Assuming that the robot is a point with perfect positioning, we can do some analysis about upper/lower bounds on the path length:

$$D, \text{ for lower bound} \tag{13.a}$$

$$D + 1,5 \sum_i \frac{n_i}{2} p_i, \text{for upper bound} \tag{13.b}$$

where D denotes the straight-line distance from start to goal and P_i denotes the perimeter of the ith obstacle. n_i represents the number of straight-line intersections of the ith obstacle.

The Tangent Bug algorithm shows a better performance compare to Bug 1 and Bug 2 as it determines a shorter path to the goal. However, this algorithm requires a 360° contact sensor.

This algorithm contains two main steps:

- Step 1: Head toward goal on the m-line—straight line from start to goal.
- Step 2: If an obstacle is encountered, use sensor to determine which point on the boundary that can minimize the total path to the goal. This step should be repeated until there is no obstacle detected (Howard Choset et al. 2005).

3   Experiments and Results

These following tasks needed to be archived:

- Connect and synchronize two robots via Bluetooth by using HM-10 Bluetooth modules.
- Path planning for individuals (i.e. kinematic model of robot) that calculates individuals' instantaneous position in global coordination from goal's position message.
- Cooperate between two individuals in path planning task by exchanging each other position data based on the previous task.

3.1 Setting up Electronics Components

At first, HM-10 modules had to be initialized by AT command sets of the manu-facturer: They had been connected to a PC using the FT232RL circuit to establish an UART connection and Hercules application to configure this module for our target of the experiment:

– Set Baud rate: AT + BAUD2 for baud rate of 38,400.
– Set Role: AT + ROLE1 for Master or AT + ROLE2 for Slave.
– Get MAC address (i.e. get specific HM-10 identities): AT + ADDR?

3.2 Connect and Synchronize Two Robots

After setting up each HM-10 module for RP6 Master and RP6 Slave, the connection between these two modules has been configured. This procedure is simple with the help of AT commands that Master firstly requests connection to Slave with the command:

$$\text{``}AT + CONx\text{''}$$

where x is the MAC Address of HM-10 module that received in the previous initialization step.

On successful connection, Master module will reply:

$$\text{``}OK + CONNA\text{''}$$

Otherwise, if the Slave module has already been connected to another device or the MAC Address is invalid, the Master module will respond after 10 s:

$$\text{``}OK + CONNF\text{''}$$

In Fig. 3, two main steps of this task are shown:

– Step 1: Sensor's state is processed by Master robot this information is contin-uously sent via Bluetooth to Slave robot.
– Step 2: Slave robot expresses received information using LEDs lines.

The connection between two modules has been tested by synchronizing envi-ronment data from sensors. There are two IR sensors of the Anti-Collision System (ACS) on the front left and front right of Master robot (Fig. 4). The input values from these sensors will be recognized as obstacle detection, then signals are sent to Slave with the rate of 2 Hz. The Slave will react to these signals by lighting up 2 lines of LEDs corresponding to which side is the obstacle detected by Master robot.
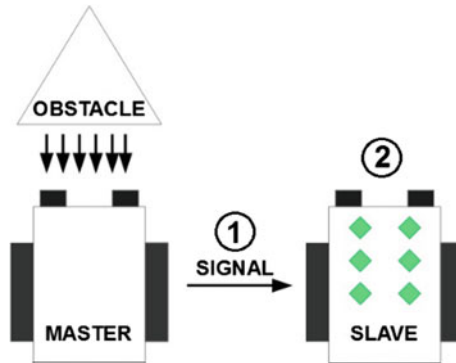
Fig. 3 Connect and synchronize two robots
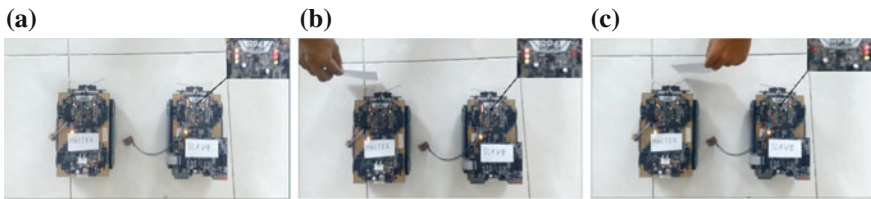
Fig. 4 IR sensors for detection of obstacles on left and right side. **a** No object detected. **b** Object on left side. **c** Object on right side

Consequently, almost everything in this task works well in normal conditions. However, we must concern about:

– RP6's IR sensors: the detecting range of these sensors varies because of many factors such as light intensity, environment temperature or the material of obstacle's surface.
– HM-10 Bluetooth Modules: Connection range between Bluetooth modules also varies from no obstacle workspace in which the maximum range is more than 30 m, to obstacle occupied workspace in which the range decreases dramatically to lower than 10 m.

3.3 Path Planning for Individuals

In this task, we use one robot to simulate basic path planning task. We will introduce the method to navigate our robot and its traveling path in two different coordinate systems: global and local.

3.3.1 Message of Goal's Position

Firstly, we should develop a suitable protocol for each sent data string. The position data should contain:

- x-value and y-value of the goal's position in global coordinate system. Thesevalues should be divided by a comma ",".
- A mark so that the receiver will know where the position string end. We have chosen the dot "." to end the string.

For example, the position of the goal at $(x,y) = (30, -20)$ should be sent by the string:

$$\text{"30, } -20.\text{"}$$

3.3.2 Introduce Global Coordinate System and Local Coordinate System

Global coordinate system: has origin $O_G = \{x_G, y_G\}$,which is unmovable (Fig. 5). Therefore, this system is also unmovable and is used as a very effective tool to locate the instantaneous position of robots during their movement.

Local coordinate system: is a movable coordinate system that has the origin $O_L = \{x_L, y_L\}$ attached to the robot and x-axis will always be robot's front side. This system is used as a mediate tool to navigate the robot.

We define the origin of the unmovable global coordinate $O_G = \{x_G, y_G\}$ that is very effective in locating positions of the robots during their movement, and the origin of movable local coordinate system $O_L = \{x_L, y_L\}$,which is attached to the robot and x-axis will always be perpendicular with the robot's front side.
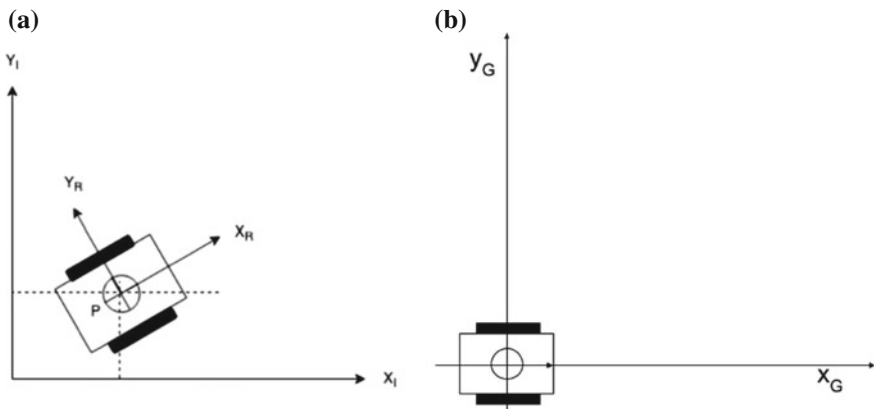
Fig. 5 Relationship between coordinate systems and robot. **a** Local coordinate attached to the robot. **b** Global origin is robot's first position

In the experiment, assuming that initial position is the global origin, after some movements, the instantaneous position of our robot is:

$$^{G}P_{R} = \begin{bmatrix} ^{G}x_{R} \\ ^{G}y_{R} \\ ^{G}\theta_{R} \end{bmatrix} \tag{14}$$

in global coordinate. If we send the next goal position which is

$$^{G}P_{G} = \begin{bmatrix} ^{G}x_{G} \\ ^{G}y_{G} \end{bmatrix} \tag{15}$$

in global coordinate, the robot has to recalculate this goal position to a position in local coordinate

$$^{L}P_{G} = \begin{bmatrix} ^{L}x_{G} \\ ^{L}y_{G} \end{bmatrix} \tag{16}$$

After the robot knows the goal position relative to itself, the traveling path can be calculated directly. As for swarm robots, we will always assume that the global origin is the initial position of one individual.

3.3.3 Process Received Goal Position to Create New Traveling Path

We can divide this step into two smaller tasks:

a. **Calculate goal position in local coordinate system**

After receiving next goal's position in global coordinates, our robot needs to recalculate it into local coordinate. We have a developed function to convert position of a point from global coordinate to local coordinate:

$$^{L}P_{G} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} ^{G}x_{G} - ^{G}x_{R} \\ ^{G}y_{G} - ^{G}y_{R} \end{bmatrix} \tag{17}$$

with:

- $(^{G}x_{G}; {}^{G}y_{G})$ is the goal's position in global coordinate.
- $(^{G}x_{R}; {}^{G}y_{R})$ is the robot's position in global coordinate, corresponding to the position local coordinate's origin.
- θ is the angle between robot and global coordinate's x-axis, corresponding to the angle between two coordinate systems.

The robot decides its traveling path through two steps (Fig. 6). Firstly, it calculates the goal position in local coordinate, this step makes robot easier to decide

(a) (b)
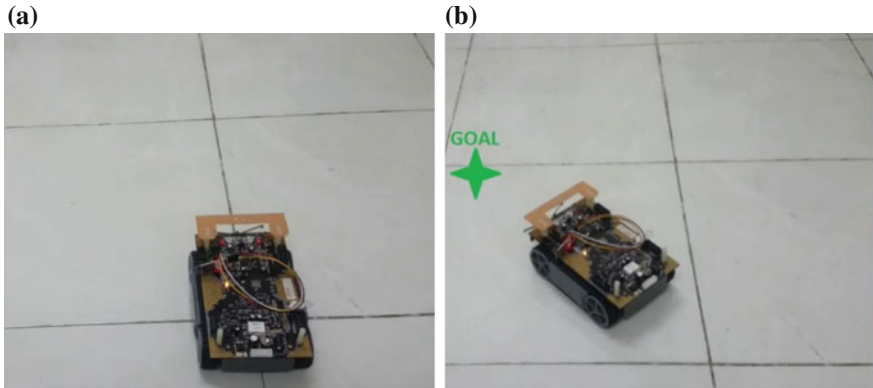
Fig. 6 Task realization. **a** Waiting for goal's position. **b** Traveling to goal

travelling path to goal in its reference coordinate. Secondly, in order to reach the desired position, we need two intermediate steps:

b. **Calculating angle of rotation** using trigonometric functions.
c. **Calculating robot's instantaneous location** to check if goal is reached in order to know when the robot reach goal using the motor encoder attached to the wheels.

3.3.4 Results

In practice, despite the fact that our robot is well programmed and can go to the correct direction, there are still some deviations between the expected goal's position and the final position of the robot. We also calculate the deviation by using the equation:

$$D = \frac{Measured\ Value - Expected\ Value}{Expected\ Value} \times 100\,\%$$

Five tests have been conducted in order to estimate the deviation:

No.	x-coordinate			y-coordinate		
	Expected	Measured	Deviation (%)	Expected	Measured	Deviation (%)
1	20	20	0	20	17	15
2	30	28	6.7	20	23	15
3	40	41	2.5	50	45	10
4	35	32	8.6	35	29	17.1
5	25	25	0	30	26	13.3

As we can see, in most of the cases in test process, the deviation in y-coordinates is up to 17.1 % while it is lower in x-coordinates (8.6 %). This is caused by the power different between robot's left and right motors.

However, the obstacles are not neglectable in the path of movement. One object is considered as an obstacle if and only if it locates on the expected path prevent the robot from going straight forward to the goal. In that case, we need to develop one extra algorithm for our robot to detect and avoid detected obstacle.

3.3.5   Suggestion to Avoid Obstacles Using Bug 2 Algorithm

We can extend this task and make the robot's path to the goal better by adding obstacle avoiding features to it. Up to now, we have studied the Bug algorithms in obstacle avoiding.

The solution for obstacles avoidance using Bug 2 Algorithm is shown in the block diagram of Fig. 7: Follow the obstacle's boundary from the Hit Point; then continue going on the expected path at the Leave Point, where robot's position locates on the expected path again.

Fig. 7 Block diagram of bug 2 algorithm
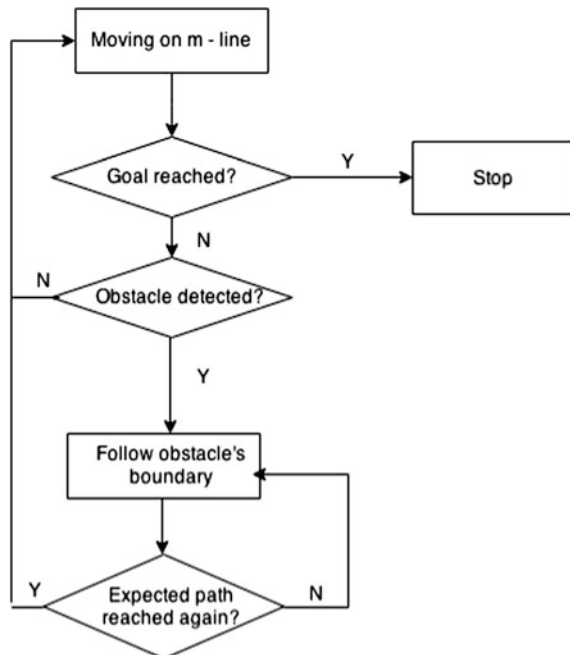
3.4 Cooperate Between Two Swarm Robot Individuals in Path Planning Task

In this scenario for swarm robot cooperation the master robot navigates to a desired goal. Next, the slave robot receives the master's position in order drive to the master robot and support it (Fig. 8).

3.4.1 Task Description

There will be three different parts in this task:

- **Monitor**: This node will send goal's position to Master Robot. In reality, this node will be the user who control the whole system. In our project, a HM-10 Module connected to PC and controlled by Hercules application will present this node.
- **Master Robot**: This robot will first receive goal's position from user or monitor. After that, it will try to reach the goal and from this point, there will be two possible situations: Goal can be reached or obstacle detected.

+ If goal is reached, Master will call assistants (Slaves) to continue with main task (bring object, scan for objects with suitable conditions…).

+ If obstacle detected, Master will ask for help in obstacle avoiding or obstacle destroying (acts that need cooperation between robots).

In both situations, Master should send its instantaneous position to Slave that it request cooperation.

Slave Robot: In general, Slave robot will receive Master's position, then create a plan to reach that goal. The biggest difference in path planning method comparing with Master robot is that at the beginning, Slave's position will not be at the original of global coordinate system. That means Slave robot's local coordinate system will not coincide with global coordinate system at the beginning moment.
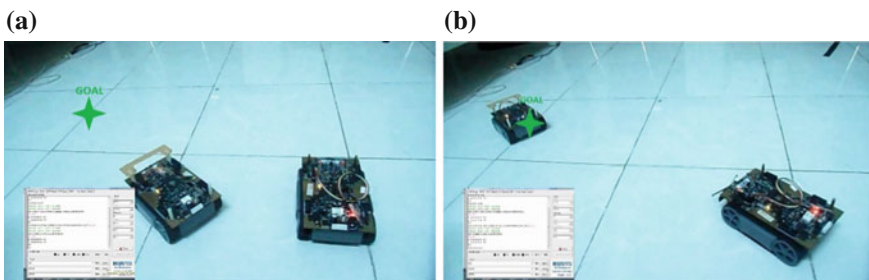
(a)　　　　　　　　　　　**(b)**

Fig. 8 Task realization. **a** Master travels to goal. **b** Goal is reached, slave travels to master

3.4.2 Task Programming

Because Master Robot will communicate with two different nodes, we will set its HM-10 to central, and the others should be peripherals.

– **Monitor**: There are not so much to do with this part, we will only set its HM10's role to *peripheral* using AT command:

$$AT + ROLE0$$

and get its MAC address using AT command:

$$AT + ADDR?$$

This address is stored in order to program Master Robot to initialize connection between them.

– **Master Robot** Comparing to the program in Path planning for individuals' task, there are some tiny changes that help Master Robot to connect to the Monitor and receive the goal's location.

As we have changed Monitor's role to peripheral, and Master's role to central, we have to program Master request connection itself instead of waiting for Monitor to request connect as before:

$$AT + CON\ Monitoraddress$$

After that, the program will jump to a function that contain a while-loop, waiting for goal's position from Monitor. This while-loop should end after our Master receive a complete message of goal's position, which means the character "." is received.

– **Slave Robot**: The program on this robot is almost the same as the last task: process received goal position to create new traveling path. The only different is that instead of having first location that standing at the origin of global coordinate system, Slave robot will have another first location.

4 Conclusion

Swarm robots are getting more and more important thanks to their advantages in getting a better life for human beings. These robots will help us solving problems in wide range with shorter time and cheaper cost. Important tasks that make swarm robot outperform others are in medical field, disaster response and rescue tasks, where time costs by lives.

In this paper we proved that even with simple technology swarm robots helping each other can be developed—Although we could not cover everything about swarm communication and path planning algorithms, it is a good very first step for our laboratories to research deeper in this field.

Concerning about our swarm robot system RP6, there are some improvements that can be done in order to make the system more reliable and can work with more features: (a) Replace Bluetooth communication with another, so that the individuals can connect with more robots at the same time; (b) Adding some position sensors such as a digital compass or a GPS module that robot can get its instantaneous position more precisely.

References

Canudas de Wit, C., Siciliano, B., & Bastin, G. (Eds.). (1996). *Theory of robot control*. London: Springer.

Choset, H., et al. (2005). *Principles of robot motion: Theory, algorithms, and implementations*. The MIT Press.

Dudek, G., & Jenkin, M. (2010). *Computational principles of mobile robotics*. Cambridge University Press.

Michalek, R., & Tarantello, G. (1988). Subharmonic solutions with prescribed minimal period for nonautonomous Hamiltonian systems. *Journal of Dynamics and Differential Equations, 72*, 28–55.

Rabinowitz, P. H. (1980). On subharmonic solutions of a Hamiltonian system. *Communications on Pure Applied Mathematics, 33*, 609–633.

Siegward, R., & Nourbakhsh, I. R. (2004). *Introduction to autonomous mobile robot*. The MIT Press.

Tarantello, G. (1988). *Subharmonic solutions for Hamiltonian systems via a $\mathbb{Z}_p$ pseudoindex theory*. Annali di Matematica Pura. Reprint.

Ying, T., & Zhong-yang, Z. (2013). *Research advance in swarm robotics*. Beijing: Defence Technology.

Indoor Localization for Swarm Robotics with Communication Metrics Without Initial Position Information

Türker Türkoral, Özgür Tamer, Suat Yetiş and Levent Çetin

Abstract Swarm robotics is an emerging research field with many scientific and commercial application areas. Swarm robotic systems are composed of simple robots cooperatively accomplishing the given task. The cooperation and the organization of the robots require the location information of the robots; so that each robot will be able to achieve the task it is responsible. The main of this work is an indoor positioning for swarm robotic applications using standard metrics for Bluetooth or Wi Fi communication infrastructures. The main work includes fusion of several position estimation methods by setting a proper weight to each method; this paper presents the work on estimating the location by using TDoA and RSSI parameters.

Keywords Swarm robotics · Positioning · Localization

1 Introduction

Swarm robotics has been paid much attention in recent years. A swarm robotic system, inspired by social insects like ants, includes a large number of robots which are coordinated in a distributed and decentralized way. Members of the team are simple and identical robots and achieve a complex task by employing local and simple rules (Navarro and Matia 2013).

T. Türkoral (✉) · S. Yetiş
Mechatronics Engineering Department, Dokuz Eylül University, İzmir, Turkey
e-mail: turker.turkoral@nucleo.com.tr

Ö. Tamer
Electricaland Electronics Department, Dokuz Eylül University, İzmir, Turkey

L. Çetin
Mechatronics Engineering Department, İzmir Katip Çelebi University, İzmir, Turkey

Each member of a swarm robotic team is assigned a task due to some factors. Location of the member is one of the important factors for assigning proper tasks to the robots (Pugh et al. 2009, p. 151–162). For example; if the mission includes a task as travelling through a path and pushing a button at the end of that path, this task is assigned to the closest robot to the path.

We may classify localization systems as absolute and relative positioning systems. Absolute localization schemes depend on a global coordinate framework and needs a master reference point to localize the robots. For example an overhead camera based robot localization system is an example to an absolute localization scheme (Krajnik et al. 2013). Most of the research for swarm robot localization is on absolute localization systems.

Relative positioning schemes only depend on the relative positions of the robots with respect to each other. Even though absolute positioning provides a more precise and valuable information, for some tasks only relative position information might be needed to fulfill the task.

If we know the distance between each member of a swarm robotic team we may use the triangulation technique for a simple relative positioning scheme. Triangulation is simply estimating the relative position of each robot by evaluating the intersection points of distance circles. To get position information by using the triangulation, we first need to get distance information between robots.

The basis of positioning with standard communication metrics is triangulation (Liu et al. 2007, pp. 1067–1080). Triangulation is positioning the location of a point by using the distance information from three other initially known points. In the work of Liu et al., triangulation is emphasized as an indoor positioning technique (Liu et al. 2007, pp. 1067–1080). However, to apply this method, the locations of three reference points have to be given initially. Laaraiedh et al., utilized RSSI (Received Signal Strength Indicator) and TDoA (Time Difference of Arrival) for positioning and compared the results (Laaraiedh et al. 2011). In addition, the usage of both these weighted metrics is also analyzed. Furthermore, Perkins compared RSSI and TDoA for estimating distances of observation points (Perkins et al. 2011, pp. 1984–1987). In the paper of Kovacs et al., Bluetooth is preferred for the communication of autonomous swarm robots (Kovacs et al. 2011, pp. 530–542). In Winfield and Holland's work, WLAN is used for distributed robotic systems (Winfield and Holland 2010, pp. 597–607). Galvan-Tejada et al., used Bluetooth, WLAN as the communication system and results are compared (Galvan-Tejada et al. 2012, pp. 101–108).

The proposed algorithm in this paper differs from triangulation with using no initial conditions. The locations of the points are to be found with using the communication metrics (TDoA and RSSI) provided by the wireless communication protocols (Liu et al. 2007; Perkins et al. 2011; Hara and Anzai 2008). The metrics TDoA and RSSI are provided by WLAN and Bluetooth communication infrastructures. In this work, positioning process will be done by using TDoA and RSSI metrics individually, while the weighted mean of these metrics are also taken into account. In the resulting system, wireless communication protocols will be used separately, however, both WLAN and Bluetooth results are used together according

to the environmental factors and the application area by weighting each of the metrics taken from each of the protocols.

2 Methodology

The work presented in this paper depends on triangulation by estimating the distance with respect to three or more reference points. The reference locations are neither fixed nor initially known so the estimated location is a relative location with respect to the reference locations.

Triangulation is estimating the location by using the distance of three reference points and is a widely used method. Distance with a single reference point helps us to estimate the location of an object on a circle with a radius equal to the distance, but when we know the distance of the object with respect to two reference points our location is more precise; two intersection points of the two circles is our probable location. The distance information with a third reference points intersects with only one of the previous intersection points which theoretically is the exact estimation of the location of the object. This basic approach of triangulation is presented in Fig. 1. However to make the estimation more precise we may use more reference points.

2.1 Positioning with Metrics

Triangulation based location estimation will be performed by the use of RSSI and TDoA information provided by the utilized wireless communication infrastructures; WLAN and Bluetooth.

Fig. 1 Positioning with triangulation with respect to two points
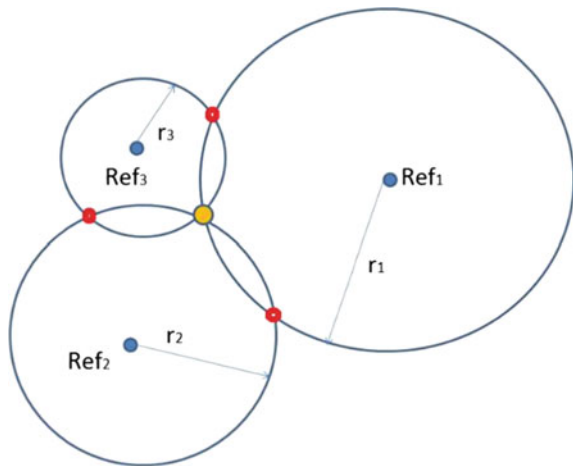
RSSI, as mentioned before, represents the strength of the arrived signal. The basic conversion of this metric to distance data can be done with Friis equation (Eq. 1).

$$\frac{Pr}{Pt} = \left(\frac{\lambda}{4\pi R}\right)^2 D_t(\theta_t\varphi_t)D_r(\theta_r\varphi_r) \tag{1}$$

In Friis equation, Dt, Dr, Pt, Pr, λ, R represent the transmitter antenna gain, receiver antenna gain, the strength of the signal sent by transmitter antenna, the strength of the signal arrived to receiver antenna (provided by RSSI metric), wavelength at center frequency and the distance between the receiver and the transmitter (the distance is to be found) respectively. Furthermore, with the measurements in different environments, different distances, this received data will be able to calibrated and converted into distance data.

RSSI metric can be transformed into distance data with the use of Friis equation and the table that will be experimentally derived.

Even though the Friis equation is valid only in free space and there exists other signal models for indoor propagation environments, it is used as an initial choice because of its simplicity. Further studies will include adapting indoor signal models.

TDoA represents the received signal time difference of arrival, can be transformed into distance data with calculating the time stamp differences between the sent signal and the received signal. If the distance between 2 robots is R, the signal covers a distance of 2R when transmitted from one point to other and come back to the initial transmitter. Considering the speed of electromagnetic wave equals to speed of light in air, TDoA-distance data conversion equation derived as Eq. 2.

$$2R = TDoA * c \tag{2}$$

$$c = speed\ of\ light \tag{3}$$

The signal travels the distance between the transmitter and the receiver twice while a minor time delay may also occur between the arrival of the signal to the receiver and the retransmission of it from the receiver. To minimize the error in the proposed system, with adding the mentioned delay to the formula, Eq. 4 is revealed.

$$2R = (TDoA - delay) * c \tag{4}$$

By the measurements, between any of the 2 points taken one by one, the distances will be revealed; hence, positions will be determined over the distances.

Proposed algorithm uses no initial conditions, by that, the coordinate plane will be defined on the first 2 robots' locations ('*a*' and '*b*' points). While measuring, 1st robot's location is always set at (0, 0), and the 2nd robot is always set at (0, L_{ab}), regarding to the distance from the 1st robot's location, '*a*' (L_{ab}) (Fig. 2a).
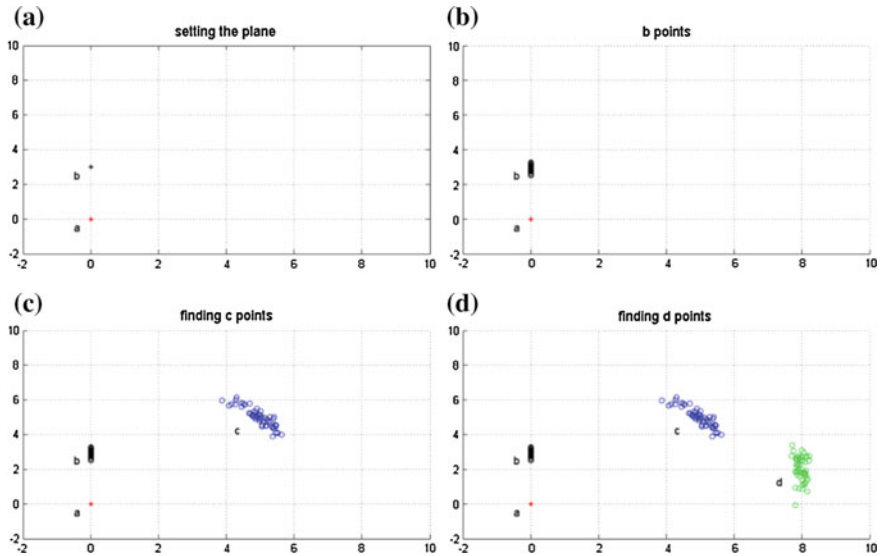
Fig. 2 Estimation of Robot locations

Noise in the medium, reflection and other environmental effects decrease the quality of the received signal, hence the measurements may be faulty. Therefore, the positioning process will not be made by one measurement set, the measurements will be taken more than one (n times), positions will be calculated over the average of these measurement sets. The aim is to minimize the error in the system. In that case, 'a' is again set (0, 0), 'b' (taken by the measurements, L_{ab}) are set on Y axis. Then the average of these robot location estimates on both X and Y axis (for point 'b', all X axis measurements are 0) will be taken and the calculated, average 'b' point will be revealed (Fig. 2b).

The position of the 3rd robot will be obtained by using the locations of first and second robots. The location of the third robot will be derived by using the measurement sets taken n times between robots 1 and 3, and 2 and 3 rather than calculating from robot 1 location and the average value of robot 2 location as presented in Fig. 2c.

By using the sample sets, when calculating the latter points, main aim is not to transfer the positioning error to the next estimation.

When calculating the location of robot at 'c', 2 possibilities occur. Choosing the right position of the robot by using only the distance measurements is not possible. So, the position of the third robot, 'c', will be chosen as its X plane value is positive. All calculations are made according to this assumption.

After the 3rd robot is located, the same process is applied for all other robots. The location of the 4th robot for example 'd', alike the calculations of robot at 'c', will be made by using the locations of the first 2 robots, 'a' and 'b'.

2.2   Realization Number and the Noise Effect

The positions of the points are calculated over a set of measurements taken n times. Both noise and this number n effects the system performance directly. Decreasing the realization number n increases the possibility of error. Also the noise in the medium affects the base metrics during the measurements. To avoid this, before the measuring process, the noise in the medium will be measured several times in different time intervals, then the average noise value will be calculated over these noise measurements and it is planned to dissociate this resulting noise value from all the metric measurements. Also with the increment of the realization number n, the amount of error is thought to be decreased.

3   Results

The simulation of the proposed algorithm has been carried out in MATLAB. As a result of the simulation, the position of each robot location has been calculated at various distances according to their actual position by depending on the sampling rates and the signal to noise ratio. Each simulation result has been observed according to the actual position, average of calculated position, the point which is farthest relative to the actual position, the most corrupted result by the noise and the variance of the error of the samples.

The simulation results which has been calculated under different SNR (Signal to Noise Ratio) values and sampling rates are given in Figs. 3 and 4. Also, the effect of
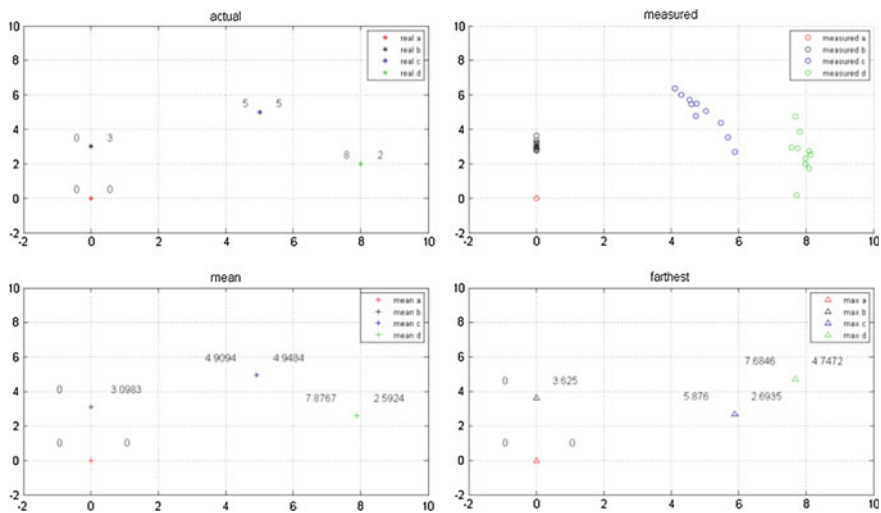
Fig. 3   Simulation result for RNUM = 10, SNR = 10
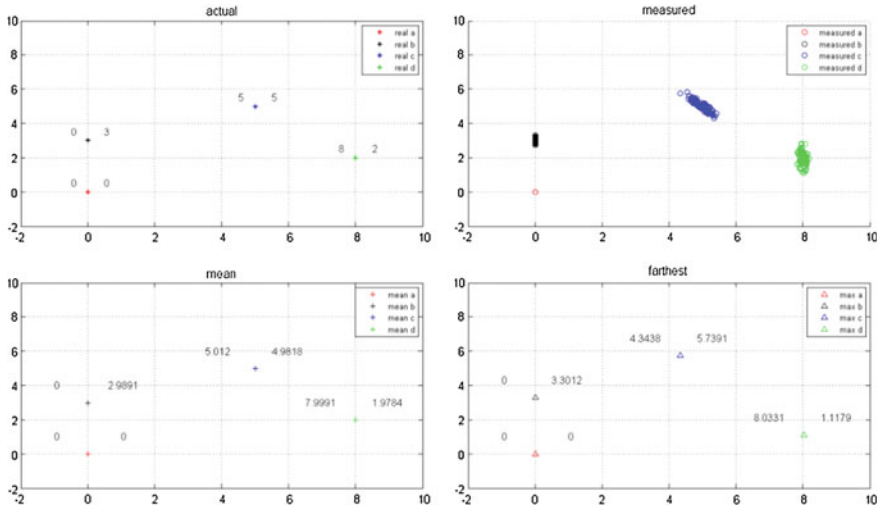
Fig. 4 Simulation result for RNUM = 100, SNR = 20

Table 1 Numerical analysis of noise

rnum snr	Points	Actual position		Average position		Farthest position		Variance
		x pos.	y pos.	x pos.	y pos.	x pos.	y pos.	
rnum	a	0	0	0.0000	0.0000	0.0000	0.0000	0.0000
10	b	0	3	0.0000	3.0983	0.0000	3.6250	0.0324
snr	c	5	5	4.9094	4.9484	5.8760	2.6935	0.5149
10	d	8	2	7.8767	2.5924	7.6846	4.7472	0.7582
rnum	a	0	0	0.0000	0.0000	0.0000	0.0000	0.0000
20	b	0	3	0.0000	3.0274	0.0000	2.2953	0.0396
snr	c	5	5	4.7323	4.9901	3.3421	7.0271	0.6739
10	d	8	2	7.8781	1.8697	7.2482	−0.6199	0.4331
rnum	a	0	0	0.0000	0.0000	0.0000	0.0000	0.0000
30	b	0	3	0.0000	3.0803	0.0000	3.9988	0.0493
snr	c	5	5	4.9484	4.8518	2.4152	6.6814	0.4719
10	d	8	2	7.9457	1.6429	7.9060	−1.2232	0.5424
rnum	a	0	0	0.0000	0.0000	0.0000	0.0000	0.0000
50	b	0	3	0.0000	3.0163	0.0000	2.3683	0.0323
snr	c	5	5	4.8416	5.0423	2.9402	6.8588	0.3683
10	d	8	2	7.9086	2.1524	7.1755	5.1487	0.4669
rnum	a	0	0	0.0000	0.0000	0.0000	0.0000	0.0000
100	b	0	3	0.0000	3.0025	0.0000	4.0961	0.0506

Table 1 (continued)

rnum snr	Points	Actual position		Average position		Farthest position		Variance
		x pos.	y pos.	x pos.	y pos.	x pos.	y pos.	
snr	c	5	5	4.8315	5.1121	2.9421	7.1852	0.4382
10	d	8	2	7.8775	2.0682	6.8397	5.6429	0.7959
rnum	a	0	0	0.0000	0.0000	0.0000	0.0000	0.0000
100	b	0	3	0.0000	2.9891	0.0000	3.3012	0.0043
snr	c	5	5	5.0120	4.9818	4.3438	5.7391	0.0414
20	d	8	2	7.9991	1.9784	8.0331	1.1179	0.0472

the noise and the sampling rate to the system are calculated numerically. The numerical results can be seen in Table 1 (RNUM = Realization Number 'n').

4 Conclusion and Future Work

According to the simulation results, when number of samples and SNR is high, rate of reaching the actual position is also high as expected.

On the other hand, accuracy of the location of the observation point can be increased by employing more estimated distance as a parameter to the algorithm. In other words, despite using L_{ad} and L_{bd} distances only for estimating the position of the 4th point, L_{cd} might also be used to improve the performance of the algorithm. The number of positions to be used in the algorithm is an optimization issue, since increasing the number of distances as a parameter also increases the operational load of the algorithm and might also decrease performance.

As introduced previously, the system is a kind relative positioning approach, as a future work, after determining the position of every robot on the relative coordinate system, it is planned to transform it to an absolute coordinate system by using rotation, transformation or Denavit-Hartenberg matrix (Siciliano et al. 2010).

References

Galvan-Tejada, I., Sandoval, E. I., &Brena, R. (2012). Wifibluetooth based combined positioning algorithm. *Procedia Engineering*, *35*, 101–108.

Hara, S., &Anzai, D. (2008). Experimental performance comparison of RSSI- and TDOA-based Location Estimation Methods. In *Vehicular Technology Conference*, *VTC 2008*, *IEEE* (pp. 2651–2655).

Kovacs, T., Pasztor, A., & Istenes, Z. (2011). A multi-robot exploration algorithm based on a static Bluetooth communication chain. *Robotics and Autonomous Systems*, *59*(7–8), 530–542.

Krajnik, T., Nitsche, M., Faigl, J., Duckett, T., Mejail, M., & Preucil, L. (2013). External localization system for mobile robotics. In *2013 16th International Conference on Advanced Robotics (ICAR) (pp. 1–6), 25-29 Nov. 2013*.

Laaraiedh, M., Yu, L., Avrillon, S., & Uguen B. (2011) Comparison of hybrid localization schemes using RSSI, TOA, and TDOA. In *Wireless Conference 2011-Sustainable Wireless Technologies (European Wireless), 11th European VDE* (pp. 1–5).

Liu, H., Darabi, H., Banerjee, P., & Liu, J. (2007). Survey of wireless indoor positioning techniques and systems. *IEEE Transactions on Systems, Man, and Cybernetics Part C: Applications and Reviews, 37*(6), 1067–1080.

Navarro, I., Matía, F. (2013). An introduction to swarm robotics. *ISRN Robotics,* Article ID 608164, 10 p.

Perkins, C., Lei, L., Kuhlman, M., Lee, T. H., Gateau, G., Bergbreiter, S., et al. (2011). Distance Sensing for mini-robots: RSSI vs. TDOA. In *Circuits and Systems (ISCAS), 2011 IEEE International Symposium on* (pp. 1984–1987), 15–18 May 2011

Pugh, J., Raemy, X., Favre, C., Falconi, R., & Martinoli, A. (2009). A fast onboard relative positioning module for multirobot systems. *Mechatronics, IEEE/ASME Transactions on, 14*(2), 151–162.

Siciliano, B., Sciavicco, L., Villani, L., & Oriolo, G. (2010). Robotics; modelling, planning and control. *Procedia Engineering*.

Winfield, A. F., & Holland, O. E. (2010). The application of wireless local area network technology to the control of mobile robots. *Microprocessors and Microsystems*, *23*(10), 597–607.

Multi-objective Optimization of a Parallel Fine-tuning Manipulator for Segment Assembly Robots in Shield Tunneling Machines

Guohua Cui, Haidong Zhou, Yanwei Zhang and Haibin Zhou

Abstract A new 6-DOF serial/parallel hybrid segment assembly robot that includes a 3-DOF redundant parallel orientation fine-tuning manipulator is developed. Redundancy can, in general, improve the ability and performance of parallel manipulators by implementing the redundant degree of freedom to optimize the objective function. A multi-objective optimization model of the proposed manipulator is likewise established with consideration of kinematic and dynamic dexterity, load-bearing capacity, stiffness, and actuating force equilibrium. To accomplish optimization, a new multi-objective optimization method is proposed using the application Isight, which integrates various engineering softwares. Results indicate that this method can automatically accomplish the calculation of the kinematic and dynamic performance index, modeling, and solving using Isight. Thus, decision makers can select the proper optimal solution based on the obtained Pareto plot and Pareto frontier. Finally, optimizing its design variables considerably improves the performance of the proposed parallel manipulator.

Keywords Segment assembly robot · Multi-objective optimization · Parallel fine-tuning manipulator · Pareto frontier

1 Introduction

Parallel robot manipulators have many applications in engineering because of their advantages, which include high rigidity, high accuracy, and good load-bearing capacity. Tunnel segment assembly robots, which are a key sub-device in shield

G. Cui · H. Zhou · Y. Zhang (✉)
College of Equipment Manufacturing, Hebei University of Engineering, Handan, China
e-mail: zhangyanwei@hebeu.edu.cn

H. Zhou
Zhejiang Provincial Special Equipment Inspection and Research Institute, Hangzhou, China

tunneling machines, can automatically assemble segments. An increasing number of parallel manipulators are used in the segment assembly robots to improve their efficiency and precision. According to Tanaka (1995) and Kosuge et al. (1996), a parallel link robot called Stewart platform is used for the assembly of the segments of a shield tunnel excavation system. Zyada et al. (2003) introduced the experimental results for the assembly of the multi-directional segments of a shield tunnel excavation using a hydraulic Cough-Stewart Platform. A 3-SPS-1-S parallel manipulator used as an orientation fine-tuning manipulator for segment assembly robots in shield tunneling machines was put forward by Wu et al. (2011). Shi et al. (2009) developed a 6-DOF shield tunneling machine controlled by electro-hydraulic proportional systems. Cui et al. (2010) used a 6-DOF serial and parallel manipulator as the mechanism of a segment erector, as well as in performing kinematics analysis.

The present research focuses on a novel 4-SPS-S parallel manipulator, which can be used as an orientation fine-tuning manipulator for segment assembly robots. In addition, we will conduct multi-objective dimension optimization. Parallel manipulator has multiple performances, and thus, its dimension synthesis constitutes a multi-objective optimization problem that has to take into consideration several kinematic and dynamic performances based on the characteristics of the structure and the applications. Numerous studies have been conducted on the dimensional synthesis of parallel manipulator. Gao et al. (2010) introduced the Pareto multi-objective optimization theory into the performance optimization of parallel manipulator. To optimize the parallel manipulator 6-RSS, Liu et al. (2012) used the improved multi-objective particle swarm optimization algorithm based on stochastic ranking. Chen et al. (2008) optimized a new type of parallel manipulator 2-PRRR with two degrees of freedom by genetic algorithm. In general, the afore-mentioned methods depend on the complex mathematical deduction process, which includes building a mathematical model and optimizing the model through an optimization algorithm. At present, Isight is used in numerous fields, such as aviation and automobile, to complete large-scale optimization problems.

The performance indexes, including kinematic dexterity, dynamic dexterity, load-bearing capacity, stiffness, and actuating force equilibrium, were defined by considering the functional requirements of segment assembly robot. In addition, we established a multi-objective optimization model of the parallel manipulator. A multi-discipline and multi-objective optimization design platform was built via Isight to integrate several application softwares. This new method, which allows for the accomplishment of the multi-objective optimization for parallel manipulator, can realize the design processes, such as automatic modeling, automatic analysis, automatic simulation, and automatic optimization. This method can then extract, analyze, and manage the design results. Overall, the kinematic and dynamic performance of the proposed parallel manipulator is considerably improved owing to the multi-objective optimization process.

2   Description of Architectures

This section conducts analysis of shield tunneling machine and describes the architecture of the 3-degree of freedom parallel manipulator. Inverse kinematic problem, as the basic kinematics issue of parallel/hybrid manipulator, is also addressed. Simultaneously, forward direct kinematics of the passive leg is also derived in detail.

2.1   *Manipulator of a Segment Assembly Robot*

As shown in Fig. 1, the shield in a typical slurry shield tunneling machine is pushed forward by the thrust from a group of hydraulic actuating cylinders. As one of the key sub-devices in a shield tunneling machine, a segment assembly robot is employed to erect segments and form the segment lining automatically. Given that a concrete segment is typically heavy, achieving its precise required pose can be difficult. As illustrated in Fig. 2, the segment must precisely reach its required position and orientation; that is, the segment assembly robots must accurately achieve six degrees of freedom (Lifting, Revolving, Sliding, Roll, Yaw, and Pitch).

As shown in Fig. 3a, a 6-DOF series-parallel manipulator is developed after considering the characteristics of serial and parallel manipulators, the operating requirements for high assembly accuracy in heavy segments, and the limited space in shield tunneling machines. The segment assembly robot has dual cylinders that lift the segments to a vertical direction. A sliding cylinder pulls or pushes an end-effector to a horizontal direction. Moreover, hydraulic motor drives the gear ring, thereby enabling the whole robot to reach every position in the circumferential direction. Near the bottom of the robot is an orientation fine-tuning manipulator that

Fig. 1   A typical shield tunneling machine
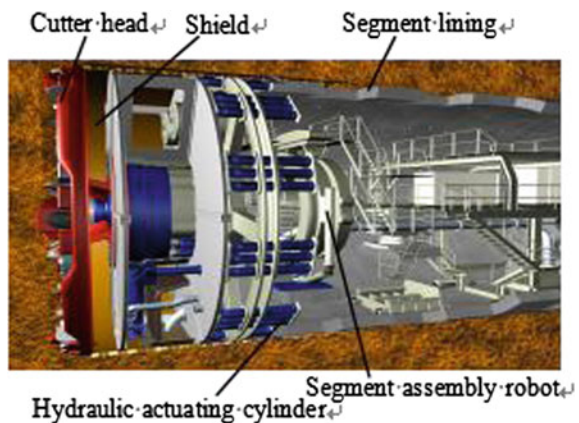
Fig. 2 Process of a segment
assembly
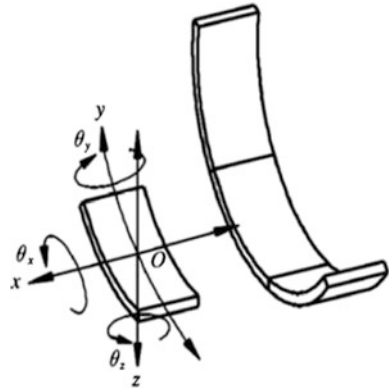
can perform the roll, pitch, and yaw motions that will maintain the segment in the
expected attitude.

2.2 Model of the Parallel Orientation Fine-tuning Manipulator

Figure 3b illustrates the kinematic architecture of a novel 3-DOF parallel orienta-
tion fine-tuning manipulator. The moving platform is connected to the fixed plat-
form by four identical SPS legs and one passive S joint, hence the manipulator is

Fig. 3 Model of segment assembly robot. **a** Kinematic structure of a segment assembly robot.
b Kinematic structure of parallel fine-tuning manipulator

Fig. 4 Kinematic schematic diagram of 4-SPS-S manipulator. **a** Main view. **b** Top view

named 4-SPS-1-S parallel manipulator. The four P joints are actuated. The moving platform can rotate only along each axis that passes through the center of the passive S joint. Wu et al. (2007) pointed out that due to the extra degrees, kine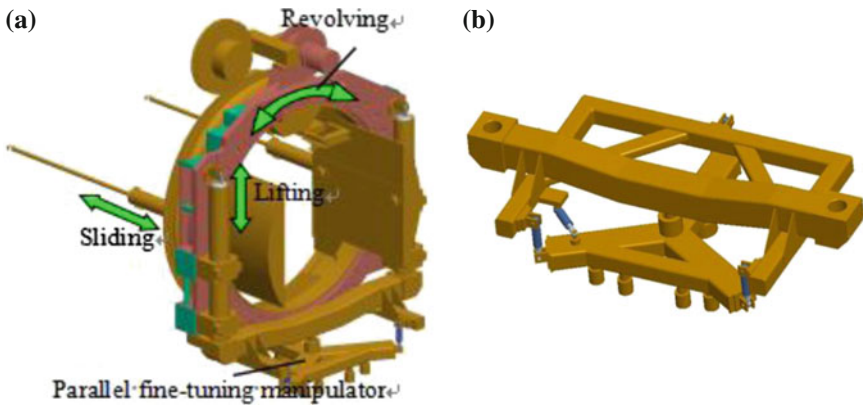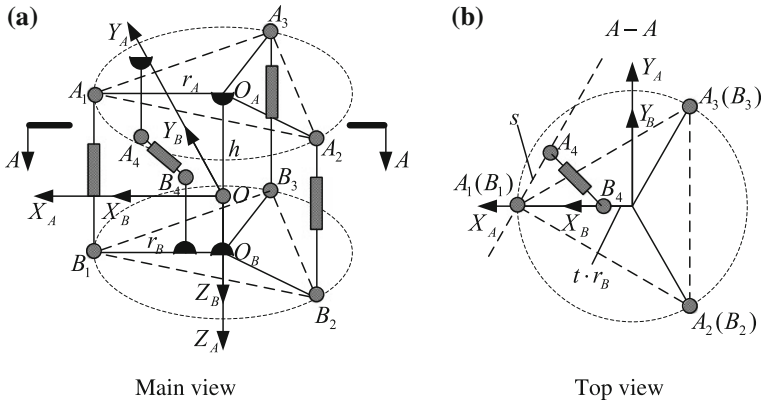matically redundant parallel manipulator is inherently capable of more dexterous manipulation, and can not only execute the original output task but also perform additional tasks such as singularity elimination, workspace enlargement, dexterity improvement, obstacle avoidance, force transmission optimization and unexpected impact compensation. The center is defined as the rotation center of the moving platform. As shown in Fig. 4, A_i and B_i ($i = 1, 2, 3, 4$) are the centers of the S joints that are attached to the base and moving platform, respectively. The passive leg has a spherical joint, marked as O, which is fixed to the base and moving platform at their centers O_A and O_B. The origins of the fixed coordinate system and moving coordinate system, i.e., $O-X_A Y_A Z_A$ and $O-X_B Y_B Z_B$, lie at point O. These systems are coincident at the initial position. The initial position indicates that the configuration $A_i B_i$ ($i = 1, 2, 3$) is perpendicular to the moving platform and the base. At this pose, the distance between the two platforms is H, here, $H = 310$ mm. The orientations of the axis are shown in Fig. 3. The distance between O and the base is h. $\triangle A_1 A_2 A_3$ and $\triangle B_1 B_2 B_3$ are equilateral triangles, with circumcircles' radius r_A and r_B, respectively. Here, $r_A = r_B = 1050$ mm. A_4 is in the plane parallel to line $O_A A_3$ and perpendicular to the base. The distance between A_4 and base is h_A, here $h_A = 155$ mm. The projector distance on the base of A_1 and A_4 is s. B_4 is in the plane $X_B O Z_B$, and its distance to Z_B-axis is ($t \cdot r_B$). The distance from B_4 to the moving platform is h_B, here, $h_B = 155$ mm.

2.3 Inverse Kinematics and Velocity Analysis

The 4-SPS-S parallel manipulator has three rotational degrees of freedom. This manipulator is a redundant drive parallel manipulator, which can decrease or eliminate singularity, increase manipulability, increase payload and acceleration of the mobile platform, further improve the efficiency and reliability by eliminating actuator singularity or force-unconstrained configurations, and improve the repeatability by controlling the direction of internal forces to reduce the effects of joint backlashes. Actuation redundancy can improve the transmission properties and yield an optimal load distribution among actuators by increasing the homogeneity of force transmission and manipulator stiffness. It can also optimize the actuator torque and make up the actuator fault.

On the basis of the combined transformation RPY, the rotation matrix between moving and fixed coordinate system can be expressed as follows:

$$
{}_B^A\mathbf{R} = \mathbf{RPY}(\phi, \theta, \psi) = \begin{bmatrix} c\phi c\theta & c\phi s\theta s\psi - s\phi c\psi & c\phi s\theta c\psi + s\phi s\psi \\ s\phi c\theta & s\phi s\theta s\psi + c\phi c\psi & s\phi s\theta c\psi - c\phi s\psi \\ -s\theta & c\theta s\psi & c\theta c\psi \end{bmatrix} \quad (1)
$$

where ${}_B^A\mathbf{R}$ is the rotation transformation matrix. ψ, θ, and ϕ are the angles of yaw, pitch, and roll of the moving platform, respectively. Here, $\sin\phi$ and $\cos\phi$ are written as $s\phi$ and $c\phi$.

The output parameters ψ, θ, and ϕ are chosen as the generalized coordinates made as $\mathbf{q} = [\psi \quad \theta \quad \varphi]^T$. Thus, generalized velocity can be expressed as $\dot{\mathbf{q}} = [\dot{\psi} \quad \dot{\theta} \quad \dot{\varphi}]^T$. Using the coordinate transformation theory, the following equation is obtained:

$$
\mathbf{l_i} = {}_B^A\mathbf{R} \cdot {}^B\mathbf{p_{Bi}} - {}^A\mathbf{p_{Ai}} \quad (i = 1, 2, 3, 4) \quad (2)
$$

where $\mathbf{l_i}$ is the vector of limbs in the fixed coordinate system, ${}^A\mathbf{p_{Ai}}$ is a vector of A_i in fixed coordinate system $O-X_AY_AZ_A$, ${}^B\mathbf{p_{Bi}}$ is the vector of B_i in the moving coordinate system $O-X_BY_BZ_B$, that is,

$$
\begin{cases} {}^A\mathbf{p_{A1}} = \begin{bmatrix} r_A & 0 & -h \end{bmatrix}^T \\ {}^A\mathbf{p_{A2}} = \begin{bmatrix} -r_A/2 & -\sqrt{3}r_A/2 & -h \end{bmatrix}^T \\ {}^A\mathbf{p_{A3}} = \begin{bmatrix} -r_A/2 & \sqrt{3}r_A/2 & -h \end{bmatrix}^T \\ {}^A\mathbf{p_{A4}} = \begin{bmatrix} r_A - s/2 & \sqrt{3}s/2 & h_A - h \end{bmatrix}^T \end{cases} \quad (3a)
$$

$$
\begin{cases} {}^B\mathbf{p_{B1}} = \begin{bmatrix} r_B & 0 & H - h \end{bmatrix}^T \\ {}^B\mathbf{p_{B2}} = \begin{bmatrix} -r_B/2 & -\sqrt{3}r_B/2 & H - h \end{bmatrix}^T \\ {}^B\mathbf{p_{B3}} = \begin{bmatrix} -r_B/2 & \sqrt{3}r_B/2 & H - h \end{bmatrix}^T \\ {}^B\mathbf{p_{B4}} = \begin{bmatrix} r_B \cdot t & 0 & H - h - h_B \end{bmatrix}^T \end{cases} \quad (3b)
$$

$\mathbf{l_i}$ can be obtained from Eq. (2), whereas ψ, θ, and ϕ are known. The square of the length of the chain can be derived as

$$l_i^2 = \|\mathbf{l_i}\|^2 = \left\|{}^{\mathbf{A}}\mathbf{p_{Bi}} - {}^{\mathbf{A}}\mathbf{p_{Ai}}\right\|^2 = \Theta_i(\psi, \theta, \phi) \tag{4}$$

where l_i is the length of the active legs. The following equation can be obtained by differentiating Eq. (4) with respect to time:

$$\mathbf{A\dot{l}} = \mathbf{B\dot{q}} \tag{5}$$

Thus,

$$\mathbf{\dot{l}} = \mathbf{A}^{-1}\mathbf{B\dot{q}} \tag{6}$$

where $\mathbf{J} = \mathbf{A}^{-1}\mathbf{B}$ is the Jacobian matrix,

$$\mathbf{A} = \begin{bmatrix} 2l_1 & 0 & 0 & 0 \\ 0 & 2l_2 & 0 & 0 \\ 0 & 0 & 2l_3 & 0 \\ 0 & 0 & 0 & 2l_4 \end{bmatrix} \tag{7}$$

$$\mathbf{B} = \begin{bmatrix} \frac{\partial \Theta_1(\psi,\theta,\phi)}{\partial \psi} & \frac{\partial \Theta_1(\psi,\theta,\phi)}{\partial \theta} & \frac{\partial \Theta_1(\psi,\theta,\phi)}{\partial \phi} \\ \frac{\partial \Theta_2(\psi,\theta,\phi)}{\partial \psi} & \frac{\partial \Theta_2(\psi,\theta,\phi)}{\partial \theta} & \frac{\partial \Theta_2(\psi,\theta,\phi)}{\partial \phi} \\ \frac{\partial \Theta_3(\psi,\theta,\phi)}{\partial \psi} & \frac{\partial \Theta_3(\psi,\theta,\phi)}{\partial \theta} & \frac{\partial \Theta_3(\psi,\theta,\phi)}{\partial \phi} \\ \frac{\partial \Theta_4(\psi,\theta,\phi)}{\partial \psi} & \frac{\partial \Theta_4(\psi,\theta,\phi)}{\partial \theta} & \frac{\partial \Theta_4(\psi,\theta,\phi)}{\partial \phi} \end{bmatrix}$$

Suppose the segment center of gravity is G. Given that the segment is fixed with the moving platform during the work of the fine-tuning manipulator, the position vector of G with respect to the moving coordinate system ${}^{\mathbf{B}}\mathbf{p_G}$ is solvable. The following can be obtained based on rotational transformation:

$$^{\mathbf{A}}\mathbf{p_G} = {}^{\mathbf{A}}_{\mathbf{B}}\mathbf{R} \cdot {}^{\mathbf{B}}\mathbf{p_G} = \begin{bmatrix} X(\psi, \theta, \varphi) & Y(\psi, \theta, \varphi) & Z(\psi, \theta, \varphi) \end{bmatrix}^T \tag{8}$$

where ${}^{\mathbf{A}}\mathbf{p_G}$ is the position vector of G with respect to the fixed coordinate system, and is a function of ψ, θ, and ϕ.

By differentiating Eq. (8) with respect to time, the linear velocity of G can be expressed as

$$\mathbf{v} = \mathbf{J_v} \cdot \mathbf{\dot{q}} \tag{9}$$

where

$$\mathbf{v} = \begin{bmatrix} \dot{X} & \dot{Y} & \dot{Z} \end{bmatrix}^T, \mathbf{J_v} = \begin{bmatrix} \frac{\partial X(\psi,\theta,\phi)}{\partial \psi} & \frac{\partial X(\psi,\theta,\phi)}{\partial \theta} & \frac{\partial X(\psi,\theta,\phi)}{\partial \varphi} \\ \frac{\partial Y(\psi,\theta,\phi)}{\partial \psi} & \frac{\partial Y(\psi,\theta,\phi)}{\partial \theta} & \frac{\partial Y(\psi,\theta,\phi)}{\partial \varphi} \\ \frac{\partial Z(\psi,\theta,\phi)}{\partial \psi} & \frac{\partial Z(\psi,\theta,\phi)}{\partial \theta} & \frac{\partial Z(\psi,\theta,\phi)}{\partial \phi} \end{bmatrix} \quad (10)$$

Based on the physical significance of ψ, θ, and ϕ in the combined transformation RPY, the angular velocity of the center of gravity of the segment with respect to the fixed coordinate system can be expressed as

$$\boldsymbol{\omega} = R_\psi \cdot \dot{\psi} + R_\theta \cdot \dot{\theta} + R_\phi \cdot \dot{\phi} = \mathbf{J_\omega} \cdot \dot{\mathbf{q}} \quad (11)$$

where

$$\dot{\mathbf{q}} = \begin{bmatrix} \dot{\psi} & \dot{\theta} & \dot{\phi} \end{bmatrix}^T, \mathbf{J_\omega} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (12)$$

2.4 Forward Direct Kinematics of the Passive Leg

The central passive leg of 4-SPS-S manipulator is an open-loop chain with three rotation degrees of freedom, whereas the spherical joint can be regarded as three revolute joints. Denavit-Harteriberg (D-H) parameter can be used to analyze the passive leg. Therefore, the rotation angles φ_1, φ_2, and φ_3 can sufficiently describe the pose. The schematic view of the central leg is shown in Fig. 5. To provide a clearer explanation, the coincident origin of the coordinate system is dispersed.

The transformation matrix of the coordinate system of the two adjacent couplers can be denoted as

$$\mathbf{A_i} = \begin{bmatrix} c\theta_i & -s\theta_i c\alpha_i & s\theta_i s\alpha_i & a_i c\theta_i \\ s\theta_i & c\theta_i c\alpha_i & -c\theta_i s\alpha_i & a_i s\theta_i \\ 0 & s\alpha_i & c\alpha_i & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (13)$$

where θ_i represents the angle of the coordinate system rotations around z_{i-1}-axis to make x_{i-1}-axis and x_i-axis parallel; d_i represents the distance of the coordinate system translated through z'_{i-1}-axis to make x'_{i-1}-axis and x_i-axis collinear; a_i represents the coordinate system translated through x''_{i-1}-axis to make the origin of the adjacent couplers' coordinate system coincident; and α_i represents the angle of the coordinate system rotating around x'''_{i-1}-axis to make z'''_{i-1}-axis and z_i-axis collinear.

Table 1 lists the D-H parameters for the middle passive leg of 4-SPS-S. The transformation matrix between coordinate system $O-X_A Y_A Z_A$ and $O-X_B Y_B Z_B$ can be calculated by

Fig. 5 D-H notation for passive leg
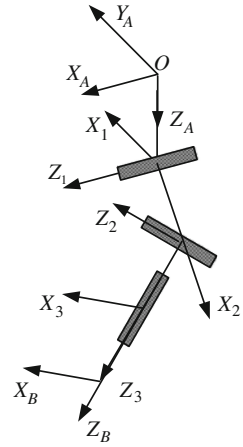
$$\mathbf{T} = \mathbf{A_1 A_2 A_3} \tag{14}$$

The relation between $\varphi_1\ \varphi_2\ \varphi_3$ and ψ, θ, ϕ can be obtained by combining Eqs. (1) and (14) as follows:

$$\varphi_i = \Gamma_i(\psi, \theta, \phi) \quad (i = 1, 2, 3) \tag{15}$$

By differentiating Eq. (16), we obtain

$$\dot{\mathbf{\Psi}} = \mathbf{J_s} \cdot \dot{\mathbf{q}} \tag{16}$$

where

$$\dot{\mathbf{\Psi}} = [\,\dot{\varphi}_1 \quad \dot{\varphi}_2 \quad \dot{\varphi}_3\,]^T, \ \mathbf{J}_s = \begin{bmatrix} \frac{\partial \Gamma_1(\psi,\theta,\phi)}{\partial \psi} & \frac{\partial \Gamma_1(\psi,\theta,\phi)}{\partial \theta} & \frac{\partial \Gamma_1(\psi,\theta,\phi)}{\partial \varphi} \\ \frac{\partial \Gamma_2(\psi,\theta,\phi)}{\partial \psi} & \frac{\partial \Gamma_2(\psi,\theta,\phi)}{\partial \theta} & \frac{\partial \Gamma_2(\psi,\theta,\phi)}{\partial \varphi} \\ \frac{\partial \Gamma_3(\psi,\theta,\phi)}{\partial \psi} & \frac{\partial \Gamma_3(\psi,\theta,\phi)}{\partial \theta} & \frac{\partial \Gamma_3(\psi,\theta,\phi)}{\partial \varphi} \end{bmatrix} \tag{17}$$

$\mathbf{J}_s$ is the Jacobian matrix between the input angular velocity of the middle passive chain and the generalized output velocity.

3 Performance Indexes of the Orientation Fine-tuning Manipulator

The parallel fine-tuning manipulator requires enough stiffness and load-bearing capacity to cope with the heavy segment. To enhance the accuracy and efficiency of segment assembly, the fine-tuning manipulator requires greater acceleration ability.

Table 1 D-H parameter for passive leg

	θ	d	a	α
1	$90° + \varphi_1$	0	0	$90°$
2	$90° + \varphi_2$	0	0	$90°$
3	φ_3	0	0	0

The transmission between input and output should not be distorted as well. This study describes the distortion relation between the input and output with the kinematics dexterity. Meanwhile, dynamic dexterity was proposed to measure acceleration performance of a specific pose. The rotation angle of the fine-tuning manipulator is small around the three axes, which is approximately three to five degrees. In addition, the spherical joints of the parallel manipulator can work normally within the scope of the corner. Therefore, the workspace of the parallel manipulator will not be taken into account. Likewise, the interference problem of the link will not be considered. In conclusion, several performances of the parallel fine-tuning manipulator should be considered, such as kinematic dexterity, dynamic dexterity, bearing capacity, stiffness, and actuating force equilibrium.

3.1 Kinematic Dexterity

Condition number of Jacobian matrix is used to measure the dexterity, which can be expressed as

$$\kappa_J = \sqrt{\lambda_{\max}(\mathbf{J}^T\mathbf{J})}/\sqrt{\lambda_{\min}(\mathbf{J}^T\mathbf{J})} \tag{18}$$

where $\kappa_{\mathbf{J}}$ is kinematic dexterity, $\lambda_{\max}(\mathbf{J}^T\mathbf{J})$ and $\lambda_{\min}(\mathbf{J}^T\mathbf{J})$ are the maximum and minimum eigenvalue, respectively.

When $\kappa_J = 1$, the manipulator owns the best transmission performance. In this case, the configuration is isotropic. When the condition number approaches infinity, the manipulator is in a singular configuration. Hence, the minimum condition number is needed.

According to Gosselin and Angeles (1989, 1991), to evaluate the kinematic performance in the entire workspace, the global dexterity index can be used, which is given by the mean value of condition number in the whole workspace:

$$\eta_J = \frac{\int_W \kappa_J dW}{\int_W dW} \tag{19}$$

where $\eta_{\mathbf{J}}$ is the the global dexterity, W is the workspace of the parallel manipulator.

3.2 Dynamic Dexterity

To measure the acceleration performance of a specific pose in the workspace, the generalized inertia ellipsoid proposed by Asada (1983) and Staicu (2009) is adopted to establish the dynamic dexterity. Dynamic dexterity refers to the rate of the square of the ellipsoid's short axis and the square of the ellipsoid's long axis, namely, the rate of the minimum and maximum eigenvalues of the mass matrix.

$$\kappa_M = \lambda_{\min}(\mathbf{M})/\lambda_{\max}(\mathbf{M}) \tag{20}$$

where $\kappa_{\mathbf{M}}$ is dynamic dexterity, $\mathbf{M}$ is the mass matrix of the operating space.

For the $n \times n$ mass matrix, the following quadratic equations can be constructed to denote generalized inertia ellipsoid GIE in the n-dimensional space:

$$x^T \mathbf{M} x = 1 \tag{21}$$

The length of the long axis is the square root of the maximum eigenvalue of the mass matrix $\mathbf{M}$. The length of the short axis is the square root of the minimum eigenvalue of the mass matrix $\mathbf{M}$. The closer the ellipsoid to the sphere—i.e., the long axis and short axis close to equal—the better the dynamic performance is. When $\kappa_M = 1$, the configuration is in the dynamic isotropy. Thus, we need the minimum κ_M.

According to the Eq. (20), the global dynamic dexterity performance can be established as

$$\eta_M = \frac{\int_W \kappa_M dW}{\int_W dW} \tag{22}$$

Dynamic model of the fine-tuning manipulator can be built through the Lagrange method to solve the mass matrix $\mathbf{M}$.

The Lagrange equation is defined as

$$\frac{d}{dt}\left(\frac{\partial K(\mathbf{q}, \dot{\mathbf{q}})}{\partial \dot{\mathbf{q}}}\right) - \frac{\partial K(\mathbf{q}, \dot{\mathbf{q}})}{\partial \mathbf{q}} + \frac{\partial P(\mathbf{q})}{\partial \mathbf{q}} = \mathbf{Q} \tag{23}$$

where $\mathbf{Q}$ is the generalized force corresponding to the generalized coordinate; $P(\mathbf{q})$ is the potential energy of the system; and $K(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2}\dot{\mathbf{q}}^T \mathbf{M}\dot{\mathbf{q}}$ is the kinetic energy of the system.

The standard form of the dynamic equation can be achieved by expanding the Lagrange equation, given by:

$$\mathbf{M}\ddot{\mathbf{q}} + V(\mathbf{q}, \dot{\mathbf{q}}) + G(\mathbf{q}) = \mathbf{Q} \tag{24}$$

where $\mathbf{M}$ is the mass matrix of the operation space, $V(\mathbf{q}, \dot{\mathbf{q}})$ is the Coriolis term, and $G(\mathbf{q})$ is the gravity coefficient.

Considering that the large mass of the segment mass of the other parallel manipulator components is negligible with respect to segment mass, the kinetic energy of the segment should be considered when measuring the kinetic energy of the parallel manipulator.

The kinetic energy of the segment, which rotates around the spherical joint on the passive chain, can be described as

$$T = \frac{1}{2}(m\mathbf{v}^T\mathbf{v} + \boldsymbol{\omega}^T\mathbf{I}\boldsymbol{\omega}) \tag{25}$$

where m denotes the mass of segment; $\mathbf{v}$ is the linear velocity of the center of segment G with respect to the fixed coordinate system; $\boldsymbol{\omega}$ is the angular velocity of segment around the center; and $\mathbf{I}$ is the inertia matrix in the coordinate system, in which the origin lies on the center of gravity and three axis are parallel to the fixed coordinate system.

According to Kelaiaia et al. (2012), kinetic energy can be calculated as

$$\mathbf{I} = {}_{\mathbf{B}}^{\mathbf{A}}\mathbf{R}[\mathbf{I_c}]_{\mathbf{B}}^{\mathbf{A}}\mathbf{R}^T \tag{26}$$

where $[\mathbf{I_c}]$ is the principal inertia matrix of segment.

Substituting Eqs. (9), (11), and (26) into Eq. (25), and expressing the latter in quadratic form, we obtain

$$T = \frac{1}{2}\dot{\mathbf{q}}^T(m\mathbf{J}_{\mathbf{v}}^T\mathbf{J}_{\mathbf{v}} + \mathbf{J}_{\boldsymbol{\omega}}^T\mathbf{I}\mathbf{J}_{\boldsymbol{\omega}})\dot{\mathbf{q}} \tag{27}$$

Hence, the mass matrix of the operating workspace can be calculated as

$$\mathbf{M} = m\mathbf{J}_{\mathbf{v}}^T\mathbf{J}_{\mathbf{v}} + \mathbf{J}_{\boldsymbol{\omega}}^T\mathbf{I}\mathbf{J}_{\boldsymbol{\omega}} \tag{28}$$

3.3 Stiffness Index

The principle of virtual work states that

$$\boldsymbol{\tau}^T\dot{\mathbf{l}} + \boldsymbol{\tau}_v^T\dot{\boldsymbol{\Psi}} = \mathbf{F}^T\dot{\mathbf{q}} \tag{29}$$

where $\boldsymbol{\tau} = \chi\Delta\mathbf{l}$ is driving force vector; $\boldsymbol{\tau}_v = \boldsymbol{\chi}_{\mathbf{v}}\Delta\boldsymbol{\Psi}$ is the internal torque of the passive chain, which does not induce any motion of the manipulator.

Mathematically, the internal torque solution vector exists in the null space of the Jacobian. Shin et al. (2011) points out that the internal preload torque can increase the active stiffness of a redundantly actuated parallel manipulator. χ and χ_v are the equivalent stiffness values of active and passive chains, respectively.

Substituting Eqs. (6) and (17) into Eq. (29), we obtain

$$\mathbf{F} = (\mathbf{J}^T \chi \mathbf{J} + \mathbf{J}_s^T \chi_v \mathbf{J}_s) \Delta \mathbf{q} \tag{30}$$

Thus, the stiffness matrix of parallel manipulator can be solved as

$$\mathbf{K} = \mathbf{J}^T \chi \mathbf{J} + \mathbf{J}_s^T \chi_v \mathbf{J}_s \tag{31}$$

According to the calculation of the vector extreme, when $\|\Delta \mathbf{q}\| = \Delta \mathbf{q}^T \Delta \mathbf{q} = 1$, the stiffness extreme value index can be calculated by

$$K_{\max} = \sqrt{\lambda_{\max}(\mathbf{K}^T \mathbf{K})} \tag{32a}$$

$$K_{\min} = \sqrt{\lambda_{\min}(\mathbf{K}^T \mathbf{K})} \tag{32b}$$

where $\lambda_{\max}(\mathbf{K}^T \mathbf{K})$ and $\lambda_{\min}(\mathbf{K}^T \mathbf{K})$ are the maximum and minimum eigenvalues of $\mathbf{K}^T \mathbf{K}$, respectively. $K_{\max}$ and $K_{\min}$ are the maximum and minimum external forces needed for the same deformation. Therefore, greater extreme value results in better stiffness. To evaluate the stiffness of the manipulator in the whole workspace, the global stiffness index is defined as follows:

$$\eta_{K_{\max}} = \frac{\int_W K_{\max} dW}{\int_W dW} \tag{33a}$$

$$\eta_{K_{\min}} = \frac{\int_W K_{\min} dW}{\int_W dW} \tag{33b}$$

3.4 Load-Bearing Capacity Index

Similar to the stiffness evaluation index, according to Lara-Molina et al. (2010), the maximum and minimum load-bearing capacities in the workspace and evaluation index can be achieved.

The maximum and minimum load-bearing capacity can be reduced to

$$G_{\max} = \sqrt{\lambda_{F\,\max}(\mathbf{G}^T\mathbf{G})} \tag{34a}$$

$$G_{\min} = \sqrt{\lambda_{F\,\min}(\mathbf{G}^T\mathbf{G})} \tag{34b}$$

where $\lambda_{F\,\max}(\mathbf{G}^T\mathbf{G})$ and $\lambda_{F\,\min}(\mathbf{G}^T\mathbf{G})$ are the maximum and minimum eigenvalues of the matrix $\mathbf{G}^T\mathbf{G}$ respectively, where $\mathbf{G} = \mathbf{J}^T$.

Global load-bearing capacity index will be computed as

$$\eta_{G\,\max} = \frac{\int_W G_{\max}dW}{\int_W dW} \tag{35a}$$

$$\eta_{G\,\min} = \frac{\int_W G_{\min}dW}{\int_W dW} \tag{35b}$$

3.5 Actuating Force Equilibrium Index

This study uses MD ADAMS software to solve the actuating force for parallel manipulator. The primary steps taken to solve driving are given as follows:

(1) Given the output angles, the inverse solution of the parallel manipulator is achieved using Eq. (4); thus, the input parameters of four active chains. The movement track equation of the three output angles change with time was shown as below:

$$\begin{cases} \psi = 3° \times \sin(\pi \times time) \\ \theta = 3° \times \sin(\pi \times time) \\ \phi = 5° \times \sin(\pi \times time) \end{cases} \tag{36}$$

(2) With the actuating parameters achieved in Step 1, spline curves are generated in ADAMS, and subsequently used in the AKISPL function to drive the active limbs. The format of the AKISPL function is AKISPL (x, z, modelname. curvename, order), which argument x can be defined time or a function that returns a real number, and z must be a function that returns a real number. Otherwise, 0 is applied. Curve name is the name or argument number of the spline data. Order is the interpolation method for the functions (return the value if 0, return calculation for 1st order differential equation if 1, and return calculation for 2nd order differential equation if 2).

(3) The gravity of the segment is expressed as the force with direction based on the system gravity. The actuating force curve can be extracted after simulation.

(4) Calculate the maximum value of each active limb, then solve the mean value $\bar{a}$ of four maximum values. Finally, the quadratic sum σ of the difference between the four maximum actuating forces is calculated.

(5) $\bar{a}$ is defined as the index of the maximum input, and σ is defined as the actuating force equilibrium index of four active limbs.

4 Multi-objective Optimization Model for Fine-tuning Manipulator

Since the independent optimization process of each single index may sacrifice the others, the simultaneous optimization of global dexterity index, global dynamic dexterity index, global stiffness index, global load-bearing capacity index and actuating force equilibrium index, as the key indexes of parallel manipulator, is critical to improve the realistic performance of parallel manipulator. A multi-objective optimization issue will be studied here to search the optimal solution. Generally, this issue is addressed by considering a series of selected points which represent the tradeoff solutions between different objective functions.

According to Cui et al. (2013), Sun et al. (2012) and Alici and Shirinzadeh (2004), the traditional optimization methods usually apply the local search by convergent stepwise procedure which possibly falls into local optimal solution. If the complex function to be optimized does not possess convexity characteristics that essentially satisfy that the local extreme point is a global optimum, a global optimization algorithm is required.

4.1 Design Variable

Based on the application and structure configuration of the parallel fine-tuning manipulator, three parameters are selected as design variables (Fig. 3), which can be expressed as

$$X = \begin{bmatrix} h & s & t \end{bmatrix}^T \tag{37}$$

where the parameter h is used to determine the position of spherical joint of the passive leg, and s and r can describe the specific spatial locations of the redundant limb.

4.2   Objective Function

The optimization objective function of the parallel fine-tuning manipulator can be expressed as $\min P$, and

$$P = [\,\eta_J \quad -\eta_M \quad -\eta_{K\max} \quad -\eta_{K\min} \quad -\eta_{G\max} \quad -\eta_{G\min} \quad \overline{a} \quad \sigma\,]^T \qquad (38)$$

4.3   Constraint Condition

Owing to the limited rotation angle of the parallel fine-tuning manipulator, the workspaces, rotation angle of kinematic pairs, and interference between chains are not considered. Considering that the redundant limb can completely eliminate the singularities of parallel manipulator, the effect of singular configuration need not be considered. The main constraint conditions of this parallel manipulator to be considered include the range of the design variables.

Subsequently, the following multi-objective optimization design model can be established:

$$\begin{cases} \min f_1(X) = \eta_J(h\,s\,t) \\ \max f_2(X) = \eta_M(h\,s\,t) \\ \max f_3(X) = \eta_{K_{\max}}(h\,s\,t) \\ \max f_4(X) = \eta_{K_{\min}}(h\,s\,t) \\ \max f_5(X) = \eta_{G_{\max}}(h\,s\,t) \\ \max f_6(X) = \eta_{G_{\min}}(h\,s\,t) \\ \min f_7(X) = \bar{a}(h\,s\,t) \\ \min f_8(X) = \sigma(h\,s\,t) \\ s.t. \ \ 150\ \text{mm} \ \le\ h\ \le\ 310\ \text{mm} \\ 300\ \text{mm} \ \le\ s\ \le\ 600\ \text{mm} \\ 0.4 \le t \le 0.8 \end{cases} \qquad (39)$$

5   Multi-objective Optimization Design Based on Isight

The manipulator performance is closely related to its geometric parameters. Dimension optimization enables the manipulator to achieve good performance in the orientation workspace. In this article, we accomplished the multi-objective optimization design for the parallel manipulator based on Isight, which can integrate several software by Simcode component, while the software are connected to one another by files. Isight manages the data files and invokes the software to complete the function. As shown in Fig. 6, the basic process of the parallel
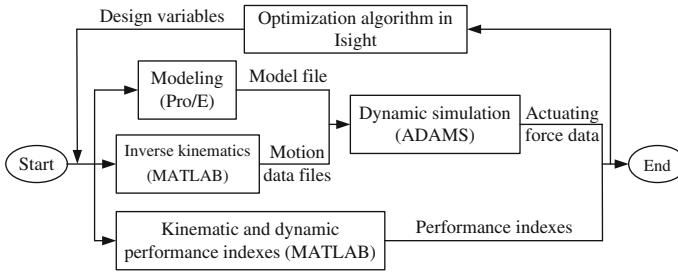
Fig. 6 The design flow of the parallel manipulator optimization based on Isight

manipulator optimization can be described through the following steps proposed by Hao and Merlet (2005):

Step 1   Choose the proper design variable, constraint condition, and optimization objective. Subsequently, establish the multi-discipline and multi-objective optimization model for the parallel fine-tuning manipulator.

Step 2   The optimization algorithm in the Isight generates and provides a group of variables to each component.

Step 3   Pro/E receives the design variables to rebuild the parameterized model and output the model. Meanwhile, MATLAB completes the inverse kinematics for the parallel manipulator and outputs the motion data files. ADAMS uses the new model and data files to accomplish the dynamic simulation, and then outputs the data files of driving force of the chains.

Step 4   In another branch, MATLAB solves the index of the kinematic and dynamic performance, including kinematic dexterity, dynamic dexterity, load-bearing capacity, and extreme stiffness value, among others.

Step 5   The output data files will be sent to the optimization algorithm after accomplishing all the branches. Then, Isight provides the components to the next group variable to perform the next loop or to the end of the optimization and output the results.

As shown in Fig. 7, an optimization design platform for the parallel manipulator is established using Isight, which integrates MATLAB, Pro/E, and ADAMS. A component named coordinate point calculation is used to calculate the manipulator's indirect parameters, as well as the position and orientation of the changing joints involved in ADAMS. Four Simcode components integrate MATLAB, Pro/E, and ADAMS to solve the performance indexes. The script component named analysis of driving force processes the data on driving force output by ADAMS to obtain the maximum driving force of each chain. A calculator component named evaluation of the maximum driving force is used to solve the mean $\bar{a}$ of the maximum driving force and quadratic sum σ of the difference between maximum driving forces and mean. Moreover, the Task Plan component is used to integrate DOE (i.e., Design of Experiments, which is a mathematical statistics method
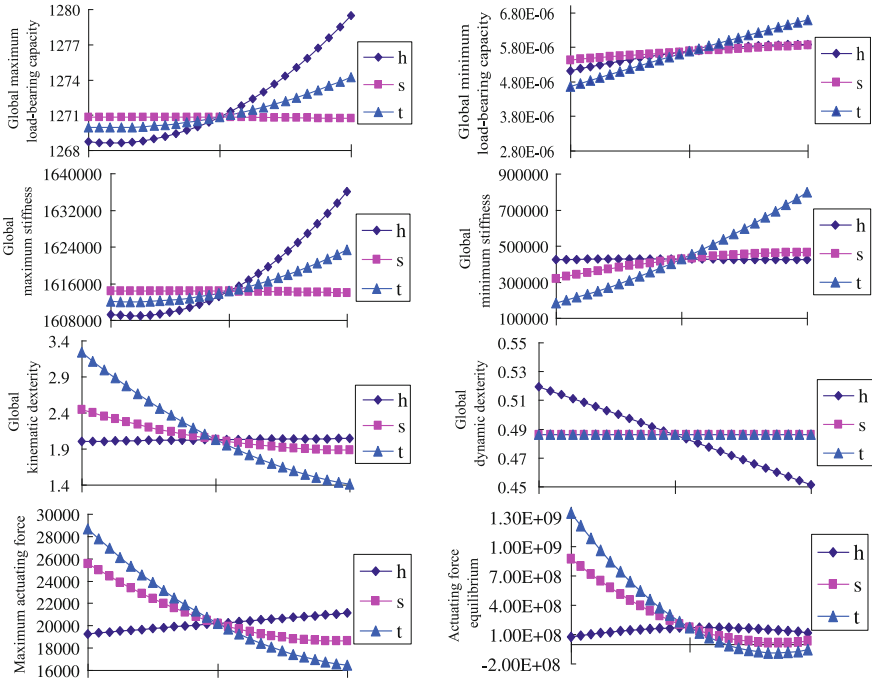
Fig. 7 Main effects plot for the manipulator performance indexes

applied to experimental arrangement and experimental data analysis) and optimization design. These two components will complete the operation based on the process established. DOE can decipher the near-optimal solution. The optimization component will accomplish the multi-objective optimization based on the near-optimal solution. Hence, the optimization cost will be visibly reduced.

6 Optimization Design Results and Analysis

In this example, normalization method is adopted to optimize the fine-tuning parallel manipulator. The proper design parameters from Pareto set are chosen as the optimal solution. Taking into consideration the complexity of the model and computational cost, optimal Latin hypercube design is chosen as the algorithm for DOE. The number of points is set as 20. We choose the design variable as factors and the object as response. In the optimization component, nondominated sorting genetic algorithm II (NSGA-II) is selected as the optimization algorithm which is the improved version of NSGA, the population size is set as 20, and genetic algebra is set as 20.

Table 2 Design variables and performance indexes before and after optimization

	h	s	t	$\eta_{G\ max}$	$\eta_{K\ max}$	η_M
Before optimization	195	370	0.6	1269.22	1,610,360	0.5007
After optimization	307	574	0.7991	1284.43	1,649,110	0.4524
Rate of change (%)	57.4	55.1	33.2	1.2	2.4	−9.6

6.1 Result of the Experiment Design

Experiment design based on Isight can conveniently determine the influence of the design variables on the responsive variables. Figure 7 shows the main effects plot for each response. Taking the global maximum load-bearing capacity as an example, the influence of the design variable s on the maximum bearing capacity is the smallest, which is basically linear. Simultaneously, the influence of h is the greatest, which is a non-linear variable. This result indicates that the position of the spherical joint on the passive leg significantly influences the load-bearing capacity. Hence, during the design of parallel fine-tuning manipulator, the distance between the spherical joint on the middle chain, the fixed platform, and the spherical joint installation error should be reasonably controlled to ensure the requirements of the maximum load-bearing capacity. We used unit force when we deduced the maximum and minimum stiffness, so the load bearing capability is relatively small.

6.2 Result of Multi-objective Optimization

(1) Multi-objective optimization design can determine the design feasibility (that is, an analysis and evaluation of a proposed project to determine if it is feasible within the estimated range. Design feasibility is advantageous for the selection of the optimal solution). Likewise, this design can obtain the Pareto sets and Pareto frontier.

The design feasibility of multi-objective optimization is shown in Fig. 8. In a manner of speaking, all the Pareto sets are optimal solutions. However, the Pareto optimal solutions cannot constantly satisfy all the objectives. Therefore, decision makers should choose a compromise based on actual needs when developing a solution.

As shown in Fig. 9, dynamic dexterity, actuating force equilibrium, maximum load-bearing capacity, and maximum stiffness are chosen to draw the Pareto frontier through Isight. The graph indicates that dynamic dexterity, maximum load-bearing capacity, and stiffness are contradictory. That is, increasing dynamic dexterity will result in the reduction of the maximum load-bearing capacity and maximum stiffness. Therefore, decision makers need to choose these factors based on the actual situation after the Pareto set and frontier are achieved.
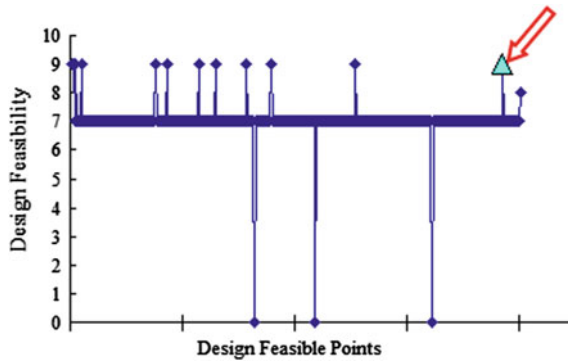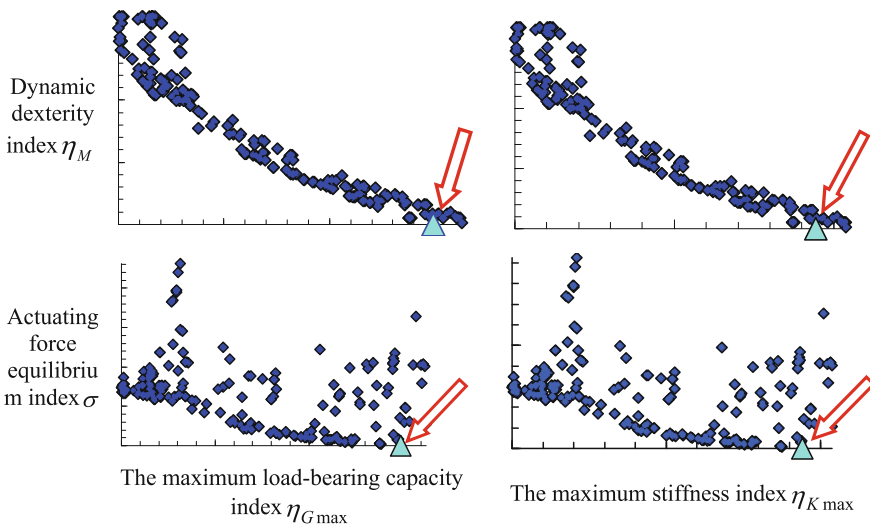
Fig. 8 Design feasibility

Fig. 9 Pareto frontier of four performance indexes

(2) Analysis of the optimization results of the segment fine-tuning manipulator.

The parallel fine-tuning manipulator requires sufficient stiffness and load-bearing capacity to handle the heavy segment. Hence, the load-bearing capacity and stiffness of this manipulator need to be strengthened, and an optimal design for the fine-tuning manipulator should be achieved.

The increase in the load-bearing capacity and stiffness will reduce dynamic dexterity, as shown in Fig. 10. The segment needed to be assembled weighs approximately four tons. After having selected the condition of the driving motor, the point with the arrow is chosen as the optimal solution, as shown in Fig. 9, the
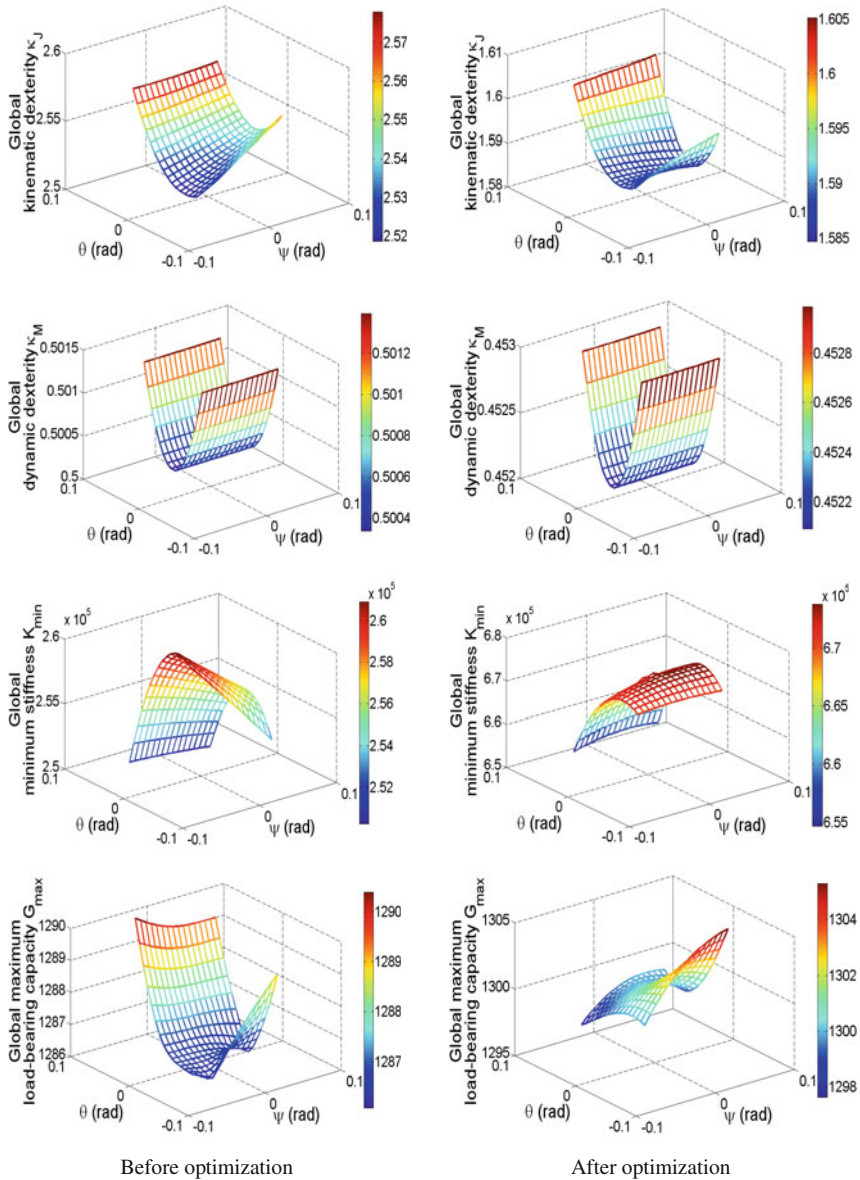
Before optimization   After optimization

Fig. 10 Comparison of the indexes before and after optimization

feasibility of which is nine (Fig. 8). Table 2 provides a comparison of the design variable and performance indexes before and after optimization.

Table 2 indicates the improvement of the load-bearing capacity and stiffness and decrease in dynamic dexterity, which indicate the decrease in acceleration

performance. The result indicates that the assembly efficiency is sacrificed to increase the load-bearing capacity and segment assembling accuracy.

Figure 10 shows the workspace distribution maps of the performance indexes, such as kinematic dexterity, dynamic dexterity, load-bearing capacity, and stiffness before and after optimization.

The maps, shown in Fig. 10, indicate the significant improvement of the optimized kinematic dexterity, stiffness, and load-bearing capacity and the slight decrease of dynamic dexterity. The overall analysis show that although the acceleration performance decreased, most performance indexes improved with different degrees. This observation indicates that the result of the selected optimization method can meet the requirement. If the drive motor can provide enough load-bearing capacity and the stiffness meets the requirements, we can increase the dynamic dexterity to improve work efficiency.

7   Conclusion

This study focuses on the multi-objective dimension optimization of the 3-DOF orientation fine-tuning manipulator for segment assembly robots in shield tunneling machine. A multi-objective optimization model of the 4-SPS-S actuation redundant parallel manipulator is established. This model takes into consideration five performance indexes, namely, kinematic, dynamic dexterity, load-bearing capacity, stiffness, and actuating force equilibrium. A multi-objective optimization method for parallel manipulator is proposed using the application named Isight, which integrates various engineering software. Results show that the proposed method can automatically accomplish the calculation of kinematic and dynamic performance index, modeling, and solving by Isight. Decision makers can thus select a proper optimal solution based on the achieved Pareto sets and Pareto frontier. The performance of the proposed parallel manipulator is significantly improved by optimizing the design variables. The dimension optimization method introduced in this paper can be expanded to other applications in designing other orientation fine-tuning manipulators or parallel manipulators.

References

Alici, G., & Shirinzadeh, B. (2004). Topology optimisation and singularity analysis of a 3-SPS parallel manipulator with a passive constraining spherical joint. *Mechanism and Machine Theory, 39*(2), 215–235.

Asada, H. (1983). A geometrical representation of manipulator dynamics and its application to arm design. *ASME Journal of Dynamic Systems, Measurement, and Control, 105*, 131–135.

Chen, J., & Liu, Q. (2008). Optimization design of new typed 2-DOF parallel mechanism based on genetic algorithm. *Journal of Machine Design, 25*(2), 21–23.

Cui, G. H., Zhang, Y. W., Zhang, Y. S., & Wu, H. M. (2010). Configuration design and kinematic analysis of serial-parallel manipulator with six degrees of freedom. *Transactions of the Chinese Society for Agricultural Machinery, 26*(1), 155–159.

Cui, G. H., Zhou, H. D., Wang, N., & Zhang, H. Q. (2013). Multi-objective optimization of 3-UPS-S parallel mechanism based on Isight. *Transaction of the Chinese Society for Agricultural Machinery, 44*(9), 261–266.

Gao, Z., Zhang, D., & Ge, Y. J. (2010). Design optimization of spatial sex degree-of-freedoms parallel manipulator based on artificial intelligence approach. *Robotics and Computer-Integrated Manufacturing, 26*(2), 180–189.

Gosselin, C. M., & Angeles, J. (1989). The optimum kinematic design of a spherical three degree-of-freedom parallel manipulator. *Journal of mechanisms, Transmissions, and Automation in Design, 111*, 202–207.

Gosselin, C. M., & Angeles, J. (1991). A global performance index for the kinematic optimization of robotic manipulators. *Journal of Mechanical Design, 113*, 220–226.

Hao, F., & Merlet, J. P. (2005). Multi-criteria optimal design of parallel manipulators based on interval analysis. *Mechanism and Machine Theory, 40*(2), 157–171.

Kelaiaia, R., Company, O., & Zaatri, A. (2012). Multiobjective optimization of a linear delta parallel robot. *Mechanism and Machine Theory, 50*, 159–178.

Kosuge, K., Takeo, K., Taguchi, D., Fukuda, T., & Murakami, H. (1996). Task-oriented force control of parallel link robot for the assembly of segments of a shield tunnel excavation system. *IEEE/ASME Transactions on Mechatronics, 1*(3), 250–258.

Lara-Molina, F. A., Rosario, J. M., & Dumur, D. (2010). Multi-objective design of parallel manipulator using global indices. *The Open Mechanical Engineering Journal, 4*, 37–47.

Liu, L. Q., Zhang, X. L., Xie, L. M., Wen, S. H., & Du, J. (2012). Multi-objective design for 6-RSS parallel robot mechanism. *Journal of Taiyuan University of Science and Technology, 33* (4), 281–285.

Shi, H., Gong, G. F., Yang, H. Y., & Zhou, R. L. (2009). Position and attitude precision analysis of segment erector of shield tunneling machine. *International Conference on Intelligent Robotics and Applications* (pp. 355–363).

Shin, H., Lee, S. C., In, W., & Jeong, J. I. (2011). Kinematic optimization of a redundantly actuated parallel mechanism for maximizing stiffness and workspace using Taguchi method. *Journal of Computational and Nonlinear Dynamics, 6*(1), 011017.

Staicu, S. (2009). Dynamics of the spherical 3-UPS/S parallel mechanism with prismatic actuators. *Multibody System Dynamics, 22*(2), 115–132.

Sun, T., Song, Y. M., Dong, G., Lian, B. B., & Liu, J. P. (2012). Optimal design of a parallel mechanism with three rotational degrees of freedom. *Robotics and Computer-Integrated Manufacturing, 28*(4), 500–508.

Tanaka, Y. (1995). Automatic segment assembly robot for shield tunneling machine. *Computer-Aided Civil and Infrastructure Engineering, 10*(5), 325–337.

Wu, C., Liu, X. J., Wang, L. P., & Wang, J. S. (2011). Dimension optimization of an orientation fine-tuning manipulator for segment assembly robots in shield tunneling machines. *Automation in Construction, 20*(4), 353–359.

Wu, J., Wang, J. S., Li, T. M., & Wang, L. P. (2007). Performance analysis and application of a redundantly actuated parallel manipulator for milling. *Journal of Intelligence & Robotic System, 50*, 163–180.

Zyada, Z., Hasegawa, Y., & Fukuda, T. (2003). Multi-directional assembly of tunnel segments using a force controlled parallel link robot with fuzzy compensation. *SICE annual conference in Fukui* (pp. 2044–2049). August 4–6, 2003.

An Imitation Framework for Social Robots Based on Visual Input, Motion Sensation, and Instruction

Mohsen Falahi, Faraz Shamshirdar, Mohammad Hosein Heydari and Taher Abbas Shangari

Abstract In this paper a new framework with three classes of inputs, including "motion sensation", "visual", and "instruction" inputs is introduced. This framework provides a fluent platform for a safe Human Robot Interaction (HRI) thanks to imitation which can be used widely in social robotics. The proposed framework benefits from the ability of mutual communication between robot and the demonstrator which facilitates the interaction and cause in removing all ambiguities in complex, new, or unknown tasks. This framework considers both levels of imitation and can be trained with different teaching methods based on the inputs. For low-level imitation, OFTM is used which is a fast and accurate method for learning trajectories. For the high-level imitation, Kragic's method is employed which enables the robot to infer based on seen sequences of actions. Moreover, feeding the framework by an instruction which is provided by demonstrator guides the robot to figure out how it can combine learnt actions in different imitation levels. It means that this framework enables the robot to combine different actions and concepts in both levels of imitation. To clarify how this framework can be used, a scenario of cooking is introduced and the role of all involved blocks is described in each section of learning and execution. This framework is implemented on a 6DoF robotic arm for experiments.

Keywords Learning by imitation · Social robots · OFTM · Human robot interaction

1 Introduction

It is becoming increasingly difficult to ignore the role and need of social robots in human daily life due to their abilities in assisting people with their tasks and helping people with social deficits (see Fong et al. 2003; Kim et al. 2013; Kanda et al. 2004).

M. Falahi (✉) · F. Shamshirdar · M.H. Heydari · T.A. Shangari
Cognitive Robotics Lab, Amirkabir University of Technology, Tehran, Iran
e-mail: mfalahi@gmail.com

These robots should be able to interact safely with human in non-engineered environment (Schaal 1999). As the environment and the way by which people may interact with the robot cannot be completely modeled, a social robot has to adopt a safe method for its interactions (Billard et al. 2008). Based on imitation features which guarantee this safety, learning by imitation is widely used for teaching social robots (Breazeal et al. 2005; Falahi et al. 2014).

Basically, learning by imitation accomplishes in two levels; high and low (Billard et al. 2008). In low-level imitation, an entity only reproduces observed action without considering the purpose of the action. However, in high-level imitation, concepts and intentions are imitated even if the reproduction be completely different from demonstrated one.

One of the most important factors which increases the efficiency and performance of imitation in robotics, independent to the level of imitation, is the framework. A framework applies rules on robot's interactions and clarifies the relationship between demonstrator and imitator which plays the central role in imitation. A considerable amount of literature has been published on imitation frameworks. These studies can be divided in two main categories. A group of researchers tried to propose a general framework by which robots can learn a variety of tasks based on imitation (Bandera et al. 2012; Tan 2012). These studies use different teaching methods including vision based, kinesthetic teaching etc. On the other side, in some researches on robot learning by imitation, based on the assigned task, a framework is customized (Cho 2012; Mohan et al. 2011; Kormushev et al. 2011).

Proposing a framework for learning is not narrowed by engineers and computer scientists' field of interest and has been investigated even more by phycologists. Bandura (1969) provided four crucial components for social learning consist of "attention", "memorization", "motivation" and "generation". The interaction and the way these elements work together create the basis of learning framework in general. As in this study we are proposing a framework for robots, we do not discuss the details of his study and continue this section by introducing robotic frameworks. However, there are a large number of attempts in robotics to embed all these four elements in a single framework. Nonetheless, they usually ignore some parts in their implementations (see Bandera et al. 2012).

Bandera et al. (2012) have a comprehensive review on component based learning frameworks for imitation which presents a comparison of Schaal (1999), Derimis (2002), Breazeal et al. (2004), Mohammad (2009) and Mühlig et al. (2009) frameworks and proposes a vision based framework.

The major drawback of these frameworks is that some of them are based on only one teaching protocol (as an example Bandera's framework is specialized for vision based teaching). This deficiency causes in ignoring other abilities of the robot and disusing other existing sensors during the teaching phase. In addition, previous frameworks suffer from the disability of combining different levels of imitation (they even are not able to perform in both levels of imitation concurrently or separately). As another example, these frameworks do not consider any instruction-like input in the framework for accomplishing algorithmic tasks.

In this paper, we have introduced a new framework for learning by imitation. The key features of this framework are as follow:

- Supporting two mostly used teaching protocols; kinesthetic and vision-based
- The ability of learning in both imitation levels
- The feasibility of combining low and high level primitive actions for accomplishing a complex task
- Providing an input for demonstrator's instructions and needed algorithm
- Considering a mutual communication between robot and demonstrator in the case of any ambiguity.

2 Proposed Framework

The proposed framework consists of six main components; "Input", "Perception", "Learning", "Knowledge", "Generation", and "Output" (see Fig. 1). Each component has some subparts which will be introduced in the following text.

2.1 Inputs

In this framework three kinds of inputs are considered; "visual input", "motion sensation", and "instruction". In the visual input all data from robot's camera can be
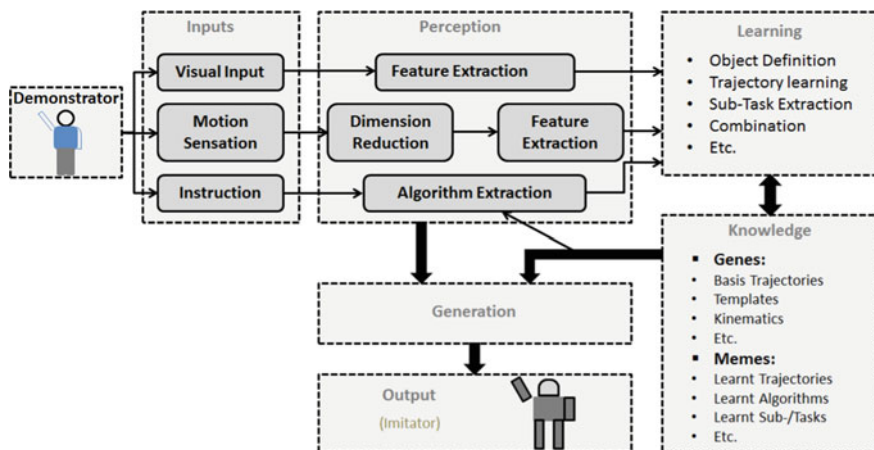

Fig. 1 This figure shows all blocks of proposed framework including "Input", "Perception", "Learning", "Knowledge", "Generation", and "Output". The Demonstrator is a human or robot which is supposed to teach the imitator. The *arrows show* how these components interact with each other

feed to the architecture which can be acquired from a monocular or stereo vision. In the motion sensation the data of encoders, gyroscopes and all related sensors comes in. As the last part, the demonstrator can mention his own instruction for accomplishing a task or even the algorithm of performing a certain task. This input helps the demonstrator and imitator to interact more fluently, especially in sequential tasks like arranging a desk.

"Input" block communicates with perception block and provides all received signals from the environment and demonstrator.

2.2 *Perception*

As it mentioned, input block does not apply any changes on the acquired data and only play the role of an interface between the framework and surrounding environment. Hence, the perception block should reduce the amount of data by mapping them to a simpler space without losing important information. Based on the nature of data, i.e., where the data is coming (from motion, instruction, or visual input), different algorithms are applied to them. For visual data, needed changes are applied to the frames for obtaining the distance of the objects, their color etc. On the other side, for the data provided by motion sensation, firstly, the data is mapped from joint space to end-effector position to obtain the trajectory. In the "instruction perception", the fed instruction is changed to a practical algorithm which defines the sequence of actions. In addition, using this input enables the demonstrator to attach labels to all objects, actions, and even introduced concepts. It means that this part is not only a one-way input and robot can interact with the demonstrator in the case that the new seen action/object/task is new for the robot or there is any ambiguity for the robot in other parts of perception. To clarify the role of this element, see experiments provided in "Results" section which shows how this component enables robot to provide related questions for demonstrator.

2.3 *Learning*

The learning component is in charge for classification of receipted information. This block is connected to perception, generation, and knowledge blocks. When the robot receives any command or demo from the demonstrator, in the first step, "learning block" checks if the robot knows that or not. In the case that it was completely new for the robot, it will be encoded in the knowledge block as a new action/object/task/concept. It means that this block checks the similarity of seen demo with previously known actions and then classifies them.

The learning process accomplishes in both levels of imitation. For the low-level imitation which is related to reproducing trajectories, learning block tries to decompose seen trajectories and recompose it based on known trajectories in

knowledge block (using OFTM, see Falahi and Jannatifar 2013; Falahi et al. 2015). On the other side, in high-level imitation, this block starts to decompose a given task to sub-tasks and complete the task by combination of known sub-tasks [for high-level imitation Kragic's method is used (see Ekvall and Kragic 2006)].

The "learning block" receives information about the definition of objects and learns trajectories by extracting sub-trajectories stored in "knowledge". In addition, this block is in charge with checking the feasibility of producing a task by combining known sub-tasks. It is clear that this block has the permission to access and manipulate information in knowledge block (the access level is introduced in next subsection).

2.4 Knowledge

There are two types of information in knowledge block; information related to the learnt tasks and previously encoded information (known as mems and genes respectively). Obviously the first type of information in this block are acquired from robot's experiences, however, genes are encoded as previous knowledge. An information which is encoded as gene remains intact and the learning process cannot change it. In our implementation we have used OFTM method which is firstly proposed by Falahi et al. (2013). Hence, the orthogonal basis functions and templates are saved as genes and robot uses them during learning process (in decomposing a trajectory and recomposing it before "action generation"). In addition, the kinematics of the robot is saved in knowledge block as a gene.

Basically, this block plays the role of database in the proposed architecture. It means that all basis trajectories, templates (frequently repeated or important trajectories (see Falahi et al. 2015), learnt trajectories, seen and defined objects, learnt instructions/algorithms, and tasks/sub-tasks are stored in this block. When demonstrator finishes his teaching, and the robot is asked to execute a task, the interaction of this block with algorithm extraction part in perception block cause in execution of the task in the case of complete knowledge of the task (otherwise, the robot asks for more information from demonstrator).

2.5 Generation

This component consists of two major parts; "Action Generator" and "Interaction Command Generator". This block gathers data from perception, learning, and knowledge blocks. Based on these data, this block can prepare needed trajectories ("action generation") for the robot or generate inquiries from demonstrator (Interaction Command Generator) as its output.

"Action Generator" produces motion points only in the case that the wanted task is completely known. In this case, based on the given instruction from the

demonstrator it starts to generate all needed trajectories. For this purpose it has a read-only access to the knowledge. Moreover, this block checks the validity of needed trajectory which means that if there is any part of the task that the robot cannot execute it (because of its kinematics or dynamics), it would not send any execution command to the robot's joints. As an example, suppose that the robot should grip a glass and the demonstrator showed that to the robot. In this case the robot has learnt this task and has enough information, hence, when the robot is asked to reproduce it, action generation block checks the position of the glass and if it was not reachable due to limitations on robot's kinematics, it will not start the task (it is the same for dynamics as well). Additionally, during task execution this block checks if all subtasks are completed correctly or not and updates commands based on the situation.

Interaction Command Generator, handles commands regarding asking for more information, demonstration, labels etc. it means that when the robot do not know all sub-tasks or even a trajectory to complete the given task, it sends a command to the output to start a new interaction with the demonstrator.

2.6 Output

This part includes whole body of the robot with all its actuators and sensors. When the action generator sends a command to this block, the robot executes it. As we are trying to have a social robot, it is highly recommended to use a human-like body for the robot to increase the acceptability of the robot in human daily life (Breazeal et al. 2008). Additionally, this block handles the low-level controllers of robot's joints for controlling the whole robot's body.

3 Scenario and Experimental Results

In this section a comprehensive example is introduced which allows us to discuss highlighted features of the proposed framework. In the beginning, the task is described and then the role and performance of each component in the framework will be discussed. As the aim of this study is presenting a general framework for learning by imitation, it is clear that our experiments would not cover all abilities of the framework. However, we have tried to choose a task which shows the most important and distinguishing features of the framework. The considered task in this experiment is cooking which consists of a wide range of subtasks. Before going to the details of the task, our experimental setup is presented.

3.1 Experimental Setup

The implementation of the proposed framework is done on a 6DoF robotic arm (see Fig. 2). Additionally, a camera is located on top of the table which provides images of the table and the arm.

Each joint of this arm is equipped with a Dynamixel Smart Motor which can control position, torque, and the velocity of attached joint. To have a proper control on these joints a PID controller is applied. Moreover, a GUI is designed as an interface for the demonstrator which facilitates attaching labels to the seen actions, objects, and introducing new concepts for the robot (see Fig. 3). This interface enables the demonstrator to feed the instruction and monitor the learning process. Through this interface the demonstrator has access to the knowledge in the case of any error during teaching.

3.2 Task Structure and Demos

In the considered task, a person starts to interact with the robot in order to teach cooking. The interaction is as follow; firstly, the demonstrator shows each object to the robot (in this experiment there are a pan, heater, and stirrer). After labeling these objects, demonstrator starts defining subtasks; gripping the stirrer, reaching the pan, gripping the pan, two types of stirring (circular and zigzag), releasing the pan in a free space, and putting the pan on the heater. All these subtasks will be encoded in the "knowledge block" of the framework as mems. It is clear that all trajectories are considered to be reproduced based on low-level imitation rules using OFTM.

Furthermore, valid combinations of these subtasks are introduced by the demonstrator as macros. Here, three macros are defined; "Put the pan on heater",

Fig. 2 This figure shows the robotic arm which is used for implementation of the proposed framework and executing cooking tasks
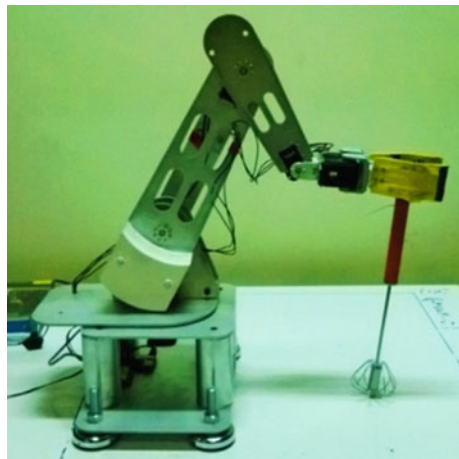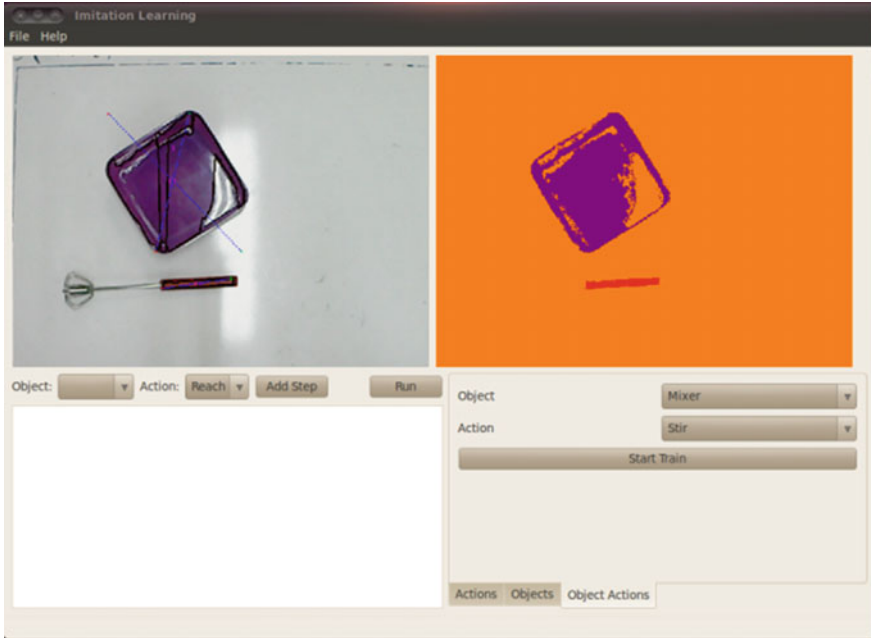
Fig. 3 This figure shows a view of the interface which is designed for the demonstrator to feed the instruction and monitor the learning process

"Stir the food circularly", and "Stir the food with zigzag pattern". The first one is consisted of "reach the pan", "grip it", and "put the pan on heater". The second and third macros are "reach the stirrer", "grip the stirrer", "stir food circularly—or—with zigzag pattern".

Based on the level of complexity of the trajectories, the demonstrator can use different teaching methods provided in this framework. In this experiment, for artfully and tiny motions (like stirring food), the demonstrator used the kinesthetic teaching because of difficulties in capturing small movements by a typical camera. On the other hand, other goal directed actions (like reaching the pan) are taught to the robot via vision based teaching.

In addition to these separated subtask demos and macros, the demonstrator provides two different instructions for accomplishing the "**same**" task as follows (this task is labeled as "Cooking"):

- Put the pan on heater, stir food circularly, and release the stirrer in a free space.
- Stir food circularly, put the pan on heater, stir food with zigzag pattern, and release the stirrer in a free space.

The whole process until now in imitation parlance is considered as "teaching phase" and after this, we will show what would happen if the robot be asked to

perform a task. The following section explains the robot behavior in different situations which is usually known as "reproduction phase".

3.3 Results

In this section we show how the robot accomplishes tasks based on what it learnt from demonstrator. In each task, involved components of the framework and their role will be introduced. It is clear that as in this research we are proposing a framework, we would not go to the details of mathematics of used approaches. We are suggesting the readers to go through the related references in order to have a profound understanding of each method performance. Basically, in this section our focus is on the abilities of proposed framework rather than comparing it with aforementioned frameworks while none of them covers these abilities and this framework is the first interactive framework with multimodal input in robotics (see Sect. 1).

3.3.1 First Experiment

In the first task, the robot should reproduce a series of subtasks; "reach the stirrer", "grip it", and "stir the food with zigzag pattern". Involved elements of the framework are highlighted in Table 1.

As it is shown in Table 1, the demonstrator only provides the instruction for the robot ("instruction" is highlighted in the input). Then, the "algorithm extractor" decomposes the instruction to a sequence of actions. Learning block checks (using "checking") with "knowledge" block if the robot knows all these actions or not. Since this task constructed by known actions, the learning block sends an acknowledgement to the "action generation" (actions are encoded as "mems" during the learning phase). Finally, the action generator produces related trajectories and sends them to the output for execution.

Table 1 This table shows which components of main blocks are involved in reproducing the task consist of "reach the stirrer", "grip it", and "stir the food with zigzag pattern"

Input	Perception	Learning	Knowledge	Generation	Output
Visual input	Feature extraction1	Low-level	Genes	Action	Body movement
Motion sensation	Dimention reduction	High-level	Mems	Command	Ask for more info
Instruction	Algorithm extraction	Checking			

The only input in this task is the instruction. As the robot has learnt all related subtasks, the instruction will be executed by the robot

3.3.2 Second Experiment

In this experiment, the robot is supposed to accomplish a more complex task. For this purpose the demonstrator asks the robot to start "cooking". As it mentioned in the previous section, there are two different demos for cooking task and in this experiment, the demonstrator does not specify which demo should be reproduced. Hence, the robot, based on the extracted rules from the demos, produces a new unseen sequence of actions (this reasoning is based on Kragic's method). The procedure is illustrated in Fig. 4.

Based on Kragic's method for high level imitation, the robot can infer another unseen instruction according to the similarities and differences of previously fed instructions. Here, as the zigzag stirring was not mentioned in both demos, the robot concluded that it is not important and eliminated it from the new sequence. In addition, as the action of putting the pan on heater can be done in the first or second order, it infers that this action can be executed before or after stirring the food. In this experiment you see how the framework enables the robot to make decisions based on the information provided in the input. The only difference in this experiment with the first one is the fact that the interaction of high level imitation block and the knowledge block causes a reasoning process, while other involved components are the same for this experiment.

3.3.3 Third Experiment

As the last experiment in this study, we want to show what happens if the demonstrator asks the robot a new task which is partially learnt before, not completely. For this purpose, here, the demonstrator proposes a task as below.

- Put the pan on heater, stir food circularly, release the stirrer in a free space, and turn the heater on.
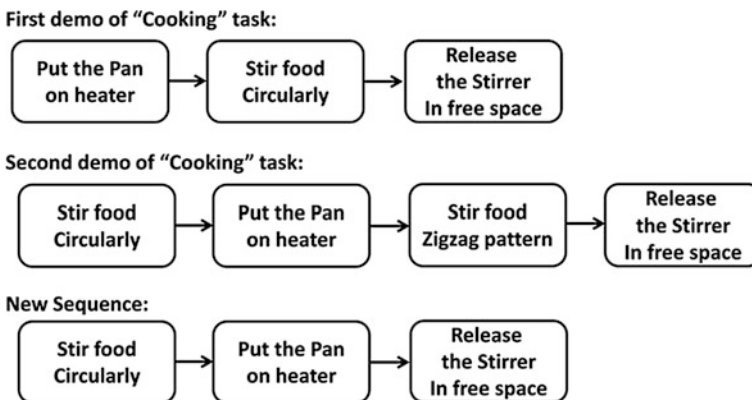

Fig. 4 This figure shows how the Kragic's method can cause in a new unseen execution sequence of known task, based on previous demos

Table 2 This table shows which components of main blocks are involved in accomplishing a task with one unknown sub task

Input	Perception	Learning	Knowledge	Generation	Output
Visual input	Feature extraction1	Low-level	Genes	Action	Body movement
Motion sensation	Dimention reduction	High-level	Mems	Command	Ask for more info
Instruction	Algorithm extraction	Checking			

The task in the third experiment consists of "put the pan on heater", "stir food circularly", "release the stirrer in a free space", and "turn the heater on". However, the robot does not know how accomplish the last subtask

It is clear that the robot does not know how to turn the heater on; hence, the involved components in this task are as Table 2.

Since the instruction in this task is not covered by the knowledge of the robot, when the learning block checks the availability of all subtasks, it understands that "turn the heater on" is missing. Hence, in the generation block a command will be generated according to the missing part. Based on the command the output block will ask the demonstrator about the unknown subtask.

In this experiment it is shown that when the robot is not able to accomplish a task, it would not produce any trajectory and asks the demonstrator for more demos or information.

All in all, to enable the robot to accomplish the given task, after receiving the message from the robot regarding the unknown subtask, the demonstrator starts to teach that part to the robot. It is clear that adjusting a heater knob is very important for cooking and as the robot gripper should execute the task with a high level of accuracy, the demonstrator taught this subtask using kinesthetic teaching (for more information regarding teaching phase, see the "task structure and demos" section).

As a conclusion, based on the acquired results, it is clear that how the proposed framework enables the robot to benefit from both levels of imitation and accomplishes tasks constructed by multiple subtasks in different levels. In addition, in the case of any ambiguity in demanded task, the robot can ask the demonstrator to provide more demos for the robot which is facilitated by this framework.

4 Conclusion

In this paper a new framework introduced which benefits from multimodal inputs including; visual input, motion sensation, and instruction. Using visual input and motion sensation enables the robot to learn through kinesthetic and vision-based teaching. In addition, using this framework, the robot can learn tasks in both levels

of imitation (high and low level). As OFTM is used for low-level imitation, the robot can learn trajectories in real-time and encode them compactly. On the other side, using Kragic's method for high level imitation enables the robot to reproduce a sequence of actions which can be different from seen action sequences. In this framework the generation block provides the ability of asking for more demos in the case that robot could not execute a task based on its knowledge. It means that this framework creates a mutual communication between demonstrator and imitator which is in the core of human robot interaction. The functionality of this framework is shown in different scenarios related to cooking tasks and implemented on a robotic arm. This study is somehow the first study on combining different levels of imitation in a unified framework for accomplishing complex tasks in robotics.

References

Bandera, J. P., Rodriguez, J. A., Molina-Tanco, L., & Bandera, A. (2012). A survey of vision-based architectures for robot learning by imitation. *International Journal of Humanoid Robotics, 9*(01), 1250006.

Bandura, A. (1969). Social-learning theory of identificatory processes. *Handbook of Socialization Theory and Research, 213*, 262.

Billard, A., Calinon, S., Dillmann, R., & Schaal, S. (2008). Robot programming by demonstration. *Springer handbook of robotics* (pp. 1371–1394). Berlin, Heidelberg: Springer.

Breazeal, C., Brooks, A., Gray, J., Hoffman, G., Kidd, C., Lee, H., et al. (2004). Humanoid robots as cooperative partners for people. *International Journal of Humanoid Robotics, 1*(2), 1–34.

Breazeal, C., Buchsbaum, D., Gray, J., Gatenby, D., & Blumberg, B. (2005). Learning from and about others: Towards using imitation to bootstrap the social understanding of others by robots. *Artificial Life, 11*(1–2), 31–62.

Breazeal, C., Takanishi, A., & Kobayashi, T. (2008). Social robots that interact with people. *Springer handbook of robotics* (pp. 1349–1369). Berlin, Heidelberg: Springer.

Cho, S., & Jo, S. (2012). Incremental motion learning through kinesthetic teachings and new motion production from learned motions by a humanoid robot. *International Journal of Control, Automation and Systems, 10*(1), 126–135.

Derimis, Y., & Hayes, G. (2002) Imitations as a dual-route process featuring predictive and learning components: A biologically plausible computational model. *Imitation in Animals and Artifacts*, 327–361.

Ekvall, S., & Kragic, D. (2006). Learning task models from multiple human demonstrations. In *ROMAN 2006. The 15th IEEE International Symposium on Robot and Human Interactive Communication* (pp. 358–363).

Falahi, M., & Jannatifar, M. (2013). Using orthogonal basis functions and template matching to learn whiteboard cleaning task by imitation. In *2013 3th IEEE International eConference on Computer and Knowledge Engineering* (*ICCKE*) (pp. 289–294).

Falahi, M., Mashhadian, M., & Tamiz, M. (2013). Learning complex trajectories by imitation using orthogonal basis functions and template matching. *Social robotics (ICSR), 2013 international conference on* (pp. 576–579). Berlin, Heidelberg: Springer.

Falahi, M., Shangari, T. A., Sheikhjafari, A., Gharghabi, S., Ahmadi, A., & Ghidary, S. S. (2014). Adaptive handshaking between humans and robots, using imitation: Based on gender-detection and person recognition. In *2014 Second RSI/ISM International Conference on Robotics and Mechatronics* (*ICRoM*) (pp. 936–941). IEEE.

Falahi, M., Sobhiyeh, S., Rezaie, A. H., & Motamedi, S. A. (2015). Wavelet based OFTM for learning stirring food by imitation. In *2015 International Conference on Advanced Robotics (ICAR)* (pp. 573–580). IEEE.

Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and Autonomous Systems, 42*(3), 143–166.

Kanda, T., Hirano, T., Eaton, D., & Ishiguro, H. (2004). Interactive robots as social partners and peer tutors for children: A field trial. *Human-Computer Interaction, 19*(1), 61–84.

Kim, E. S., Berkovits, L. D., Bernier, E. P., Leyzberg, D., Shic, F., et al. (2013). Social robots as embedded reinforcers of social behavior in children with autism. *Journal of Autism and Developmental Disorders*, 43(5), 1038–1049.

Kormushev, P., Nenchev, D. N., Calinon, S., & Caldwell, D. G. (2011). Upper-body kinesthetic teaching of a free-standing humanoid robot. In *2011 IEEE International Conference on Robotics and Automation* (*ICRA*) (pp. 3970–3975).

Mohammad, Y., & Nishida, T. (2009). Interactive perception for amplification of intended behavior in complex noisy environments. *AI & Society, 23*(2), 167–186.

Mohan, V., Morasso, P., Zenzeri, J., Metta, G., Chakravarthy, V. S., & Sandini, G. (2011). Teaching a humanoid robot to draw 'Shapes'. *Autonomous Robots, 31*(1), 21–53.

Mühlig, M., Gienger, M., Hellbach, S., Steil, J. J., & Goerick, C. (2009). Task-level imitation learning using variance-based movement optimization. In *IEEE International Conference on Robotics and Automation* (*ICRA'09*) (pp. 1177–1184).

Schaal, S. (1999). Is imitation learning the route to humanoid robots? *Trends in Cognitive Sciences, 3*(6), 233–242.

Tan, H. (2012). *Implementation of a framework for imitation learning on a humanoid robot using a cognitive architecture.* INTECH Open Access Publisher.

Part II
Mechanical Engineering and Power System

New Reactionless Spatial Grasper Design and Analysis

Dan Zhang and Bin Wei

Abstract A new reactionless spatial grasper is proposed and designed based on the principal vector linkage. In the literature, few dynamic balanced spatial grasp mechanisms can be found and here a new reactionless spatial grasper is designed. Through using the pantographs, the center of mass of the grasper mechanism is fixed at a still point, and through symmetrical structure design of the four fingers, it is moment balanced. The advantages of the proposed reactionless grasper mechanism and the design process are discussed and the principal dimensions are derived.

Keywords Reactionless grasper · Principal vector linkage · Symmetrical design

1 Introduction

Dynamic balancing for mechanisms has become an important part of mechanism design and development. When mechanisms move, due to the fact that the center of mass is not fixed and angular momentum is not constant, they often produce vibrations in the base, which greatly deteriorate the accuracy performance. Dynamic balancing can also be called shaking force balancing and shaking moment balancing, i.e. reactionless condition. Shaking force balancing can be achieved by making the center of mass of mechanism be fixed, i.e. linear momentum constant, normally the linear momentum is set to zero for the ease of analysis. Shaking moment balancing can be achieved by making the angular momentum constant, normally the angular momentum is set to zero.

Grasper mechanisms can be used in many industrial areas. There are two main types of grasper mechanisms, planar grasper mechanism and spatial grasper mechanism. Planar grasper mechanism can be put as follows: the fingers are in the same plane or planes that are parallel with each other (Dollar and Howe 2005a, b; Kragten 2011; Stavenuiter 2013). Spatial grasper mechanism can be put as follows: the fingers

D. Zhang · B. Wei (✉)
University of Ontario Institute of Technology, Oshawa, ON, Canada
e-mail: Bin.Wei@uoit.ca

are not in the same plane. The purpose of graspers is to grasp objects. Planar grasper mechanisms and spatial grasper mechanism can both grasp objects, but generally speaking, planar graspers can grasp regular shapes of objects, but for the very complex shapes of objects, spatial graspers is better than planar graspers in terms of steadiness and easiness. Most of the graspers in use nowadays are not dynamic balanced, and some grasper mechanisms are medical based (Remirez et al. 2014; Hanley and Rosario 2011), i.e. they can only use in medical arena, not in the industry areas. Since the grasper mechanism is not dynamic balanced, the whole system will swing and vibrate when the grasper mechanism is grasping object, which will severally affect the overall performance of the grasping process. For this reason reactionless spatial grasper mechanism is developed based on the pantograph and principle vector linkage. The pantographs are used to trace the center of mass of the whole system and let it be fixed at a still point, and the grasper mechanism is designed in such way that the four fingers are symmetric so that the shaking moment of each opposite symmetrical finger can be balanced out each other. This will be the first dynamic balanced spatial grasper mechanism. The advantage of grasper mechanism proposed here is that firstly it is dynamic balanced, secondly, it can grasp the complex objects steadily by using four symmetrical fingers.

This paper will propose a spatial grasper mechanism that is dynamic balanced and also it is gear towards the industry arena. It is known that the principal vector linkage architecture is force balanced, and through using the pantographs, the center of mass of the grasper mechanism is fixed at still point. Through symmetrical structure design of the four fingers, it is also moment balanced.

2 New Grasper Design

Figure 1 shows the reactionless spatial grasper mechanism. First it is force balanced because the center of mass of the whole system is at still point C by using the pantograph to trace the center of mass; it is moment balanced because it is
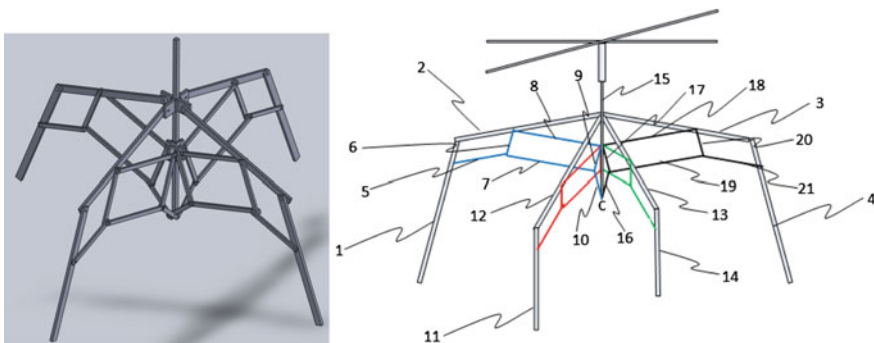


Fig. 1 Dynamic balanced grasper mechanism

symmetrical design, the shaking moment of each opposite leg can be balanced out each other. The grasper mechanism includes 8 principle links, 1, 2, 3, 4, the other perpendicular part 11, 12, 13, 14 and 24 pantograph links 5, 6, 7, 8, 9, 10, 16, 17, 18, 19, 20, 21 and their perpendicular counterparts as shown in Fig. 1. The links are all connected by revolute joint, and the center of mass of the grasper mechanism is at still point C which is connected by link 15 to the ceiling, this link 15 can move along the guide ways to be able to perform different locations of picking and placing which will be illustrated in the application section.

2.1 Principal Dimensions

The grasper mechanism can be decomposed into four segments as shown in Fig. 2, for the purpose of clearly showing the dimensions, the links are blown up. First we need to determine where to put the joint, in other words, we need to determine the principle dimensions, i.e. a1, b21 and c2 and the rest of the counterparts of other segments. The principle dimensions can be determined from the force balance condition (Wijk 2014). From the moment balance condition, we can also obtain the fact that the leg should be opposite and symmetrical, same as we intuitively design the legs. In order to have the force balance condition, we need to determine the linear momentum and make the linear momentum equal to zero. The linear momentum of the motion of link 1 are expressed as follows with respect to coordinate frame $x_0 y_0$ by fixing the link $P_1 B_1$, $B_1 C_1$ and $C_1 C$ (Wijk 2014), here we assume that the center of mass of each link is on the line that connects the joints.

$$
\begin{aligned}
&(m_2 + m_3 + m_4 + m_{18} + m_{19} + m_8 + m_{21} + m_{19} + m_{20})a_1 \\
&+ m_6 p_6 + m_9 p_9 + m_{16} p_{16} - m_1 b_1 = 0
\end{aligned}
\tag{1}
$$

Fig. 2 Segment of the grasper mechanism
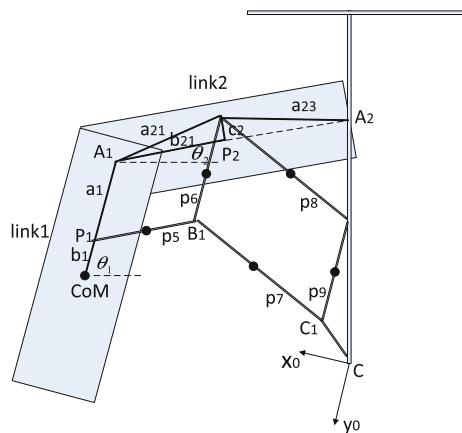
where m_i is the mass of link i, p_i is the distance from center of mass of link i to the connecting joint. From the above, the force balance condition can be derived as follows:

$$(m_2 + m_3 + m_4 + m_{18} + m_{19} + m_8 + m_{21} + m_{19} + m_{20})a_1$$
$$+ m_6 p_6 + m_9 p_9 + m_{16} p_{16} = m_1 b_1 \tag{2}$$

The dimension a_1 can be obtained as follows from the above force balance condition:

$$a_1 = \frac{m_1 s_1 - m_{12} p_{12} - m_{13} p_{13} - m_{14} p_{14}}{m_1 + m_2 + m_3 + m_4 + m_{21} + m_{22} + m_{31} + m_{41} + m_{42} + m_{43}} \tag{3}$$

The dimensions b_{21} and c_2 can be determined by using the equivalent linear momentum system (Wijk 2014) and from the theory of principle vector linkage, the dimensions b_{21} and c_2 can be derived as follows:

$$b_{21} = \frac{\left(\frac{a_{21} + p_2^a}{a_{21}} + \frac{p_2^b}{a_{23}}\right)\left(e_2 + \frac{p_2^b}{a_{23}} l_2\right)}{\left(\frac{a_{21} + p_2^a}{a_{21}} + \frac{p_2^b}{a_{23}}\right)^2} \tag{4}$$

$$c_2 = \frac{\left(\frac{a_{21} + p_2^a}{a_{21}} + \frac{p_2^b}{a_{23}}\right)\left(f_2 - \frac{q_2^b}{a_{23}} l_2\right)}{\left(\frac{a_{21} + p_2^a}{a_{21}} + \frac{p_2^b}{a_{23}}\right)^2} \tag{5}$$

where

$$a_{21} = \frac{m_2 p_2^a - m_5 p_5}{m_1}, a_{23} = \frac{m_2 p_2^b - m_{18} p_{18} - m_{19} p_{19}}{m_3 + m_4 + m_{21} + m_{20}},$$

and s_1 is the distance from the first joint A_1 to the CoM of link 1, e_2 is the distance from the first joint A_1 to the CoM of link 2 along the line that connects the joint A_1 and A_2, f_2 is the distance from the CoM of link 2 to the line that connects the joint A_1 and A_2. l_2 is the opposite side of e_2. Because it is symmetrical, the other principle dimension of the rest segments can be calculated in the same manner.

2.2 Symmetrical Design

For the moment balance, we need to first write the angular momentum. Figure 3a, b shows the separate principle vector linkages. The angular momentum about the center of mass of the grasper mechanism C can be written as follows:
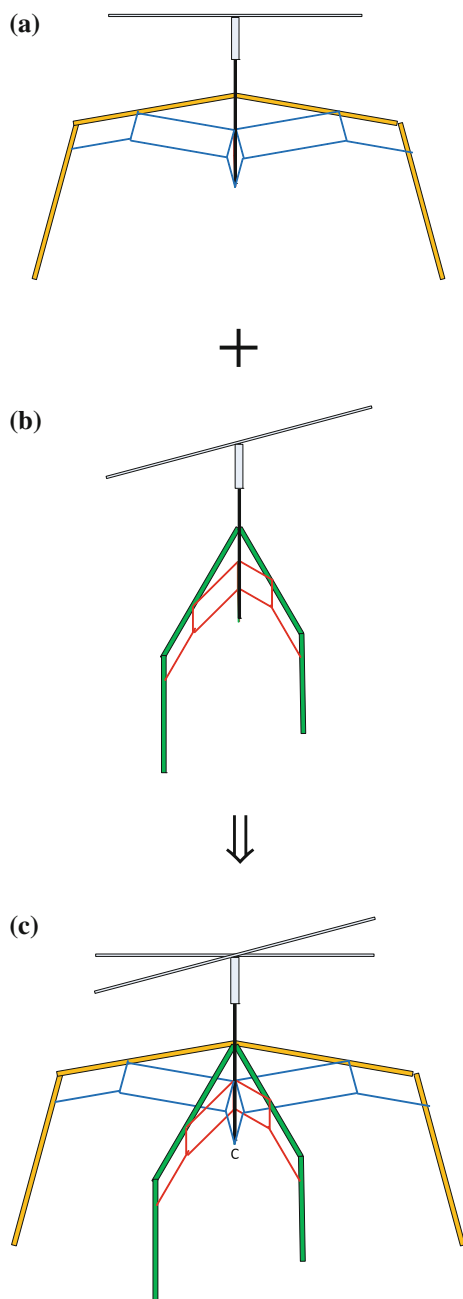
Fig. 3 Principle vector linkages and symmetrical design

$$
\begin{aligned}
A ={}& (I_1 + I_6 + I_9 + I_{16})\dot{\theta}_1 + (I_2 + I_5 + I_{18} + I_{19})\dot{\theta}_2 \\
&+ (I_3 + I_7 + I_8 + I_{21})\dot{\theta}_3 + (I_4 + I_{10} + I_{18} + I_{20})\dot{\theta}_4 \\
&+ m_1(r_1 \times \dot{r}_1) + m_2(r_2 \times \dot{r}_2) + m_3(r_3 \times \dot{r}_3) + m_4(r_4 \times \dot{r}_4) \\
&+ m_5(r_5 \times \dot{r}_5) + m_6(r_6 \times \dot{r}_6) + m_9(r_9 \times \dot{r}_9) + m_{16}(r_{16} \times \dot{r}_{16}) \\
&+ m_{18}(r_{18} \times \dot{r}_{18}) + m_{19}(r_{19} \times \dot{r}_{19}) + m_8(r_8 \times \dot{r}_8) + m_7(r_7 \times \dot{r}_7) \\
&+ m_{21}(r_{21} \times \dot{r}_{21}) + m_{20}(r_{20} \times \dot{r}_{20}) + m_{17}(r_{17} \times \dot{r}_{17}) + m_{10}(r_{10} \times \dot{r}_{10})
\end{aligned}
\tag{6}
$$

where I_i is the inertia about its center of mass of link i, r_i is the position vector of center of mass of link i relative to the center of mass of grasper mechanism C, and $\dot{\theta}_i$ is the angle from horizontal axis to the center line of link i.

Secondly substitute the position vectors of center of mass of each link relative to the center of mass of the whole mechanism in point C, its position vector derivatives and the above force balance conditions to the angular momentum equation to obtain the final form of the angular momentum, then make the angular momentum equal to zero to determine the moment balance condition, from the moment balance condition, we can derive that links 1 and 4 are symmetrical and have opposite motions, links 2 and 3 are symmetrical and have opposite motions as follows.

$$\dot{\theta}_1 + \dot{\theta}_4 = 0$$

$$\dot{\theta}_2 + \dot{\theta}_3 = 0$$

$$\theta_1 + \theta_4 = \pi$$

$$\theta_2 + \theta_3 = \pi \tag{7}$$

By vertically combining Fig. 3a, b, symmetrical design can be derived as shown in Fig. 3c.

3   Application

The proposed dynamic balanced grasper mechanism can be used in many areas, such as industrial picking and placing, space, etc. The following figure shows the general concept of grasping process in an industrial area. Because it is dynamic balanced, the whole system can be hanged by a single wire if the grasper is used in space which can make the system lighter. Future work will include designing the proposed grasper mechanism to be foldable, when the grasper mechanism is used in space, the grasper mechanism can be folded when it is sending into space and unfold it when it is in operation, and due to it is dynamic balanced and there is no gravity in space, the grasper mechanism can be hanged by a single wire rather than a rigid link in some special situations (Fig. 4).
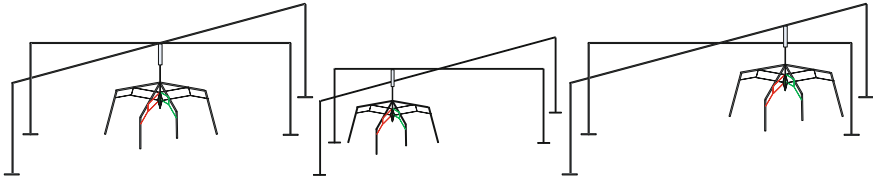
Fig. 4 Grasper mechanism used in industrial area

4 Conclusion

In this paper, a reactionless spatial grasper mechanism is proposed and designed based on the principal vector linkage. It is force balanced because the center of mass of the whole system is at still point by using the pantograph to trace the center of mass; it is moment balanced because it is symmetrical design, the shaking moment of each opposite leg can be balanced out each other. The principal dimensions and the symmetrical design are derived based on the force and moment balance conditions. In the literature, few reactionless spatial grasp mechanism can be found and here a reactionless spatial grasper mechanism is designed. The advantages of the proposed grasper mechanism are that it is dynamic balanced and it can grasp complex objects steadily by using four symmetrical fingers. The proposed dynamic balanced grasper mechanism has potential to use in many areas, such as industrial picking and placing and space.

References

Dollar, A. M., & Howe, R. D. (2005a). Design and evaluation of a robust compliant grasper using shape deposition manufacturing. In *Proceedings of ASME International Mechanical Engineering Congress and Exposition, USA*.

Dollar, A. M., & Howe, R. D. (2005b). Towards grasping in unstructured environments: Grasper compliance and configuration optimization. *Advanced Robotics, 19*(5), 523–543.

Hanley, H., Rosario, M., & Chen, Y., et al. (2011). Design and testing of a three fingered flexural laparoscopic grasper. In *Proceedings of the 2011 Design of Medical Devices Conference, USA*.

Kragten, G. (2011). Underactuated hands fundamentals, performance analysis and design (Thesis). Delft University of Technology.

Remirez, A. A., Lathrop, R. A., Russell III, P. T., & Webster III, R. J. (2014). Design of a stiff steerable grasper for sinus surgery. In *Proceedings of the 2014 Design of Medical Devices Conference, USA*.

Stavenuiter, R. A. J. (2013). Design of an underactuated grasper with adjustable compliance (Thesis). Delft University of Technology.

Wijk, V. (2014). Methodology for analysis and synthesis of inherently force and moment-balanced mechanisms—Theory and applications (Thesis). University of Twente.

Tracking and Vibration Control of a Carbon Nanotube Reinforced Composite Robotic Arm

Mohammad Azadi and Behzad Hasanshahi

Abstract This research deals with a study of tracking and vibration control of a functionally graded carbon nanotube reinforced composite robotic arm which crosses pre-specified path. The Euler-Bernoulli beam theory and Lagrange-Rayleigh-Ritz technique are employed to derive the governing equations of CNTRC robotic arm. The robotic arm is subjected to a follower force; also, in order to control the vibration, the piezoelectric layers have been used beside the active control system. The robotic arm is simulated and the results show the effectiveness of the controller algorithm.

Keywords FG-CNTRC · Robotic arm · Active vibration control · Follower force · Piezoelectric layers

1 Introduction

Considerable mechanical, thermal and electrical properties of carbon nanotubes (CNT), has attracted a great number of researcher's attention. While the high strength, stiffness, and aspect ratio, the density of this material is very low. Ke et al. (2010) investigated the nonlinear free vibration of functionally graded carbon nanotube-reinforced composite beam; they used a direct iterative method to obtain the nonlinear vibration frequencies. Formica et al. (2010) used an equivalent continuum model based on the Eshelby–Mori–Tanaka approach to study the vibration of carbon nanotube-reinforced composites. An analytical solutions for bending, buckling and vibration responses of FG-CNTRC beams which have an elastic

M. Azadi (✉)
Department of Mechanical Engineering, Marvdasht Branch,
Islamic Azad University, Marvdasht, Iran
e-mail: mazadi@shirazu.ac.ir

B. Hasanshahi
Department of Mechanical Engineering, Arsanjan Branch,
Islamic Azad University, Arsanjan, Iran

265

foundation is presented by Wattanaskulpong and Ungbhakorn (2013); several shear deformation theories are employed to study the bending, buckling and vibration problem. According to the Timoshenko beam theory along with Von-Karman geometric nonlinearity, Ansari et al. (2014) are studied the forced vibration behavior of nano-composite beams reinforced by single-walled carbon nanotubes (SWCNTs). Rokni et al. (2015) considered the free vibration of FG-CNTRC beam based on the modified couple stress theory. They proposed a new CNT dispersion function along the micro-beam length to maximize the fundamental natural frequency of the MWCNT/epoxy composite beams. On the other hand, the flutter vibration in the momentous structures is very destructive, thereby steep decline of vibration's amplitude seems serious issue in the different industries. Rafiee et al. (2013) investigated the influences of a temperature change and an applied voltage on the free vibration of FG-CNTRC beams with surface-bonded piezoelectric layers. Free vibration behavior of FG-CNTRC cylindrical panel embedded in piezoelectric layers with simply supported boundary conditions is studied by Alibeigloo (2014). Wu and Chang (2014) considered the stability of carbon nanotube reinforced composite plates with surface-bonded piezoelectric layers which subjected to bi-axial compression. Khorshidi et al. (2015) studied active vibration control analysis of a circular plate coupled with piezoelectric layers on both sides; the effect of different piezoelectric patch thickness and host patch thickness on the plate natural frequencies are discussed. Li and Lyu (2014) used the piezoelectric actuator and sensor pairs to investigate the active vibration control of lattice sandwich beams; by taking advantage from Hamilton's principle with the assumed mode method, they developed the dynamical model of the structural system.

Nowadays robotic arms became an integral part of the human's life due to their accuracy, efficiency, speed, ability to work in hazardous environments which are unsafe and inhospitable for humans; and also free from human limitations like the need to eat or sleep. Because of mechanical nature of robotic arms and computerized control, they can carry out a repetitive task with great precision in the dangerous and toxic situations such as car and aerospace industries, and steel plants. Hence, studying on the vibration control of the robotic arms and utilize the new composite material for industrial robotic arms will become justifiable as the main purpose of this research.

In this study, a robotic arm, which has been reinforced by carbon nanotubes, has been considered. An active control law is used to control both position and vibrations of this robotic arm. This robotic system has been simulated and simulation results show the high performance of this system.

2 Governing Equations of the Robotic System

Figure 1 shows an arm with prismatic joint. This arm have the length L and thickness h which subjected to follower force P. This arm has been bonded by the piezoelectric actuators/sensors patches. The arm made of the polymeric matrix and

Fig. 1 Schematic view of the CNTRC robotic arm
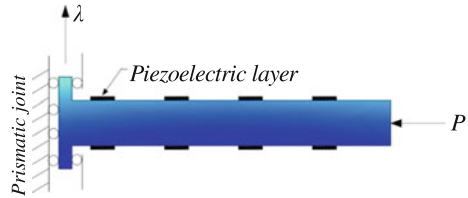
single-walled carbon nanotubes (SWCNTs) as a reinforcement that are assumed to be aligned in axial direction and functionally graded in the thickness direction of the arm. The arm has been considered as a beam.

To estimate the effective material properties of FG-CNTRC arm, The rule of mixture has been used; though, the effective Young's modulus and shear modulus of CNTRC beam can be predicted by Ke et al. (2010):

$$
\begin{aligned}
E_{11} &= \eta_1 V_{cnt} E_{11}^{cnt} + V_m E^m \\
\frac{\eta_2}{E_{22}} &= \frac{V_{cnt}}{E_{22}^{cnt}} + \frac{V_m}{E^m} \\
\frac{\eta_3}{G_{12}} &= \frac{V_{cnt}}{G_{12}^{cnt}} + \frac{V_m}{G^m}
\end{aligned}
\tag{1}
$$

where E_{11}^{cnt}, E_{22}^{cnt} and G_{12}^{cnt} are Young's and shear modulus of SWCNTs, respectively. E^m and G^m represent the corresponding properties of the isotropic matrix; η_j ($j = 1$, 2, 3) is the CNT efficiency parameters which considers the size-dependent material properties and will be calculated later by matching the elastic modulus of CNTRCs observed from the MD simulation results with the numerical results obtained from the rule of mixture. V_{cnt} and V_m are the volume fraction of carbon nanotube and matrix, respectively, which are related by $V_{cnt} + V_m = 1$. Furthermore, Poisson's ratio ν and mass density ρ of the composite beam can be presented as:

$$
\begin{aligned}
\nu &= V_{cnt} \nu^{cnt} + V_m \nu^m \\
\rho &= V_{cnt} \rho^{cnt} + V_m \rho^m
\end{aligned}
\tag{2}
$$

where ν^{cnt} and ν^m are Poisson's ratios of the carbon nanotube and the matrix, respectively; also, ρ^{cnt} and ρ^m are the mass density of the reinforcement and polymeric matrix.

The volume fraction of CNTs $\left(V_{cnt}^*\right)$ is calculated as:

$$
V_{cnt}^* = \frac{\Lambda_{cnt}}{\Lambda_{cnt} + (\rho^{cnt}/\rho^m) - (\rho^{cnt}/\rho^m)\Lambda_{cnt}}
\tag{3}
$$

Λ_{cnt} is the mass fraction of carbon nanotubes.

Based on the distribution functions of CNTs, the volume fractions V_{cnt} are defined for different distribution geometry pattern of CNTs as follows:

$$\begin{cases} V_{cnt} = V_{cnt}^* & UD \\ V_{cnt} = 4\frac{|z|}{h}V_{cnt}^* & FG-X \\ V_{cnt} = 2\left(1 - 2\frac{|z|}{h}\right)V_{cnt}^* & FG-O \end{cases} \tag{4}$$

To obtain the governing equation of motion, calculating the kinetic and potential energy is necessary; the kinetic and potential energy of the system are included the aggregate of the kinetic and potential energy of each part which presented as:

$$\begin{aligned} T &= T_{arm} + T_p \\ U &= U_{arm} + U_p \end{aligned} \tag{5}$$

where the subscripts of *arm*, and *p* refer to arm, and piezoelectric layers, respectively. The kinetic energies can estimated as follow:

$$T_{arm} = \frac{1}{2}\int_0^L \int_{-\frac{h}{2}}^{\frac{h}{2}} b\rho(z)\left(\dot{\lambda} + \dot{w}\right)^2 dzdx$$

$$T_p = \frac{1}{2}\sum_{n=1}^N \int_0^L \iint_{A_n} \rho_{p_n}(\dot{\lambda} + \dot{w})^2[H(x - x_{1_n}) - H(x - x_{2_n})]dA_ndx \tag{6}$$

where $w(x,t)$ is the displacements along the z direction, and $\lambda(z)$ is the motion along the z direction. In addition, the superscript of dot denotes a partial derivative with respect to time. In the above equations, N is the number of piezoelectric layers, ρ_{p_n} is the density of the nth piezoelectric layer, $H(x)$ is Heaviside function, also, the coordinates of the two ends of the nth piezoelectric layer is shown by x_{1_n} and x_{2_n}. Similarly for potential energy:

$$U_{arm} = \frac{1}{2}\int_v Q_{11}(z)\cdot\varepsilon_x^2 dv + \frac{1}{2}\int_0^L P\left(\frac{\partial w(x,t)}{\partial x}\right)dx$$

$$U_p = \frac{1}{2}\sum_{n=1}^N \int_{PZT_n} E_{p_n}z^2\left(\frac{\partial^2 w_{p_n}}{\partial x^2}\right)^2 dv + \sum_{n=1}^N \int_{PZT_n} ze_{31}E_{z_n}\frac{\partial^2 w_{p_n}}{\partial x^2}dv + \frac{1}{2}\sum_{n=1}^N \int_{PZT_n} E_{z_n}d_ndv \tag{7}$$

In the potential energy equations, ε_x is the strain in the x direction, P is the follower force, the electric energy stored in the piezoelectric material calculated in the last term of the piezoelectric potential energy equation, also, $Q_{11}(z)$ and the electric displacement of the nth piezoelectric pair, d_n are defined as below:

$$Q_{11}(z) = \frac{E_{11}^{cnt}(z)}{1 - v^2(z)}$$

$$d_n = \varepsilon_{pn} \frac{v_n}{h_{pn}}$$

(8)

where in the electric displacement equation, ε_p is the dielectric constant of the piezoelectric material. In the next step by using the Lagrange theory, the governing equations of motion will be obtained.

$$\frac{d}{dt}\left(\frac{\partial(T - U)}{\partial \dot{\psi}}\right) - \frac{\partial(T - U)}{\partial \psi} = Q$$

(9)

where Q is the vector of generalized forces and ψ is the vector of the generalized coordinates.

The Rayleigh-Ritz technique is employed in the present study to derive the governing equations of motion for linear forced vibration of FG-CNTRC Euler-Bernoulli arm. According to this theory, the transverse displacement will be replaced with the trial shape functions and time-dependent generalized coordinates parameters which are shown by ϕ_i and q_i, respectively.

$$w(x, t) = \sum \phi_i q_i = \mathbf{\Phi}^T \mathbf{q}$$

(10)

where $\mathbf{\Phi}^T$ and $\mathbf{q}$ are the vectors of the trial shape functions and time-dependent generalized coordinates, respectively. The general shape of governing equations will be obtained as follows:

$$\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} \ddot{\lambda} \\ \ddot{q} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{K}_b + \mathbf{K}_p + \mathbf{K}_w \end{bmatrix} = M\ddot{\psi} + K\psi = F$$

(11)

$$\mathbf{v}_s = K_{p\,elect}^{-1} K_{p\,elastelect_s}^{-1} q$$

where M and K are the mass and stiffness matrices, respectively.

$$F = \begin{bmatrix} F_s \\ -K_{p\,elastelect\,a} \mathbf{v}_a \end{bmatrix}$$

(12)

$$K_{p\,elatelect_a} \; or \; K_{p\,elatelect_s} = \begin{bmatrix} K_{pee_1} K_{pee_2} \ldots K_{pee_n} \end{bmatrix}$$

$$K_{p\,elect} = \sum_{n=1}^{N} \int_{PZT_n} \varepsilon_{pn} p_n p_n^T dv$$

(13)

where F_s is the controller load which moves the arm. $\mathbf{v}_a$ and $\mathbf{v}_s$ are the voltages of the piezoelectric actuators and sensors, respectively. $K_{P\,elatelect_a}$ and $K_{P\,elatelect_s}$ denote the matrices of the elastic–electric effect of the piezoelectric actuator and

sensor layers, respectively. In addition, K_{pelect} is a diagonal capacitance matrix of the piezoelectric patches. p_n is a $N \times 1$ vector which the entry for n is $\frac{1}{h_{pn}}$, while the other entries are zero.

$$K_{peen} = \frac{e_{31n}}{h_{pn}} \int\limits_{PZT_n} z \frac{\partial^2 \Phi}{\partial x^2} dv \tag{14}$$

where K_{pee_n} is the vector of nth column of the elastic–electric matrix.

The parameters of the mass matrix are:

$$\mathbf{M}_{11,arm} = \int\limits_0^L \int\limits_{\frac{-h}{2}}^{\frac{h}{2}} b\rho(z) dz dx + \sum_{n=1}^N \int\limits_{x_{1_n}}^{x_{2_n}} \rho_{pn} b_{pn} h_{pn} dx$$

$$\mathbf{M}_{12,arm} = \int\limits_0^L \int\limits_{\frac{-h}{2}}^{\frac{h}{2}} b\rho(z) \mathbf{\Phi}^T dz dx + \sum_{n=1}^N \int\limits_{x_{1_n}}^{x_{2_n}} \rho_{pn} b_{pn} h_{pn} \mathbf{\Phi}^T dx$$

$$\mathbf{M}_{21,arm} = \int\limits_0^L \int\limits_{\frac{-h}{2}}^{\frac{h}{2}} b\rho(z) \mathbf{\Phi} dz dx + \sum_{n=1}^N \int\limits_{x_{1_n}}^{x_{2_n}} \rho_{pn} b_{pn} h_{pn} \mathbf{\Phi} dx \tag{15}$$

$$\mathbf{M}_{22,arm} = \int\limits_0^L \int\limits_{\frac{-h}{2}}^{\frac{h}{2}} b\rho(z) \mathbf{\Phi}\mathbf{\Phi}^T dz dx + \sum_{n=1}^N \int\limits_{x_{1_n}}^{x_{2_n}} \rho_{pn} b_{pn} h_{pn} \mathbf{\Phi}\mathbf{\Phi}^T dx$$

and the stiffness matrices are:

$$\mathbf{K}_b = \int\limits_0^L P \sum_{i=1}^Y \sum_{j=1}^Y \frac{\partial \phi_i(x)}{\partial x} \frac{\partial \phi_j(x)}{\partial x} dx$$

$$+ \int\limits_0^L \int\limits_{\frac{-h}{2}}^{\frac{h}{2}} b z^2 Q_{11}(z) \left[\sum_{i=1}^Y \sum_{j=1}^Y \frac{\partial^2 \phi_i(x)}{\partial x^2} \frac{\partial^2 \phi_j(x)}{\partial x^2} \right] dz dx$$

$$\mathbf{K}_p = \sum_{n=1}^N \iiint\limits_{PZT_n} E_{pn} z^2 \sum_{i=1}^Y \sum_{j=1}^Y \frac{\partial^2 \phi_i(x)}{\partial x^2} \frac{\partial^2 \phi_j(x)}{\partial x^2} dv \tag{16}$$

$$\mathbf{K}_w = -P \sum_{i=1}^Y \sum_{j=1}^Y \phi_i(x) \frac{\partial \phi_j(x)}{\partial x} \bigg|_{x=L}$$

3   Control Design

Mechanical engineering intuition dictates the idea of canceling nonlinear terms and decoupling the dynamics of the system by using an inverse dynamics control of the form

$$F = Mu_0 + K\psi + f_d \tag{17}$$

which applied to Eq. (11), and in view of the regularity of matrix M that is a symmetric positive definite matrix which verifies

$$\lambda_m I \leq M \leq \lambda_M I \tag{18}$$

where $\lambda_m (\lambda_M < \infty)$ denotes the strictly positive minimum (maximum) eigenvalue of M, yields a set of n decoupled linear systems.

$$\ddot{\psi} = u_0 \tag{19}$$

where u_0 is an auxiliary control input to be designed. Typical choice for u_0 is:

$$u_0 = \ddot{\psi}_d + K_D(\dot{\psi}_d - \dot{\psi}) + K_P(\psi_d - \psi) \tag{20}$$

Herein the subscript d denotes the desired one. The error equation is

$$\ddot{\tilde{\psi}} + K_D \dot{\tilde{\psi}} + K_P \tilde{\psi} = 0 \tag{21}$$

where $\tilde{\psi} = \psi - \psi_d$. The error equation (21) is exponentially stable by a suitable choice of the matrixes K_P and K_D. To see this, we can rewrite Eq. (21) in its state space form, i.e., $\dot{\zeta} = A\zeta$. It is thus easy to show that if K_P and K_D are suitably chosen then the matrix A is stable. Then we can fine, for some symmetric Q > 0, a symmetric positive definite matrix P satisfying $A^T P + PA = -Q$. Global asymptotic Lyapunov stability follows. For any initial position the error will asymptotically converge to zero. The transient behavior will thus be symmetric and independent of the operation condition. The matrix gains K_P and K_D are adjusted once for all.

4   Simulation Results

In this paper, a functionally graded carbon nanotube reinforced composite robotic arm has been studied. The piezoelectric layers have been used as sensors and actuators. The governing equation of motion has been obtained, using

Lagrange-Rayleigh-Ritz technique and Euler-Bernoulli beam theory. An inverse dynamics control algorithm was used to not only control the position of the robotic arm but also damp the induced vibrations of the arm. Figure 2 shows the motion of the arm. According to this figure the robotic arm track the desired trajectory completely.

Figure 3 compares the vibration of the arm in two cases, with and without controller. It is clear that by applying voltages to the actuators the vibrations of the arm have been suppressed rapidly. The time response of the generalized coordinates have been illustrated in Fig. 4. Finally the simulation results show the high performances of the controller. For all cases the distribution patterns of carbon nanotubes are FG-X and 12 % for the volume fraction of CNTs are considered.
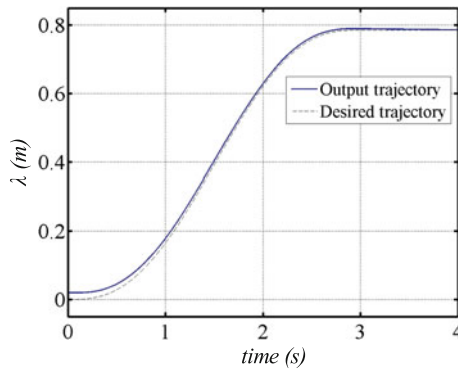
Fig. 2 Prismatic trajectory tracking of the robotic arm
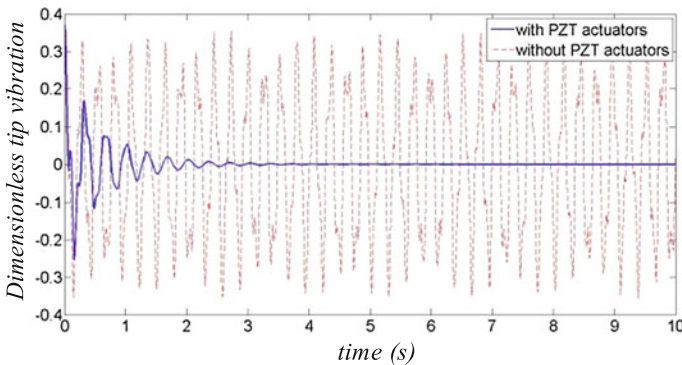
Fig. 3 Dimensionless tip vibration of the robotic arm

Fig. 4 Time history of the arm generalized coordinates. **a** First generalized coordinate, **b** second generalized coordinate, **c** third generalized coordinate
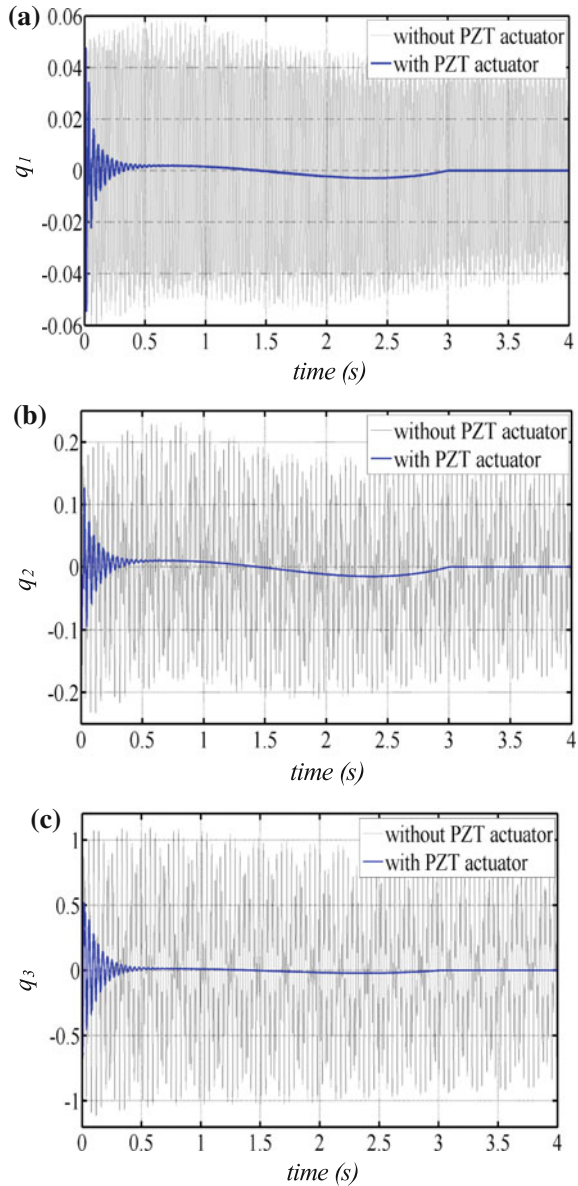
5 Conclusion

In this study, the robotic arm was tracking the pre-specified path; while this robotic structure had the motion under a specific function and it was subjected to Follower force, the vibration of the FG-CNTRC beam had been suppressed by piezoelectric

layers which used as actuators and sensors. As the results are illustrated in the previous section, the vibrations of the robotic arm were damped instantly; as can be seen, the piezoelectric layers had an important role in controlling the amplitude of flutter vibration which seems the crucial issue to consider in application of robotic arms in various industries.

References

Alibeigloo, A. (2014). Free vibration analysis of functionally graded carbon nanotube-reinforced composite cylindrical panel embedded in piezoelectric layers by using theory of elasticity. *European Journal of Mechanics A/Solids, 44*, 104–115.

Ansari, R., FaghihShojaei, M., Mohammadi, V., Gholami, R., & Sadeghi, F. (2014). Nonlinear forced vibration analysis of functionally graded carbon nanotube-reinforced composite Timoshenko beams. *Composite Structures, 113*, 316–327.

Formica, G., Lacarbonara, W., & Alessi, R. (2010). Vibrations of carbon nanotube-reinforced composites. *Journal of Sound and Vibration, 329*, 1875–1889.

Ke, L. L., Yang, J., & Kitipornchai, S. (2010). Nonlinear free vibration of functionally graded carbon nanotube reinforced composite beams. *Composite Structures, 92*, 676–683.

Khorshidi, K., Rezaei, E., Ghadimi, A. A., & Pagoli, M. (2015). Active vibration control of circular plates coupled with piezoelectric layers excited by plane sound wave. *Applied Mathematical Modeling, 39*, 1217–1222.

Li, F. M., & Lyu, X. X. (2014). Active vibration control of lattice sandwich beams using the piezoelectric actuator/sensor pairs. *Composites: Part B, 67*, 571–578.

Rafiee, M., Yang, J., & Kitipornchai, S. (2013). Large amplitude vibration of carbon nanotube reinforced functionally graded composite beams with piezoelectric layers. *Composite Structures, 96*, 716–725.

Rokni, H., Milani, A. S., & Seethaler, R. J. (2015). Size-dependent vibration behavior of functionally graded CNT-Reinforced polymer micro cantilevers: Modeling and optimization. *European Journal of Mechanics—A/Solids, 49*, 26–34.

Wattanasakulpong, N., & Ungbhakorn, V. (2013). Analytical solutions for bending, buckling and vibration responses of carbon nanotube-reinforced composite beams resting on elastic foundation. *Computational Materials Science, 71*, 201–208.

Wu, C. P., & Chang, S. K. (2014). Stability of carbon nanotube-reinforced composite plates with surface-bonded piezoelectric layers and under bi-axial compression. *Composite Structures, 111*, 587–601.

Synthesis and Analysis of Pneumatic Muscle Driven Parallel Platforms Imitating Human Shoulder

Xingwei Zhao, Bin Zi and Haitao Liu

Abstract The shoulder joint typically refers to the glenohumeral joint, which is surrounded by the rotator cuff muscles. The muscles and joints allow the shoulder to move through a remarkable range of motion, making it one of the most mobile joints in the human body. In this paper, three pneumatic muscle driven parallel platforms are synthesized for imitating the human shoulder. A hybrid tension-stiffness index is carried out and utilized for dimensions optimization of the mechanisms. Based on the index, the kinematic characteristics of the three mechanisms are investigated and compared to one another. Overall, the mechanism III has the maximum wrench-closure workspace, as well as the best tension-stiffness factor among the wrench-closure workspace.

Keywords Human shoulder · Pneumatic muscle · Parallel mechanism · Hybrid tension-stiffness index

1 Introduction

The shoulder joint typically refers to the glenohumeral joint, which is surrounded by the rotator cuff muscles. The muscles and joints allow the shoulder to move through a remarkable range of motion, making it one of the most mobile joints in

X. Zhao (✉)
Mechatronics and Machine Dynamics, Technical University of Berlin,
10587 Berlin, Germany
e-mail: zhaoxingwei001@gmail.com

B. Zi
School of Mechanical and Automotive Engineering, Hefei University of Technology,
Hefei 230009, China
e-mail: binzi.cumt@163.com

H. Liu
School of Mechanical Engineering, Tianjin University, Tianjin 300072, China
e-mail: liutju@gmail.com

the human body. In most of the previous literatures, the shoulder joint usually contains a spherical joint. Yang et al. (2005) designed a 7 degree-of-freedoms (DOFs) cable-driven humanoid arm, in which the shoulder module is an upper link connecting the lower link through a ball joint and several driving cables. A flexible shoulder structure was developed by Sodeyama et al. (2005), which consists of a shoulder blade connected to an arm through a boll-joint. Lenarčić et al. (2000, 2003) suggested that the humanoid shoulder is an extremely sophisticated system which is composed of two subsystems: a parallel mechanism which serves as the shoulder girdle; and a serial mechanism which serves as the glenohumeral joint.

Although most of the exiting robotic shoulder possess some fundamental features of a human shoulder joint, their mechanism designs as well as their driving schemes are still quite different from that of a human shoulder. The goal of this work is to synthesize pneumatic muscle driven parallel platforms (PMDPPs) that present the structure of the human shoulder. Three PMDPPs are synthesized. All of them have three rotational DOFs and are driven by four pneumatic muscles. Due to the advantages of pneumatic muscles, such as their high mass/weight ratio, their high flexibility, and their similarities to biological muscles (Zhao and Zi 2013), the PMDPPs can imitate the shoulder motion excellently.

In order to optimize the dimensions of the three PMDPPs and compare them to each other, an index is required to quantify the kinematic property of the PMDPPs. On the one hand, pneumatic muscles have a unilateral property similar to cables (i.e. they can pull, but they cannot push moving platforms), which makes it possible to transfer many of the concepts and methods used for cable-driven parallel mechanisms (CDPMs) to PMDPPs. Recently, cable-driven parallel mechanisms have received a lot of attention in robotics research and literature (Zi et al. 2008). Tadokoro et al. (2002) analyzed the reachable workspace for a 6-DOF parallel cable-driven architecture, and an optimal fundamental mechanical design is performed from the viewpoint of kinematics. Ebert-Uphoff et al. (2004) have reviewed some of the connections between CDPMs and grasping, and the transfers of the antipodal grasping theorem to planar cable-driven robots. Gouttefarde and Gosselin (2004, 2006) analyzed the wrench-closure workspace of planar 3-DOF parallel mechanisms driven by four cables. A tension factor was proposed by Pham et al. (2006, 2009), which could reflect the relative tension distribution among the driving cables for CDPMs. Whereas Hiller et al. (2005) analyzed the stiffness of CDPMs. He suggested that when some eigenvalue of the stiffness matrix vanishes, the mechanism is singular.

On the other hand, pneumatic muscles have limited workspace compared to cables, and pneumatic muscles usually co-operate with restrained joints and form rotational mechanisms. For those mechanisms, both unidirectional constraints (pneumatic muscles) and full constraints (restrained joints) work together on the platform. Therefore, the wrench-closure workspace, the tension factor, as well as the stiffness of CDPMs cannot be directly used for PMDPPs. In this paper, both the tension factor and the stiffness factor of the PMDPPs are proposed. While combining the tension factor and the stiffness factor, a hybrid tension-stiffness index is investigated. A global hybrid tension-stiffness index is carried out for the

dimensions optimization of the three PMDPPs. In addition, the kinematic characteristics of these PMDPPs are investigated and compared to one another after optimization.

2 System Description

The shoulder is mobile with a wide range of actions from the arms, as well as stable to allow for actions such as lifting, pushing and pulling. The shoulder joint is known as the glenohumeral joint, is the major joint of the shoulder shown in Fig. 1. It allows the arm to rotate in a circular fashion or to hinge out and up away from the body. The rotator cuff is a group of tendons and muscles surrounded by a shoulder joint, connecting the humerus to the scapula. The rotator cuff tendons provide stability to the shoulder, and the muscles allow the shoulder to rotate. The muscles in the rotator cuff include: the teres minor, the infraspinatus, the supraspinatus and the subscapularis.

According to the glenohumeral joint's biological structure, mechanism I is designed as a bottom platform connecting to an up platform through a spherical joint and four pneumatic muscles, as shown in Fig. 2(1). However, the maximum swing angle of the spherical joint is $\pm 30°$, which is less than the required workspace of the shoulder joint. Some alternatives to the spherical joint are given in Table 1, where three revolute joints are utilized instead of the spherical rolling joint. There are in total $A_3^3 = 6$ types of combinations. However, the mechanism IV is similar to the mechanism II, and the mechanisms V, VI and VII are similar to the mechanism III. Therefore, in this paper, only mechanism I, mechanism II and mechanism III are analyzed, as shown in Fig. 2. In Fig. 2(2), mechanism II is a bottom platform connecting to an up platform through four pneumatic muscles and three revolute joints, in terms of Rx, Rz, and Ry, successively. Rx, Ry and Rz represent revolute joints whose axes of rotation are x axis, y axis and z axis of their
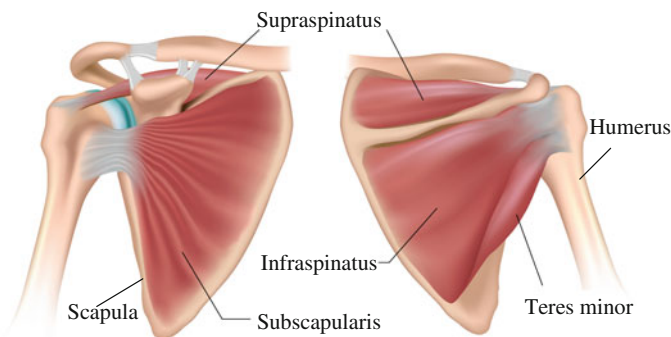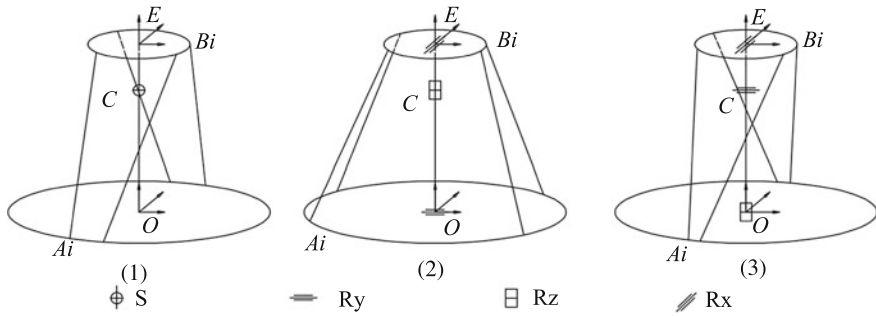

Fig. 1 Structure of the human shoulder

Fig. 2 Kinematic scheme of the mechanism I, II and III

Table 1 Synthesis of PMDPPs

Name	I	II	III	IV
Joint type	S	Rx-Rz-Ry	Rz-Rx-Ry	Ry-Rz-Rx
Name	V	VI	VII	
Joint type	Rx-Ry-Rz	Ry-Rx-Rz	Rz-Ry-Rx	

S represents a spherical joint, Rx, Ry and Rz represent revolute joints whose axes of rotation are x axis, y axis and z axis of their local coordinates

local coordinates. In Fig. 2(3), mechanism III is a bottom platform connecting an up platform through four pneumatic muscles and three revolute joints, in terms of Rz, Rx and Ry, successively.

3   Kinematic Analysis

In order to describe the mechanism conveniently, some labels for each joint and geometric centers are given as follows. The geometric center of a bottom platform is point O, the geometric center of an up platform is point E. Point C is located between point O and E. The ith pneumatic muscle is connected to the bottom platform at point Ai and with the up platform at point Bi. A global coordinate K_O is attached to the bottom platform, with its origin at point O, and a coordinate K_C is attached to the up platform, with its origin at point C. For the mechanism I, a spherical joint is located in point C. For the mechanism II, a revolute joint Rx is located in point O, a revolute joint Rz is located in point C, and a revolute joint Ry is located in point E. For the mechanism III, a revolute joint Rz is located in point O, a revolute joint Rx is located in point C, and a revolute joint Ry is located in point E.

The position vectors of $\boldsymbol{r}_{OAi}$, $\boldsymbol{r}_{EBi}$, $\boldsymbol{r}_{OC}$ and $\boldsymbol{r}_{OE}$ in their local coordinates are,

$$
\begin{aligned}
\mathbf{r}_{OAi} &= [R_a \cos \phi_i, R_a \sin \phi_i, 0]^T, \quad i = 1, 2, 3 \\
\mathbf{r}_{EBi} &= [R_b \cos \psi_i, R_b \sin \psi_i, 0]^T, \quad i = 1, 2, 3 \\
\mathbf{r}_{OC} &= [0, 0, H_1]^T, \mathbf{r}_{CE} = [0, 0, H_2]^T \\
R_a &= 0.3, R_b = 0.1, H_1 = 0.2, H_2 = 0.1 \\
\phi_1 &= -\phi_2 = v_1, \phi_3 = -\phi_4 = \pi - v_1 \\
\psi_1 &= -\psi_2 = v_2, \psi_3 = -\psi_4 = \pi - v_2
\end{aligned}
\tag{1}
$$

where v_1 and v_2 are unknown dimensions, which should be optimized later. In order to reduce the unknown dimensions, we assume the distribution of pneumatic muscles is symmetrical with x and y axis.

The position vector of $\boldsymbol{r}_{ABi}$ in the coordinate Ko of the mechanism I, II and III are,

$$
\begin{aligned}
\mathbf{r}_{ABi} &= -\mathbf{r}_{OAi} + \mathbf{r}_{OC} + \mathbf{R}_z \mathbf{R}_y \mathbf{R}_x (\mathbf{r}_{CE} + \mathbf{r}_{EBi}) \quad \text{I} \\
\mathbf{r}_{ABi} &= -\mathbf{r}_{OAi} + \mathbf{R}_x \mathbf{r}_{OC} + \mathbf{R}_x \mathbf{R}_z \mathbf{r}_{CE} + \mathbf{R}_x \mathbf{R}_z \mathbf{R}_y \mathbf{r}_{EBi} \quad \text{II} \\
\mathbf{r}_{ABi} &= -\mathbf{r}_{OAi} + \mathbf{R}_z \mathbf{r}_{OC} + \mathbf{R}_z \mathbf{R}_x \mathbf{r}_{CE} + \mathbf{R}_z \mathbf{R}_x \mathbf{R}_y \mathbf{r}_{EBi} \quad \text{III}
\end{aligned}
\tag{2}
$$

where $\boldsymbol{R}_x$, $\boldsymbol{R}_y$ and $\boldsymbol{R}_z$ are rotation matrix given as,

$$
\mathbf{R}_z = \begin{bmatrix} C\alpha & -S\alpha & 0 \\ S\alpha & C\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{R}_y = \begin{bmatrix} C\beta & 0 & S\beta \\ 0 & 1 & 0 \\ -S\beta & 0 & C\beta \end{bmatrix}, \mathbf{R}_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & C\gamma & -S\gamma \\ 0 & S\gamma & C\gamma \end{bmatrix}
\tag{3}
$$

here C denotes cosine and S denotes sine.

According to the screw theory, the actuation wrenches on the platform provided by pneumatic muscles are,

$$
\begin{aligned}
\hat{\$}_{wai} &= \begin{bmatrix} s_{ABi} \\ \mathbf{R}_z \mathbf{R}_y \mathbf{R}_x \mathbf{r}_{EBi} \times s_{ABi} \end{bmatrix} \quad \text{I} \\
\hat{\$}_{wai} &= \begin{bmatrix} s_{ABi} \\ \mathbf{R}_x \mathbf{R}_z \mathbf{R}_y \mathbf{r}_{EBi} \times s_{ABi} \end{bmatrix} \quad \text{II} \\
\hat{\$}_{wai} &= \begin{bmatrix} s_{ABi} \\ \mathbf{R}_z \mathbf{R}_x \mathbf{R}_y \mathbf{r}_{EBi} \times s_{ABi} \end{bmatrix} \quad \text{III}
\end{aligned}
\tag{4}
$$

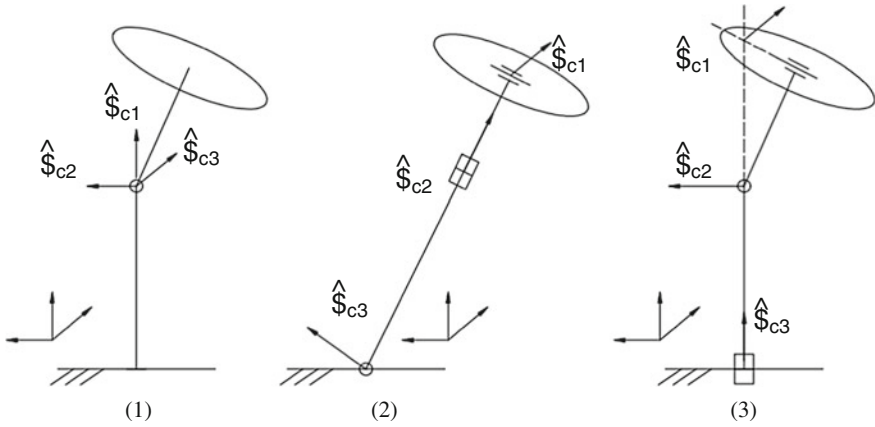where s_{ABi} is unit vector of $\boldsymbol{r}_{ABi}$.

Fig. 3 Constraint wrenches of the mechanism I, II and III

The constraint wrenches on the platform of the mechanism I, II and III are given as follows, as shown in Fig. 3 (Liu et al. 2011),

$$\hat{\$}_{c1} = \begin{bmatrix} e_z \\ -R_z R_y R_x r_{CE} \times e_z \end{bmatrix}, \quad \hat{\$}_{c2} = \begin{bmatrix} e_y \\ -R_z R_y R_x r_{CE} \times e_y \end{bmatrix}, \quad \hat{\$}_{c2} = \begin{bmatrix} e_x \\ -R_z R_y R_x r_{CE} \times e_x \end{bmatrix} \quad \text{I}$$

$$\hat{\$}_{c1} = \begin{bmatrix} R_x e_x \\ 0 \end{bmatrix}, \quad \hat{\$}_{c2} = \begin{bmatrix} R_x e_z \\ -R_x R_z r_{CE} \times (R_x e_z) \end{bmatrix}, \quad \hat{\$}_{c2} = \begin{bmatrix} R_x R_z e_y \\ -(R_x r_{OC} + R_x R_z r_{CE}) \times (R_x R_y e_y) \end{bmatrix} \quad \text{II}$$

$$\hat{\$}_{c1} = \begin{bmatrix} R_z e_x \\ H_2 \tan \gamma R_z R_x e_y \times (R_z e_x) \end{bmatrix}, \quad \hat{\$}_{c2} = \begin{bmatrix} R_z e_y \\ -R_z R_x r_{CE} \times (R_z e_y) \end{bmatrix},$$

$$\hat{\$}_{c3} = \begin{bmatrix} R_z e_z \\ -(R_z r_{OC} + R_z R_x r_{CE}) \times (R_z e_z) \end{bmatrix} \quad \text{III}$$

$$\tag{5}$$

Combining the actuation wrenches and the constraint wrenches, the wrench matrix can be formed which contained the kinematic information of the mechanisms,

$$W = \begin{bmatrix} \hat{\$}_{wa1}, \hat{\$}_{wa2}, \hat{\$}_{wa3}, \hat{\$}_{wa4}, \hat{\$}_{c1}, \hat{\$}_{c2}, \hat{\$}_{c3} \end{bmatrix} \tag{6}$$

In the wrench matrix, the first four wrenches represent actuation wrenches which are characterized by the unilateral property (can pull but cannot push). The last three wrenches represent constraint wrenches which are full constraints (can pull and push). This wrench matrix will be used for further analysis. A tension factor and a stiffness factor can be extracted from the wrench matrix.

4 Index of the PMDPPs

4.1 Tension Factor

For CDPMs, a wrench-closure workspace is the set of poses for which any platform wrench can be produced by tightening the cables. However for PMDPPs, their wrench-closure workspace is the space where any external platform wrenches can be counterbalanced by constraints wrenches working with sets of positive pneumatic muscles tension. This condition is mathematically described as,

$$\forall \mathbf{w}_p \in \mathbb{R}^6 : \exists t_1, t_2, t_3, t_4 \in [0, \infty) : \mathbf{W}\mathbf{t} = \mathbf{w}_p \tag{7}$$

where $\mathbf{t} = [t_1, t_2, t_3, t_4, t_{c1}, t_{c2}, t_{c3}]^{\mathrm{T}}$.

It should be noted that t_1–t_4 are coefficients of pneumatic muscle driven wrenches which need to be positive, and t_{c1}–t_{c3} are coefficients of constraint wrenches which can be any value.

According to the definition of the wrench-closure workspace, the necessary and sufficient condition for a pose to be inside the wrench-closure workspace is given as,

$$\exists \mathbf{t} \in \ker(\mathbf{W}) \text{ such that } \exists t_1, t_2, t_3, t_4 \in [0, \infty) \tag{8}$$

where $\ker(\mathbf{W})$ is the nullspace of $\mathbf{W}$. If the number of pneumatic muscles of the CPDPPs is equal to its DOFs plus 1, the CPDPPs falls into a fully constrained mechanism. In this case, the solution of Eq. (8) is $t = \mathrm{span}(t_0)$, where t_0 is given as,

$$t_0 = \begin{bmatrix} t_{01} \\ t_{02} \\ t_{03} \\ t_{04} \\ t_{0c1} \\ t_{0c2} \\ t_{0c3} \end{bmatrix} = \begin{bmatrix} \det\left(\left[\hat{\$}_{c3}, \hat{\$}_{wa2}, \hat{\$}_{wa3}, \hat{\$}_{wa4}, \hat{\$}_{c1}, \hat{\$}_{c2} \right] \right) \\ \det\left(\left[\hat{\$}_{wa1}, \hat{\$}_{c3}, \hat{\$}_{wa3}, \hat{\$}_{wa4}, \hat{\$}_{c1}, \hat{\$}_{c2} \right] \right) \\ \det\left(\left[\hat{\$}_{wa1}, \hat{\$}_{wa2}, \hat{\$}_{c3}, \hat{\$}_{wa4}, \hat{\$}_{c1}, \hat{\$}_{c2} \right] \right) \\ \det\left(\left[\hat{\$}_{wa1}, \hat{\$}_{wa2}, \hat{\$}_{wa3}, \hat{\$}_{c3}, \hat{\$}_{c1}, \hat{\$}_{c2} \right] \right) \\ \det\left(\left[\hat{\$}_{c3}, \hat{\$}_{wa2}, \hat{\$}_{wa3}, \hat{\$}_{wa4}, \hat{\$}_{c3}, \hat{\$}_{c2} \right] \right) \\ \det\left(\left[\hat{\$}_{c3}, \hat{\$}_{wa2}, \hat{\$}_{wa3}, \hat{\$}_{wa4}, \hat{\$}_{c1}, \hat{\$}_{c3} \right] \right) \\ -\det\left(\left[\hat{\$}_{c3}, \hat{\$}_{wa2}, \hat{\$}_{wa3}, \hat{\$}_{wa4}, \hat{\$}_{c1}, \hat{\$}_{c2} \right] \right) \end{bmatrix} \tag{9}$$

Therefore, a tension factor can be extracted from t_0 which can evaluate the quality of the wrench closure at a specific configuration,

$$\eta_t = \begin{cases} \frac{\min(t_{01})}{\max(t_{01})} & \forall t_{01} < 0, \text{ or } \forall t_{01} > 0 \quad i = 1, 2, 3, 4 \\ 0, & \text{otherwise} \end{cases} \tag{10}$$

where t_{0i} is ith element of vector $\boldsymbol{t}_0$. The tension factor is defined as the minimum tension over the maximum tension of the pneumatic muscles. It reflects the relative tension distribution among the pneumatic muscles for a specific configuration. The range of the tension factor is from zero to one. If t_{01}–t_{04} are not same sign, the tension factor is equal to zero, i.e. the platform is out of the wrench-closure workspace. If one of the pneumatic muscle tensions is close to zero, the tension factor approaches to zero, i.e. the platform is located near the workspace boundary. If the tension factor is close to 1, there is a better tension balance among the pneumatic muscles.

4.2 Stiffness Factor

The stiffness of pneumatic muscles is a function of its length and inner pressure, while the stiffness of its constraints is assumed to be a large value. Therefore a stiffness matrix for each component is given as,

$$\mathbf{k} = \mathrm{diag}(k_{a1}, k_{a2}, k_{a3}, k_{a4}, k_{c1}, k_{c2}, k_{c3}) \tag{11}$$

where k_{ai} is the stiffness of the ith pneumatic muscle, k_{ci} is the stiffness of ith constraints. Once the wrench matrix and stiffness matrix are known, the stiffness characteristics of the PMDPPs can be analyzed by the decomposition of eigen-wrenches (Ciblak and Lipkin 1999; Zhao et al. 2015),

$$\begin{aligned} \mathbf{W_E k_E W_E^T} &= \mathbf{WkW}, \\ \mathbf{W_E} &= [\hat{\$}_{wl1}, \hat{\$}_{wl2}, \hat{\$}_{wl3}, \hat{\$}_{wr1}, \hat{\$}_{wr2}, \hat{\$}_{wr3}] \\ \mathbf{k_E} &= \mathrm{diag}(k_{l1}, k_{l2}, k_{l3}, k_{r1}, k_{r2}, k_{r3}) \end{aligned} \tag{12}$$

where $\boldsymbol{W}_E$ is the matrix of eigen-wrenches; $\hat{\$}_{\mathrm{Wl}i}$ (i = 1, 2, 3) are three eigen-forces in three orthogonal directions and $\mathrm{k}_{\mathrm{l}i}$ are their linear stiffness; $\hat{\$}_{\mathrm{Wr}i}$ (i = 1, 2, 3) are three eigen-moments in three orthogonal directions and $\mathrm{k}_{\mathrm{r}i}$ are their torsional stiffness; $\boldsymbol{k}_E$ is the stiffness characteristic matrix of the PMDPP.

The linear stiffness factor and torsional stiffness factor are given as,

$$\eta_{kl} = \frac{\min(k_{li})}{\max(k_{li})}, \eta_{kr} = \frac{\min(k_{ri})}{\max(k_{ri})} \tag{13}$$

Since translation motions of the PMDPPs are restrained by joints, we only concentrated on the torsional stiffness factor η_{kr}. The torsional stiffness factor reflects the relative torsional stiffness distribution among three orthogonal directions. Its range is from zero to one. For the PMDPPs, the torsional stiffness consists of the stiffness of pneumatic muscles and the maximum torsional stiffness will not approach infinity. Therefore, the torsional stiffness factor approaches to 0 when the

minimum torsional stiffness approaches to zero. This means that the platform is uncontrollable, i.e. there is a singularity in this pose. The torsional stiffness factor approach to 1 results in a better torsional stiffness balance among three orthogonal directions.

4.3 Hybrid Tension-Stiffness Index

When we multiply the tension factor by the torsional stiffness factor, we can get a hybrid tension-stiffness index,

$$\eta_g = \eta_t \eta_{kr} \tag{14}$$

where the tension factor reflects the relative tension distribution among the four pneumatic muscles, while the torsional stiffness factor reflects the relative torsional stiffness distribution among three orthogonal directions. The tension factor close to 0 means the pose is near to the wrench-closure workspace boundary and the torsional stiffness factor close to 0 means the pose is near to singularity, and vice versa. Therefore, this index should be as large as possible.

For the optimization of the kinematic design, one may be interested in an index that represents a global property of the mechanism. A global hybrid tension-stiffness index is given as,

$$\bar{\eta}_g = \frac{\sum_{i=1}^{N} \eta_g(i)}{N} \tag{15}$$

in which $\eta_g(i)$ is the local hybrid tension-stiffness index at a particular sample point of the workspace, and N is the total number of the sample points.

5 Numerical Results and Discussion

According to Eq. (1), two dimensions v_1 and v_2 of the PMDPPs are unknown, which describe the position of Ai and Bi. The global hybrid tension-stiffness index is used to optimize those two values. Figure 4 shows the global index over v_1 and v_2 of the three mechanisms. Figure 4(1) shows the global index over v_1 and v_2, which is symmetrical with linear $v_1 = v_2$. This is due to the fact that the spherical joint of the mechanism I has an isotropy property. In contrast, the distribution of the global index of the mechanism II and III is asymmetrical, which is dependent on the arrangements of the revolute joints. The optimal dimensions of mechanism I are $v_1 = 1.3$ and $v_2 = 0.1$; the optimal dimensions of mechanism II are $v_1 = 0.2$ and $v_2 = 0.9$; the optimal dimensions of mechanism III are $v_1 = 1.5$ and $v_2 = 0.1$.
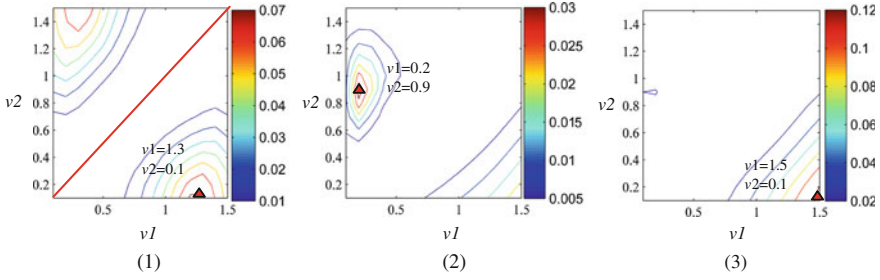
Fig. 4 Global index over v_1 and v_2 of the mechanism I, II and III
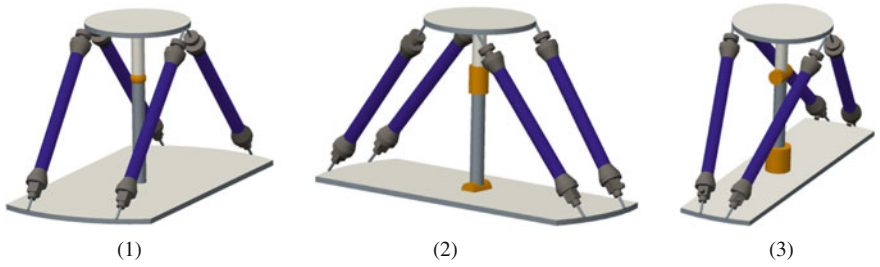

Fig. 5 Structure of the PMDPPs after optimization

Figure 5 shows the optimal structures of the mechanism I, II, and III. The distribution of pneumatic muscles over the three mechanisms is similar to the distribution of muscles in the human shoulder shown in Fig. 1. This similarity testifies to the effectiveness of the hybrid tension-stiffness index.

After optimization, the tension factor, the torsional stiffness factor, as well as the hybrid tension-stiffness index of the three mechanisms are expressed in Fig. 6. Row (a), (b) and (c) denote the tension factor, the torsional stiffness factor and the hybrid tension-stiffness index, respectively. Column (1), (2) and (3) denote the mechanism I, II and III. Compared with (a2), (a3) and (b2), (b3), one can see that the distribution of the tension factor and the torsional stiffness factor for one mechanism are different. Each of them cannot represent the whole kinematic property of the mechanism. With that said, a hybrid tension-stiffness index has an advantage which is the combination of both factors.

Figure 7 shows the wrench-closure workspace and the hybrid tension-stiffness index of the mechanism I, II and III. The boundary line of each layer is $\eta = 0.01$ and the gradient between two adjacent lines is $\eta = 0.1$. According to Fig. 7, the range of angle α of the mechanism II is from $-30°$ to $30°$, while the range of angle α of the mechanism I and III is from $-60°$ to $60°$. The mechanism II has the minimum wrench-closure workspace from all three mechanisms. The mechanism III has a larger wrench-closure workspace, as well as a better tension-stiffness factor among its wrench-closure workspace compared to the mechanism I and II. Therefore, in the
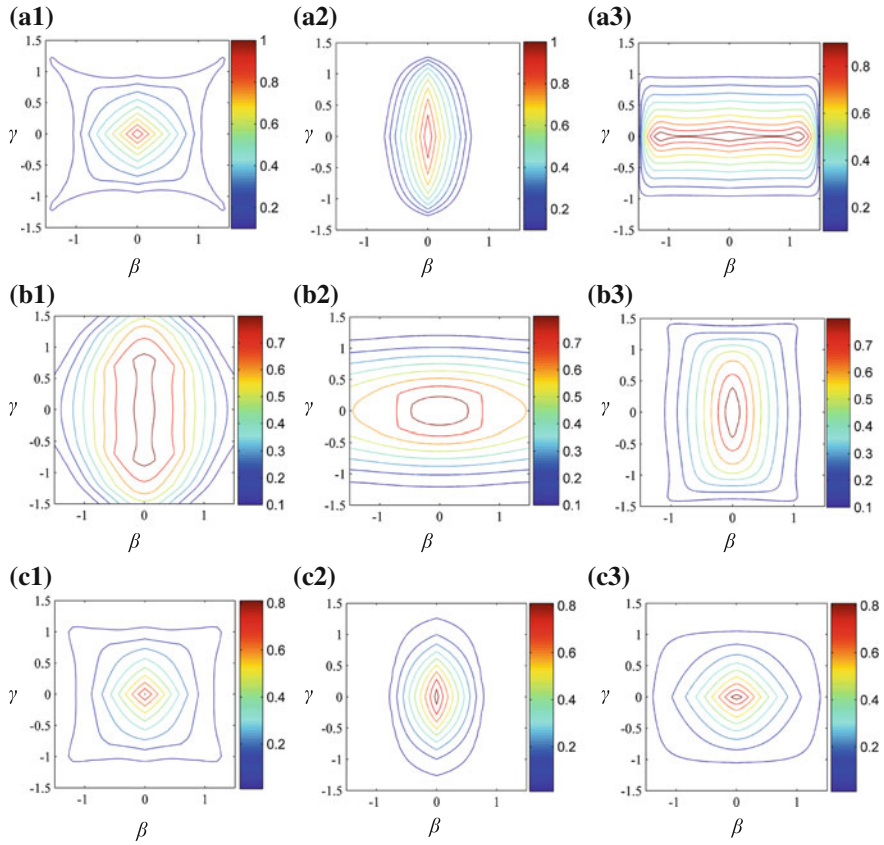
Fig. 6 Tension factor (**a1–a3**), torsional stiffness factor (**b1–b3**), and hybrid tension-stiffness index (**c1–c3**) of the mechanism I, II and III when $\alpha = 0$
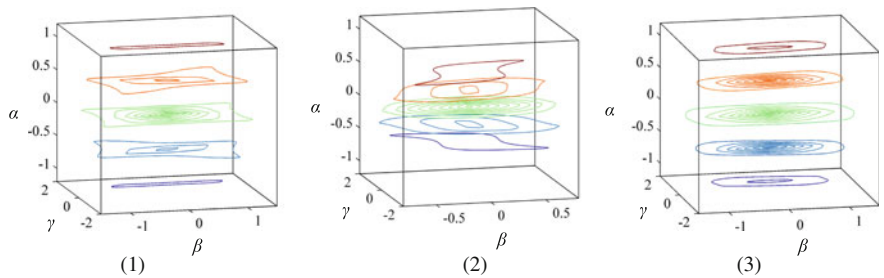

Fig. 7 Workspace and index of the mechanism I, II and III

view of the kinematic analysis, mechanism III is a better choice for imitating the human shoulder than the mechanism I and II, which has its advantages which are specifically the large wrench-closure workspace, the high tension-stiffness factor, as well as the low manufacturing cost (i.e. only revolute joints are required). Due to the space limitation the dynamic comparison of the three mechanisms will be discussed in another paper.

6   Conclusion

In this paper, pneumatic muscle driven parallel platforms imitating the human shoulder (PMDPP) have been presented as an approach to imitating the ways in which the human shoulder works. According to the structure of the human shoulder, three PMDPPs have been designed, where all of them have three rotational DOFs and are driven by four pneumatic muscles. A hybrid tension-stiffness index was proposed to analyze the property of the mechanism I, II and III. The tension factor reflects the relative tension distribution among four pneumatic muscles, while the torsional stiffness factor reflects the relative torsional stiffness distribution among three orthogonal directions. Each of them cannot represent the whole kinematic property of the mechanisms. A hybrid tension-stiffness index has advantages in combining both factors. After that, a global hybrid tension-stiffness index was used for the dimensions optimization of the three mechanisms. The optimal structures of the three mechanisms were obtained. The property of the mechanism I, II and III were compared to one another, and the mechanism III has the maximum wrench-closure workspace, as well as the best tension-stiffness factor among the wrench-closure workspace from all three of the mechanisms. In conclusion, in the view of the kinematic analysis, the mechanism III would be a better choice for imitating the motion of the human shoulder, which has a few advantages such as its large workspace, its high tension-stiffness factor, as well as its low manufacturing cost.

References

Ciblak, N., & Lipkin, H. (1999). Synthesis of Cartesian stiffness for robotic applications. *In Robotics and Automation, 1999 IEEE International Conference on* (Vol. 3, No. 1, pp. 2147–2152). USA: Michigan.

Ebert-Uphoff, I., & Voglewede, P. (2004). On the connections between cable-driven robots, parallel manipulators and grasping. In *Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE International Conference on* (Vol. 5, pp. 4521–4526). USA: New Orleans.

Gouttefarde, M. & Gosselin, M. (2004). On the properties and the determination of the wrench-closure workspace of planar parallel cable-driven mechanisms. In *ASME 2004 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (pp. 337–346). USA: Utah.

Gouttefarde, M., & Gosselin, M. (2006). Analysis of the wrench-closure workspace of planar parallel cable-driven mechanisms. *Robotics, IEEE Transactions on*, 22(3), 434–445.

Hiller, M., Fang, S., Mielczarek, S., Verhoeven, R., & Franitza, D. (2005). Design, analysis and realization of tendon-based parallel manipulators. *Mechanism and Machine Theory, 40*(4), 429–445.

Lenarčić, J., & Stanišić, M. (2003). A humanoid shoulder complex and the humeral pointing kinematics. *Robotics and Automation, IEEE Transactions on, 19*(3), 499–506.

Lenarčič, J., Stanišič, M. M., & Parenti-Castelli, V. (2000). Kinematic design of a humanoid robotic shoulder complex. *In Robotics and Automation, 2000. Proceedings. ICRA'00. IEEE International Conference on* (Vol. 1, No. 1, pp. 27–32). USA: San Francisco.

Liu, H., Huang, T., & Chetwynd, D. G. (2011). A method to formulate a dimensionally homogeneous jacobian of parallel manipulators. *Robotics, IEEE Transactions on, 27*(1), 150–156.

Pham, C. B., Yeo, S. H., Yang, G., & Chen, I. M. (2009). Workspace analysis of fully restrained cable-driven manipulators. *Robotics and Autonomous Systems, 57*(9), 901–912.

Pham, C. B., Yeo, S. H., Yang, G., Kurbanhusen, & Chen, I. M. (2006). Force-closure workspace analysis of cable-driven parallel mechanisms. *Mechanism and Machine Theory*, 41(1), 53–69.

Sodeyama, Y., Mizuuchi, I., Yoshikai, T., Nakanishi, Y., & Inaba, M. (2005). A shoulder structure of muscle-driven humanoid with shoulder blades. In *Intelligent Robots and Systems. Intelligent Robots and Systems, 2005. (IROS 2005). 2005 IEEE/RSJ International Conference on* (pp. 4028–4033). Canada: Alberta.

Tadokoro, S., Murao, Y., Hiller, M., Murata, R., Kohkawa, H., & Matsushima, T. (2002). A motion base with 6-DOF by parallel cable drive architecture. *Mechatronics, IEEE/ASME Transactions on, 7*(2), 115–123.

Yang, G., Lin, W., Pham, C., & Yeo, S. H. (2005). Kinematic design of a 7-DOF cable-driven humanoid arm: a solution-in-nature approach. *In Advanced Intelligent Mechatronics. Proceedings, 2005 IEEE/ASME International Conference on* (pp. 444–449). USA: California.

Zhao, X., & Zi, B. (2013). Design and analysis of a pneumatic muscle driven parallel mechanism for imitating human pelvis. In *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*. 0954406213489410.

Zhao X., Zi B., & Qian L. (2015). Design, analysis, and control of a cable-driven parallel platform with a pneumatic muscle active support. *Robotica*. 1–22.

Zi, B., Duan, B. Y., Du, J. L., & Bao, H. (2008). Dynamic modeling and active control of a cable-suspended parallel robot. *Mechatronics, 18*(1), 1–12.

Conceptual Design of Energy Efficient Lower Extremity Exoskeleton for Human Motion Enhancement and Medical Assistance

Nazim Mir-Nasiri

Abstract The paper describes conceptual design and control strategies for a new fully autonomous lower limb exoskeleton system. The main advantage of the system is its ability to decouple the weight/mass carrying function of the system from its forward motion function to reduce power consumption, weight and size of the propulsion motors. An efficient human machine interface has been achieved by means two sets of sensors: one (flexible sensors) to monitor subject leg's shank and ankle movements and the second to monitor subject's foot pressure. The weight is supported by a couple of passive pneumatic cylinders with electronically controlled ports. Joint motors of the exoskeleton then are only left to timely drive links of the exoskeleton when the legs take step. Therefore, motors consume less electrical energy and are small in size. In contrast to other existing exoskeleton designs, the motor batteries are able to sustain the energy supply for a longer travel distance before discharging.

Keywords Robotics · Exoskeleton · Motion enhancement · Energy efficient system

1 Introduction

Exoskeletons for human performance enhancement are wearable devices that can support and assist the user besides increasing their strength and endurance. During the last decade, researchers have focused on the development of lower limb exoskeletons that are now applied to several fields, including power augmentation for the military (Zoss et al. 2006) or medical assistance (Suzuki et al. 2007), and rehabilitation (Jezernik et al. 2003; Mao and Agrawal 2012; Jamwal et al. 2014; Iqbal and Baizid 2015). In such devices human provides control signals while the

N. Mir-Nasiri (✉)
School of Engineering, Nazarbayev University, 53, Kabanbay Batyr Ave.,
Astana 010000, Republic of Kazakhstan
e-mail: nazim.mir-nasiri@nu.edu.kz

exoskeleton actuators provide required power for performing the task. Thus human applies a scaled-down force compared to the load carried by the exoskeleton. A distinctive characteristic of exoskeletons compared to other robotic interfaces with haptic feedback is their close physical and cognitive coupling between the robot and the user (Pons 2008). The components—physical and control—that allow this physical and cognitive cooperation constitute the human-robot interface. In such design, the physical human-robot interfaces were developed, i.e. the mechanical and sensory components that mediate the transfer of physical interaction between the user and the exoskeleton (De Santis et al. 2008).

Early attempt on lower extremity exoskeleton development can be traced back to the 70's when Vukobratovic et al. (1974) studied anthropomorphic systems dynamics for development of robotic exoskeleton. The exoskeleton system developed at Mihailo Pupin Institute aims at helping paraplegic patients on rehabilitation. The device was constructed with pneumatic powered three degrees of freedom on each leg's sagittal (side) plane (hip, knee and ankle), making six pneumatic actuators in total. The hip joints are also powered for coronal (frontal) plane which allow for compensation and stability. Interface to the human was done with cuffs along the legs and a "corset" for the torso. The exoskeleton was driven by a preprogrammed walking gait sequence with compensation for external disturbances.

On lower extremity exoskeletons, most previous researchers paid their attention in developing walking aid systems for gait disorder persons or aged people (Veneman et al. 2007; Banala et al. 2009). One of those systems is HAL (Hybrid Assistive Leg) developed by Yoshiyuki Sankai of University of Tsukuba was aimed at assisting human leg muscles during walking (Kasaoka and Sankai 2001; Kawamoto and Sankai 2002, 2005). The system was based on electromyography (EMG) sensing of human muscles as the primary drive signals. The development resulted in several versions of HAL with the latest HAL-5 in 2009 (Kawabata et al. 2009). The exoskeleton was motor powered on the hip and knee joints, leaving other joints free. The system includes an embedded computer, sensor systems, motor drivers and battery pack, thus allows full mobility. The significance of their design is the implementation of EMG sensing which detects muscle activities before actual limb movement. Motor driven joints approach was taken by other researches as well (Kao and Ferris 2009). The Berkeley Lower Extremity Exoskeleton (BLEEX) developed by Kazerooni et al. (2006) and Ghan and Kazerooni (2006) was aimed at enhancing human strength and endurance for payload transport. The exoskeleton incorporates hydraulic actuation on all three sagittal joints and two coronal joints on the hip with all others joints free. The overall control design of BLEEX was to minimize interface between human and machine. Therefore, there was no sensor in direct measurement of human leg but includes all required sensors for determining the dynamics of the exoskeleton. The control system monitors the dynamics of the exoskeleton to determine operator's intention of motion. The exoskeleton has secure attachment to the operator only at the shoulder, waist and both feet. The significance of BLEEX is of the complex control network distributed throughout the exoskeleton and a custom designed onboard engine to power the hydraulic actuation system. Hydraulic actuation was

implemented by various researchers (Heng et al. 2009, 2010). The ECUST Leg Exoskeleton Robot (ELEBOT) designed by Heng et al. (2009) at East China University of Science and Technology (ECUST) shares the similar design goal as BLEEX but with a simplified system. ELEBOT has the same approach of using hydraulic system as joint actuation. However, has identified that only the knee joints would require substantial actuation support and therefore leaving all other joints free. The control of ELEBOT also came close to that of BLEEX by only monitoring stance phase and torque generated on the hydraulic actuators.

While the above exoskeleton designs require substantial power for operation on low efficiency, an exoskeleton design at Massachusetts Institute of Technology (MIT) (Walsh and Paluska 2006) attempted to lower the power requirement for load carrying. The exoskeleton has only series elastic actuation at hip sagittal joints, variable damper at the knee joints and spring at ankle sagittal joints. The control is based on a state machine and monitors forces and orientation of the exoskeleton to determine the states. The Walking Assist Device designed by Ikeuchi et al. (2009) at Honda Research and Development aims at increasing the lower extremity endurance of the elderly and those with weak legs. By partially supporting the upper body weight, the user bears less weight on the lower limbs and requires less energy for motion. The designed Walking Assist Device was a pair of non-anthropomorphic mechanical limbs attached to a seat. The whole device was fixed between the user's legs during operation. The walking assist device was only powered by electric motor at each of the knee joints and incorporates only pressure sensors beneath the shoes of the device. The control monitors user's weight applied on the pressure sensors and provides required force on both knees to achieve the predetermined weight reduction on both of the user's legs.

Various problems remain to be solved, the most daunting being the creation of a power and cost efficient system that will allow an exoskeleton to operate for extended periods without being frequently plugged into external power. The paper presents a conceptual design of a new exoskeleton that will enable to decouple the weight/mass carrying function of the system from the forward motion control which will reduce then power and size of propulsion motors and thus the overall weight, cost and required electrical power for the system. Such lighter and cheaper devices are currently important engineering research area in medicine and military (Leslie 2012).

2 Conceptual Design of Exoskeleton Mechanical System

The exoskeleton in this paper calls for an anthropomorphic design that supports human body weight and assists human lower limb muscles during motion. Figures 1, 2 and 3 shows the conceptual sketch of the proposed exoskeleton structure around a subject model. The proposed exoskeleton features a seat 1 for supporting subject body weight. Along the limbs of the exoskeleton, each exoskeleton leg has six degrees of freedom: four at the hip 2, one at the knee 3 and one at the ankle 4 to allow
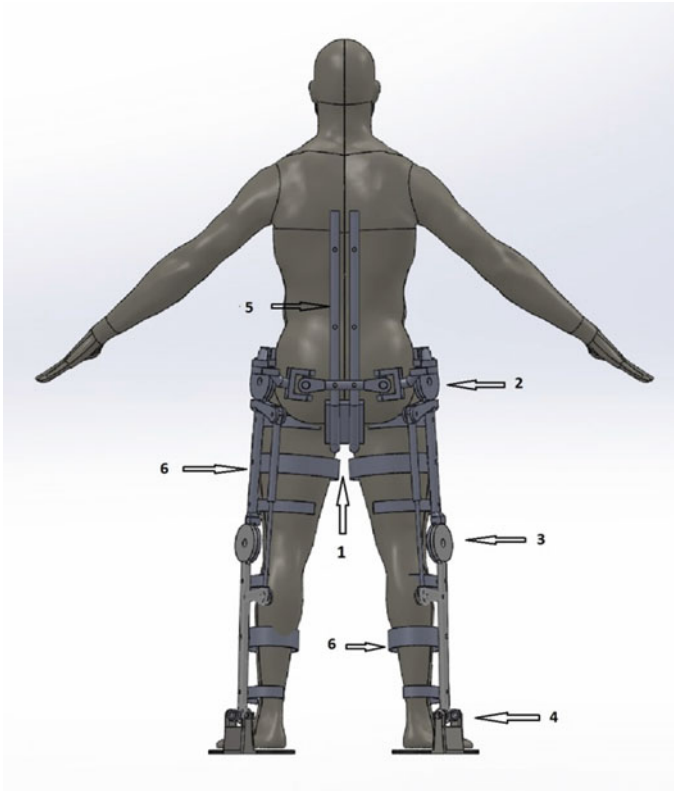
Fig. 1 Back view

as much motion freedom as possible. The design begins with a cushioned seat 1 in between subject crotch, supported by two parallel rigid pipes 5 to the back of the object. A back panel 6 mounted onto the rigid pipes serves as a platform for control and power supply mounting and also belt attachments for securing the subject. At hip level, the two parallel rigid pipes extend out to the two hip coronal joint. The link then continues to both sides of the hip 2 where sagittal and transverse joints are located, subsequently to the knee joints 3 and through the ankle joints 4 to the ground. Both exoskeleton legs are attached to subject legs using flexible cuffs 6. The three degrees of freedom linkage mechanism with rotary joints 7 at the hip level provides full freedom to maintain the lateral rotation of the exoskeleton leg around vertical axis located exactly at the center of the subject's hip. The remaining three single-degree parallel axes rotary joints at the hip 2, knee 3 and ankle 4 provide freedom of flexion at the joints. Pneumatic cylinders 8 are used to block the motion at the knee joints 3 when necessary.

Figures 4 and 5 are schematic diagrams of the exoskeleton. In these figures the telescopic members of the exoskeleton 1 that can be adjusted according to the leg
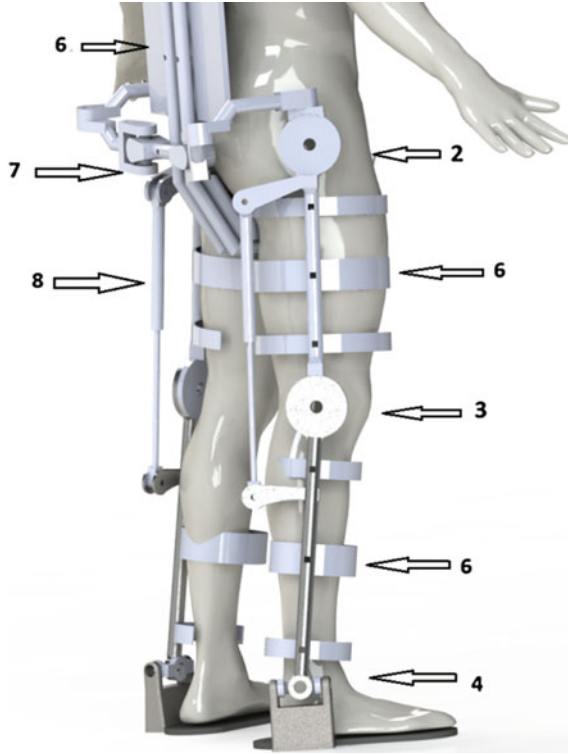
Fig. 2 Side view
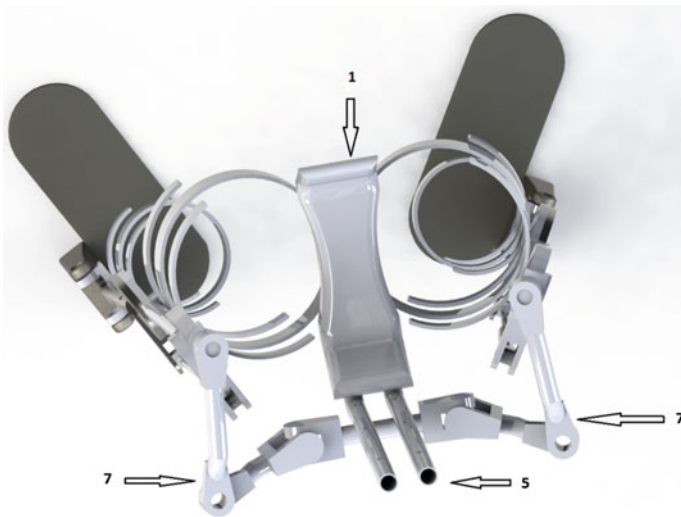
Fig. 3 Top view

Fig. 4 Schematic diagram of exoskeleton
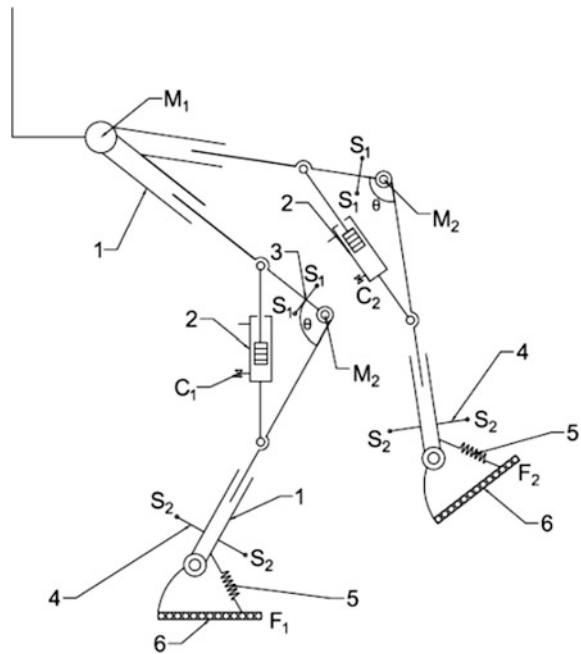
Fig. 5 Exoskeleton interface with object
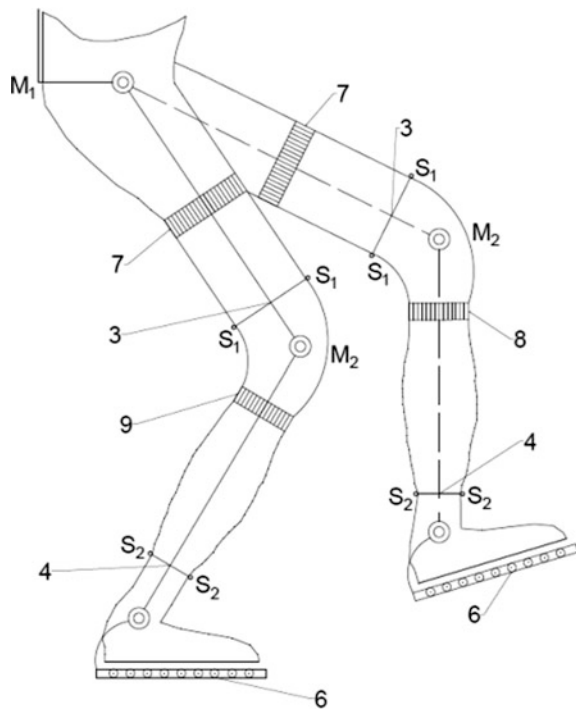
size of the subject; 2 are dummy pneumatic cylinders that are able to inhibit the motion at the knee joints if their respective valves are blocked; 3 are the sensors to detect motion of subject thigh; 4 are the sensors to detect motion of subject shank; 5 are springs to support feet of the exoskeleton; 6 are feet of the exoskeleton; 7 are flexible belts to fasten the exoskeleton to the subject thigh; 8 are flexible belts to fasten the exoskeleton to the subject shank.

In Figs. 4 and 5, M_1 and M_2 are motors driving the hip and knee joints of each leg; C_1 and C_2 are solenoid valves of the pneumatic cylinders 2 that are able to inhibit motions at the knee joints; S_1 and S_2 are flexible strips 3 and 4 with bonded strain gages that are able to sense the motions of subject limbs; F_1 and F_2 are foot pressure sensors that are able to sense the amount of pressure applied by the ground on the exoskeleton sole 6 during the walk. The pressure at the exoskeleton sole is generated due to the transmission of the weight forces via mechanical structure to the ground while the subject is resting on the seat.

3   Conceptual Design of Controller for Human-Machine Interfacing

The operation of all actuators (motors M_1 for the hip joint of each leg, motor M_2 for the knee joint of each leg and the knee motion inhibiting cylinders solenoid valves C_1 and C_2) depends on the amount of pressure felt by the feet pressure sensors F_1 and F_2 (Fig. 4). The pressure on the feet depends on actual posture of the subject and can be classified in three distinct cases. For example, if (case 1) the subject is standing on one leg, then the total subject body weight will be fully transmitted via that leg structure and the expected pressure reading from the corresponding foot sensors F will be at its maximum possible value P_b (Fig. 1) whereas the reading from other pressure sensor will be at zero value. If (case 2) the subject stands still on both legs then the total subject body weight is almost equally shared by both leg structures and the expected pressure reading from both feet sensors will be almost equal half of maximum possible value $P_b/2$. If (case 3) the subject is in the stage of transiting the weight from, e.g. leg 1 to leg 2 while standing on both legs, then the reading from sensor F_1 will gradually reduce from its half value $P_b/2$ to zero while the reading from sensor F_2 will gradually increase from its half value $P_b/2$ to its maximum possible value P_b. The pressure P_b is due to cumulative object's and exoskeleton's weights.

Since the cylinders across the knee join are able to inhibit the motion of the latter (if the inlet port of each cylinder is closed), then they can be used to assist the subject to support fully or partially the effective weight felt by muscles of the knee joint. The rules have to be established as when exactly motion of the knee joints should be inhibited and then released from inhabitation during the subject walk.

The first rule of inhibiting the motion of knee joint depends also on the minimum allowable bending angle $\Theta_{\min}$ of the knee joint. Bending angle is monitored by the

rotary sensors imbedded into the knee joint of the exoskeleton. The subject at the start of rehabilitation procedure won't be able to support its own weight while walking. Therefore, the system should allow the patient to rest on the rigid structure of the exoskeleton when this minimum allowable bending angle Θ_{min} of the knee joint has been reached. At this particular moment of time the respective pneumatic cylinder should fully inhibit the rotation at the knee joint and convert the exoskeleton into the rigid structure. The value of this bending angle Θ_{min} depends on usual walk style of the subject when he/she takes a step forward with one leg and stands on the other one which is bent at the knee joint. It can be selected based on actual numerous studies of human walk in order to not hinder the comfort of walker. This implies a stilt type of walk when the knee joint of one leg is inhibited at particular instant of motion while the user takes a step forward with another leg. While the rehabilitation is in progress the object muscles will be able with time to fully support the weight and then the bending angle Θ_{min} will be never reached. Therefore the cylinders will not then block the motion at the knee joint and thus the exoskeleton will not hinder motion of the object's leg.

The second rule obviously depends on the weight of a particular object and the minimum pressure P_{min} that would be allowed at the sole of exoskeleton foot before inhabitation can be cancelled at a particular knee joint and the exoskeleton will be allowed to activate its motors to take a step. It can be then expressed as a fraction of the total weight value, i.e. $P_{min} = P_b \cdot k$. The k can be selected in the range from 0 to 1. Besides, k should be a varying parameter which depends on rate of the patient recovery. In fact, ability of the patient to support own weight should increase gradually during the rehabilitation period and intensive exercises. Parameter k should be selected manually by the patient or the monitoring doctor manually depending on the progress of the treatment.

Based on the rules discussed above, the following control strategy for the walk is proposed:

- Motors M_1 and M_2 of each leg are actuated only and only if $\Theta \geq \Theta_{min}$ (first rule) **[AND]** pressure reading from the corresponding pressure sensors at the foot $P \leq P_{min}$ (second rule) in order to pick up the leg from the ground and take a step
- Cylinders solenoid valves C_1 and C_2 are actuated only and only if $\Theta < \Theta_{min}$ (first rule). The algorithm can be adopted for the actuators gradually to applying closing forces on the valves. It means the cylinders will be actuated when the knee has reached the minimum allowed value (first rule) which means further bending at the knee joint is not allowed

The control strategy for the motors M_1 and M_2 is aimed to make sure that the exoskeleton structure will follow the subject's leg physical motion without hindering it. The set of sensors S_1 and S_2 (flexible strips with bonded strain gages) are attached to the links of the exoskeleton (Figs. 4 and 5). When the subject limbs commence the motion the limbs will touch and bend the strips. The sensors will detect in real time the intended tiny motions of the subject's limbs and send the
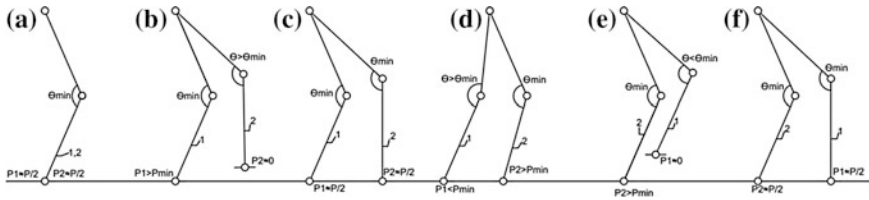
Fig. 6 Gaits of the exoskeleton

signals to the PID controller. The controller will react immediately by activating hip and knee motors M_1 and M_2 in order to move the links of the exoskeleton away from the object limbs and thus to restore the original shape of the strips. The set point of the PID controller is zero signals from the sensors. The PID controller can provide fast system response and accurate positioning of the exoskeleton links with respect to subject's limps. As a result object limp can move free with no obstruction from the exoskeleton.

The exoskeleton motion phases (strides) are demonstrated in Fig. 6. Figure 6a–d show how the exoskeleton takes a step with leg 2 from the stationary posture while leg 1 motion is inhibited by the respective cylinder. These figures show the expected instantaneous values of sensors Θ and P during the motion. Figure 6d–f show how the exoskeleton takes a step with leg 1, instead, while leg 2 motion is inhibited by the respective cylinder. Figure 6c, f are identical except the legs have interchanged their duties.

The elements of the control system are shown in Fig. 7. The microcontroller received the signals from the knee angle sensor Θ, foot pressure sensor P, strain gages S_1, S_1', S_2, S_2' and actuate either cylinder valve C or motors M_1 and M_2.

The system operational or logic flow chart is shown in Fig. 8. If the common switch is on then the system start receiving data from angle Θ and pressure P sensors. If $\Theta \leq \Theta_{min}$ and $P \neq 0$ that means exoskeleton shoe is in touch with the ground and knee angle limit has been reached. In this case cylinder valve is actuated and joint motors are deactivated. If $\Theta \leq \Theta_{min}$ and P–0 that means exoskeleton shoe is in the air (not in touch with ground). In this case cylinder valve is deactivated and the operation of motors is allowed. If $\Theta \geq \Theta_{min}$ and $P \geq P_{min}$ that means knee angle limit has been reached but the shoe pressure above the pressure limit. In this case cylinder valve is actuated and joint motors are deactivated. Finally, if $\Theta \geq \Theta_{min}$ and $P \leq P_{min}$ that means the exoskeleton is allowed to deactivate cylinder valve and actuate the reading of strain gages and operation of both motors. By comparing the reading from the pair of sensors S_1, S_1', and S_2, S_2' the sense of motors rotation can be established. For example, if the reading $S_1 > S_1'$ the rotation of M_1 can be set in clockwise direction. Conversely, if $S_1 < S_1'$ then the rotation of
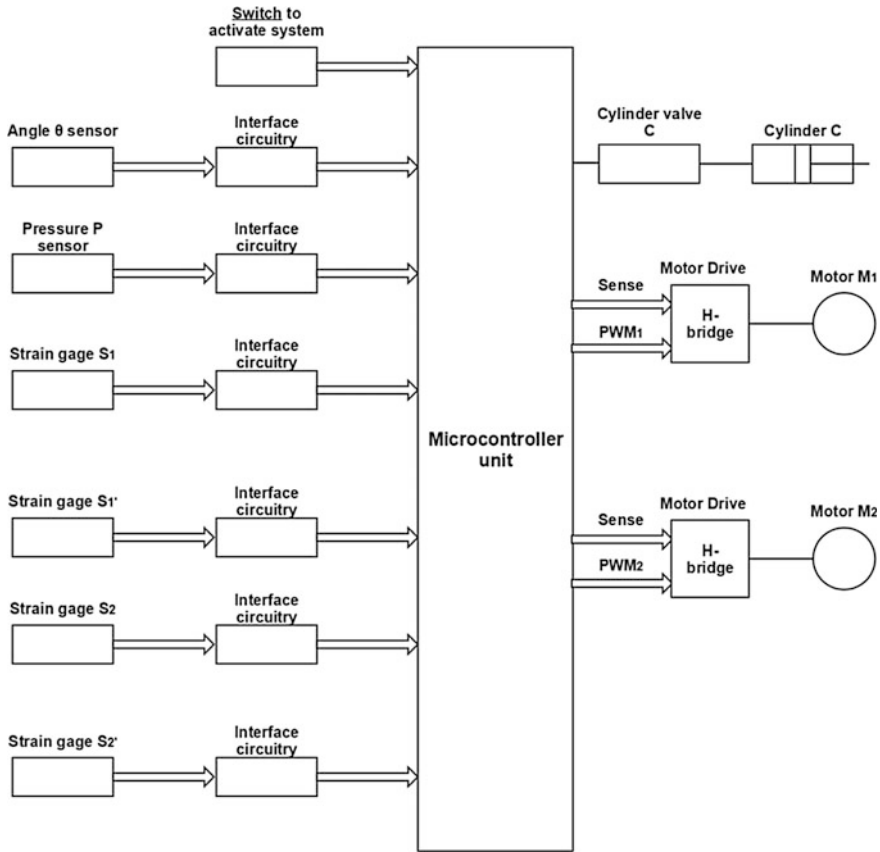
Fig. 7 Elements of the system

M_1 can be set in counterclockwise direction. Same is true for the data received from sensors S_2, S_2' that control the sense of rotation of motor M_2. The speed of motor rotation is controlled by the motor driver and PWM signal received from the microcontroller. PWM is selected to be proportional to the absolute difference between the readings of the pair of sensors, i.e. $|S_1 - S_1'| |S_2 - S_2'|$. It is very effective way of monitoring the speed of the motor response to the object intention to move a limb. The higher is the pressure applied by the user to the strip the higher is the acceleration of the motor to restore the shape of the strip with attached strain gages. It is in a way implementation of proportional control strategy for the motor speed control. The microcontroller operates in loop continuously checking status of all sensors, making decision and actuating either cylinder valves or the motors as long as the common switch is on (Fig. 8).

Fig. 8 Flow chart of
controller operation
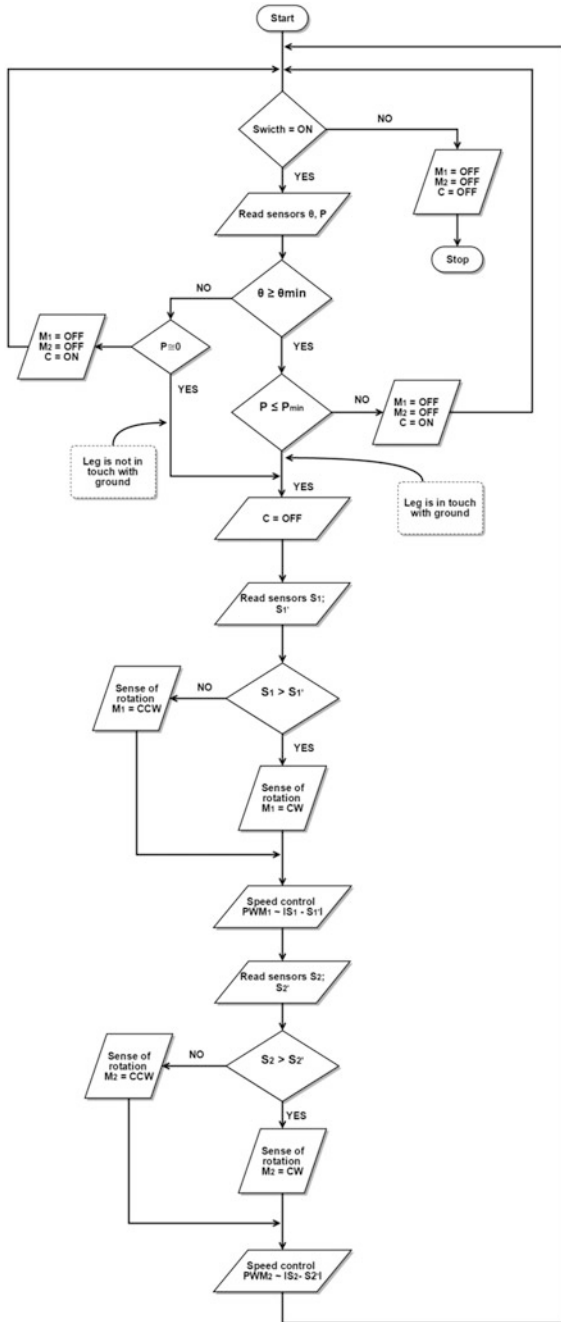
4 Conclusion

The paper describes the methodology of mechanical design and effective control of a new exoskeleton system to enhance walk capabilities of people. It also can be used for rehabilitation of people with leg injuries. The core idea is to use exoskeleton to decouple weight carrying capabilities of the legs from its body advancing capabilities. This has been done by special logic and intelligent management of electrical motors and motion inhibiting passive pneumatic cylinders operation. The operation is managed and controlled by the microcontroller which receives the necessary data from the strain gauge sensors located at the subject's thighs and shanks and the pressure sensors located at the feet. The work is being done currently is to simulate the system performance. This methodology will help people to leave the weight carrying capability to the exoskeleton and just focus on forward motion of own legs. This exoskeleton is more power efficient because electrical motors are smaller in size and did not participate in supporting the weight like in all other existing exoskeleton designs. The motors just provide a synchronous motion of the exoskeleton and human legs.

References

Banala, S., Agrawal, S., & Scholz, J. (2009). Robot assisted gait training with active leg exoskeleton (ALEX). *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 17*, 2–8.

De Santis, A., Siciliano, B., De Luca, A., & Bicchi, A. (2008). An atlas of physical human-robot interaction. *Mechanisms and Machines Theory, 43*, 253–270.

Ghan, J., & Kazerooni, H. (2006). System identification for the Berkeley lower extremity exoskeleton (BLEEX). In *Proceedings 2006 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3477–3484.

Heng, C., Jun, Z., Chunming, X., Hong, Z., Xiao, C., & Yu, W. (2010). Design and control of a hydraulic-actuated leg exoskeleton for load-carrying augmentation. In *Proceedings 2006 IEEE International Conference on Robotics and Automation (ICRA)*, Part I, LNAI 6424, pp. 590–599.

Heng, C., Zhengyang, L., Jun, Z., Yu, W., Wei, W. (2009). Design frame of a leg exoskeleton for load-carrying augmentation. In *Proceedings of IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pp. 426–431.

Ikeuchi, Y., Ashihara J., Hiki Y., Kudoh H., & Noda T. (2009). Walking assist device with bodyweight support system. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4073–4079.

Iqbal, J., & Baizid, K. (2015). Stroke rehabilitation using exoskeleton-based robotic exercisers mini review. *Biomedical Research, 26*(1), 197–201.

Jamwal, P. K., Sheng, Q. X., Shahid, H., & John, G. P. (2014). An adaptive wearable parallel robot for the treatment of ankle injuries. *IEEE/ASME Transactions on Mechatronics, 19*(1), 64–75.

Jezernik, S., Colombo, G., Keller, T., Frueh, H., & Morari, M. (2003). Robotic orthosis lokomat: A rehabilitation and research tool. *Neuromodulation, 6*, 108–115.

Kao, P., & Ferris, D. P. (2009). Motor adaptation during dorsiflexion-assisted walking with a powered orthosis. *Gait Posture., 29*, 230–236.

Kasaoka, K., & Sankai, Y. (2001). Predictive control estimating operator's intention for stepping-up motion by exo-skeleton type power assist system HAL. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1578–1583.

Kawabata, T., Satoh, H., & Sankai, Y. (2009). Working posture control of robot suit HAL for reducing structural stress. In *Proceedings of IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pp. 2013–2018.

Kawamoto, H., & Sankai, Y. (2002). Comfortable power assist control method for walking aid by HAL-3. In *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, pp. 6–12.

Kawamoto, H., & Sankai, Y. (2005). Power assists method based on phase sequence and muscle force condition for HAL. *Journal of Advanced Robotics, 19*(7), 717–734.

Kazerooni, H., Steger, R., & Huang, L. (2006). Hybrid control of the Berkeley lower extremity exoskeleton (bleex). *The International Journal of Robotics Research, 25*(5–6), 561–573.

Leslie, M. (2012). The next generation of exoskeletons. *A Magazine of the IEEE Engineering in Medicine and Biology Society, 3*(4), 56–61.

Mao, Y., & Agrawal, S. K. (2012). Design of a cable driven arm exoskeleton (CAREX) for neural rehabilitation. *IEEE Transactions on Robotics, 28*(4), 922–931.

Pons, J. L. (2008). *Wearable robots: Bio-mechatronic exoskeletons*. Hoboken, NJ, USA: Wiley.

Suzuki, K., Mito, G., Kawamoto, H., Hasegawa, Y., & Sankai, Y. (2007). Intention-based walking support for paraplegia patients with robot suit HAL. *Advanced Robotics, 21*, 1441–1469.

Veneman, J. F., Kruidhof, R., Hekman, E. E., Ekkelenkamp, R., van der Van Asseldonk, E. H., & Kooij, H. (2007). Design and evaluation of the LOPES exoskeleton robot for interactive gait rehabilitation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 15*, 379–386.

Vukobratovic, M., Hristic, D., & Stojiljkovic, Z. (1974). Development of active anthropomorphic exoskeletons. *Journal of Medical and Biological Engineering and Computing, 12*(1), 66–80.

Walsh, C.J., & Paluska, D. (2006). Development of a lightweight, underactuated exoskeleton for load-carrying augmentation. In *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3485–3491.

Zoss, A. B., Kazerooni, H., & Chu, A. (2006). Biomechanical design of the Berkeley lower extremity exoskeleton (BLEEX). *IEEE/ASME Transactions on Mechatronics, 11*(2), 128–138.

A New Algorithm for Analyzing Method of Electrical Faults of Three-Phase Induction Motors Using Duty Ratios of Half-Period Frequencies According to Phase Angle Changes

YoungJin Go, Myoung-Hyun Song, Jun-Young Kim, Wangrim Choi, Buhm Lee and Kyoung-Min Kim

Abstract A stator fault of induction motor occurs due to the breakdown of insulation, meaning the stator is directly connected with the power supply, and the direct connection is a direct cause of a major accident. For this reason, many studies are being performed to detect the faults. As for the existing studies on stator fault detection, they are being performed considering the possibility of stator fault only, excluding the possibility of rotor fault. It is necessary to identify and detect whether a stator (or rotor) fault is the cause of the electrical fault. This paper suggested a new algorithm that identifies the causes of stator faults with the use of the change in

Y. Go
Department of Electrical Automation, Suncheon JEIL College,
Deogwol-dong, Suncheon-city, Jeollanam-do, Republic of Korea
e-mail: zerojjin@suncheon.ac.kr

M.-H. Song · J.-Y. Kim
Department of Electric Control Engineering, Sunchon National University,
Maegok-dong, Suncheon-city, Jeollanam-do, Republic of Korea
e-mail: mhsong@sunchon.ac.kr

J.-Y. Kim
e-mail: rokemt21@naver.com

W. Choi
Department of Biomedical and Electronic Engineering, Chonnam National University,
Dundeok-dong, Yeosu-city, Jeollanam-do, Republic of Korea
e-mail: o2mewl@naver.com

B. Lee · K.-M. Kim (✉)
Department of Electrical and Semiconductor Engineering, Chonnam National University,
Dundeok-dong, Yeosu-city, Jeollanam-do, Republic of Korea
e-mail: kkm@jnu.ac.kr

B. Lee
e-mail: buhmlee@jnu.ac.kr

303

the duty ratio of the half-period frequency of the frequency when a phase angle change occurs at that moment. Also, by applying the algorithm to the fault of the rotor, it was also possible to grasp their fault state and to identify accurately whether a stator (or rotor) fault is the cause of an electrical fault.

1   Introduction

Most of the faults of induction motors are caused by bearing fault (44 %); wrong winding or stator fault due to a stator winding short circuit fault or a stator winding open circuit fault (26 %); and fault of the rotor (8 %). The causes of induction motors are largely categorized into mechanical faults and electrical faults—bearing faults causing mechanical faults; and stator faults and rotor faults causing electrical ones. Among the causes, a stator fault occurs due to the breakdown of insulation of the stator, a deterioration in insulation of the stator winding is caused by various reasons, such as coil movement, thermal stress, overload, and mechanical vibration, and the deterioration in insulation of the winding results in a turn to turn short.

The occurrence of a turn to turn short causes a large circulating current to occur in the faulty loop, and the heat in the winding, which is proportional to the square of the stator current, plays a role in accelerating the occurrence of a more serious fault, such as phase-to-phase short fault and a phase-earth fault, which cause a serious damage to the winding and core of the stator.

Hence this is the time when early diagnosis is desperately needed, as detecting a turn to turn short in the early stage can prevent a sequential damage to the motor and considerably reduce economic loss by minimizing the outage time of the motor, the number of additional workers, and the cost of repair.

1.1   Background

Among methods of diagnosis of ITSC(Inter-Turn Short Circuit) fault of induction motors, studies on MESA (Motor Electrical Signature Analysis) using only electrical signatures are being performed actively. Arkan et al. (2001) was researched through online stator fault diagnosis using Park's vector (PV) approach, which is a representative one; as for Song (2008), faults were identified by setting the circular distortion ratios at the threshold values with the use of DR; Filippetti (2006) and Tallam (2000) were researched through AI technique; Awadallah (2006) was researched through wevelet analysis. As for Kilman (2000) and Tallam et al. (2002), negative sequence currents were utilized; as for Kohler et al. (2002), negative sequence impedance was utilized. Lee et al. (2003) was researched through a matrix

of impedances modeling cross-coupling between positive and negative sequence components; and recently as for Bakhri (2012), studies were performed considering asymmetric elements in the machines in which a negative sequence current component is bound to occur. Go et al. (2014) was researched a stator fault diagnosis of an induction motor based on the phase angle of park's vector approach.

1.2 Problem and a Suggestion for a New Algorithm of Analysis Method

But as for the previous studies, experiments were performed considering the possibility of stator fault, an ITSC fault. Actually stator faults and rotor faults cause electrical ones, but it is necessary to identify which part has been faulty, following the occurrence of the electrical one.

As for methods using PV's DR, the occurrence of the change below the threshold value even in case of rotor fault causes a difficulty in identifying the accurate cause of the fault.

This paper suggested a fault diagnosis method using the duty ratio of half-period frequency according to phase angle change, with the consideration of the effects of negative sequence components on PV's d-q transform. By applying the method even to faults of rotors, it is possible to identify which part of the induction motor is faulty.

2 Background Theory

A three-phase induction motor can simplify MCSA by using Park's Vector Approach. The marks for a phasor space are located on the cross-section of an induction motor and are converted into arbitrary instantaneous values on the three-phase on the complex plane. The phasor space on this complex plane is rotated at the same angular rate as the angular frequency of the three-phase supply system.

Hence if the three-phase input current of an induction motor is applied to Park's Vector Approach, it is converted as follows (see the Eqs. 1 and 2 below):

$$I_d = \left(\frac{\sqrt{6}}{2}\right) i_M \sin(\omega t) \tag{1}$$

$$I_q = \left(\frac{\sqrt{6}}{2}\right) i_M \sin\left(\omega t - \frac{\pi}{2}\right) \tag{2}$$

Here i_M represents supply current's maximum value [A], ω represents supply angular frequency (rad/s), and t represents time parameter [s].
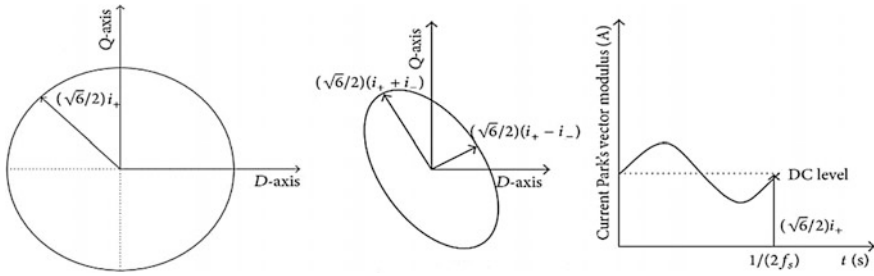
Fig. 1 Pattern of Park's vector

Figure 1 shows the trace of a circle with its center at the origin on a coordinate plane. Under normal start conditions, the left pattern in Fig. 1 is shown. If a stator winding turn fault occurs, the right pattern (oval shape) in Fig. 1 is shown. More turn faults mean more asymmetric components of supply current, and more asymmetric components of supply current mean a higher degree of distortion of the trace of the pattern of Park's Vector.

The degree of distortion of the pattern has recently been defined as *DR* (Distortion Ratio), which has been numerically calculated using the Eqs. 3 and 4.

$$r = \sqrt{i_d^2 + i_q^2} \tag{3}$$

$$DR = \frac{r_{max}}{r_{min}} \tag{4}$$

Here r represents the scalar of Park's Vector, r_{max} represents the broadest amplitude of it, and r_{min} represents the narrowest amplitude of it.

If the maximum value r_{max} is the same as the minimum value r_{min}, the induction motor is in the most ideal state (*DR*: 1). As a larger difference between them means a higher *DR*, the phase current when the difference is large is under highly unbalanced conditions due to winding turn fault.

2.1 Phase Changes

$r = \sqrt{I_d^2 + I_q^2}$ represents the maximum value of current, actually determines the shape of a circle in Park's Vector Approach, and is an important factor representing the amplitude of current. Here the components $\cos\theta$ and $\sin\theta$ of R_e and I_m determine the position of I_s.

When a fault occurred, a particular pattern was created by the actual phase angle θ although an analysis was performed only using the maximum value of r. This can be seen in Fig. 2.

Fig. 2 Different phase angle
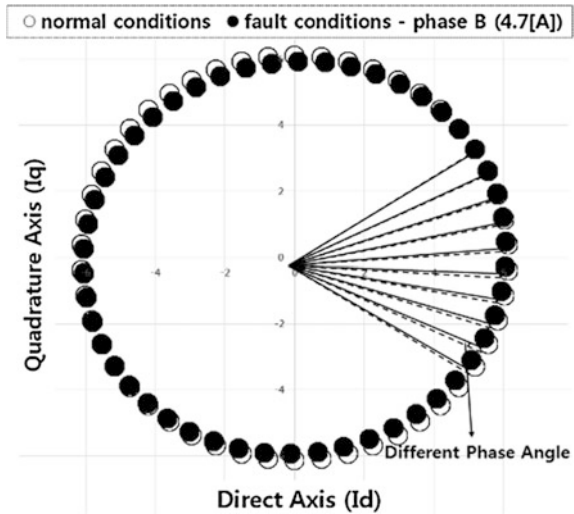
As for Fig. 2 a simulation was performed based on a normal condition current of 5[A]. A current fault signal was generated in phase B among three phases (the phase current: 4.7[A]). If the difference in fault current is large, the change of the phase angle θ is also large.

To use the characteristics of the pattern according to the change of the phase angle θ, the following Eq. 5 can be found using Euler's theorem.

$$e^{j\theta} = \frac{I_d + jI_q}{r} = \frac{I_d + jI_q}{\sqrt{I_d^2 + I_q^2}} \tag{5}$$

2.2 Negative Sequence

In steady state, the relationship between the vector operator 'a' and the phase angle ⊖ can be interpreted through a vector operator: This can be expressed as a theorem for the operator 'a' as follows:

$$a = 1\angle 120° = \cos 120° + j\sin 120° = -\frac{1}{2} + j\frac{\sqrt{3}}{2} \tag{6}$$

$$a^2 = 1\angle 240° = \cos 240° + j\sin 240° = -\frac{1}{2} - j\frac{\sqrt{3}}{2} \tag{7}$$

$$a^3 = 1\angle 360° = 1 \tag{8}$$

$$a + a^2 = -\frac{1}{2} + j\frac{\sqrt{3}}{2} - \frac{1}{2} - j\frac{\sqrt{3}}{2} = -1 \tag{9}$$

$$1 + a + a^2 = 0 \tag{10}$$

That is, in steady state, the total power of 3-phases must be zero.

Here a zero sequence component is not shown on a revolution indicator, and the component can be ignored. Therefore, when the total voltage of the positive sequence component of the stator is represented as $\tilde{V}_{1s}$ and that of the negative sequence component is represented as $\tilde{V}_{2s}$, it is possible to write the stator voltage of $\tilde{V}_s$ using the relationship $V_d + jV_q = \tilde{V}$ as follows:

$$\tilde{V}_s = \tilde{V}_{1s} + \tilde{V}_{2s} \tag{11}$$

When seeing Eq. 11 based on d-q transform, it can be expressed as follows:

$$v_{dqs} = \bar{V}^e_{1dqs}e^{-j(\theta - w_s t)} + \bar{V}^{-e}_{2dqs}e^{-j(\theta + w_s t)} \tag{12}$$

Here the formula $v_{dqs} = v_{ds} + jv_{qs}$ can be devised; it represents the positive sequence component of $\bar{V}^e_{1dqs} = \tilde{V}_{1s}$ and the negative sequence component of $\bar{V}^{-e}_{2dqs} = \tilde{V}^*_{2s}$; and t represents time and θ represents an angle of rotation.

In the Eq. 12, the domain of e is created by the positive sequence component that rotates together with the magnetic field and is expressed as such when only the positive sequence component in the 3-phase system exists; and v_{dqs} equals V_{1s} and is expressed as V^e_{1dqs}.

The domain of as $-e$ is created by the negative sequence component that rotates together with the magnetic field and is expressed as such when only the negative sequence component in the 3-phase system exists; and v_{dqs} equals V^*_{2s} and is expressed as V^{-e}_{2dqs}.

These two domains rotate in the opposite direction to which the two pairs of frequencies of d-q transform rotate at the same speed.

This can be expressed as follows:

$$V^e_{1dqs} = v^e_{1ds} + jv^e_{1qs} = \tilde{V}_{1s} \tag{13}$$

$$V^{-e}_{2dqs} = v^{-e}_{2ds} + jv^{-e}_{2qs} = \tilde{V}_{2s} \tag{14}$$

$$V^e_{dqs} = v^e_{ds} + jv^e_{qs} = V^e_{1dqs} + V^{-e}_{2dqs}e^{-j2w_s t} \tag{15}$$

In this way, it is also possible to infer currents and magnetic fluxes.

That is, it is possible to approach the interpretation of imbalance through the negative sequence component.

2.3 Imbalance Between Frequencies by Interpretation of Imbalance

When supposing that the positive sequence component, the negative sequence component, and the zero sequence component of 3-phases exist as in Fig. 3.

As you can see in Fig. 4, an imbalance between frequencies occurs due to the negative sequence component, which causes e^{-j2w_st} to affect the total phase angle θ in the element $-e$, the domain of the negative sequence component in the Eq. 15.

2.4 An Analysis of Duty Ratio of Frequency

This paper suggests a simplified algorithm of PWM (pulse width modulation) by detecting the rising times and falling times of half-period frequencies and analyzing distortions of the frequencies—the use of duty ratios of the frequencies for the analysis.

The amount of time for the half-period frequency at 60 Hz is approximately 0.0083 s, and in steady state, the falling time must be the same as the rising time: approximately 0.04167 s. But as you can see in Fig. 5, 0.0167 s (the total frequency: 60 Hz) remains the same because of the negative sequence component; but there is a time difference between the rising time and falling time in the duty ratio of the half-period frequency. In addition, the value of the duty ratio represented a change in the falling time compared to the amount of time for the half-period frequency.
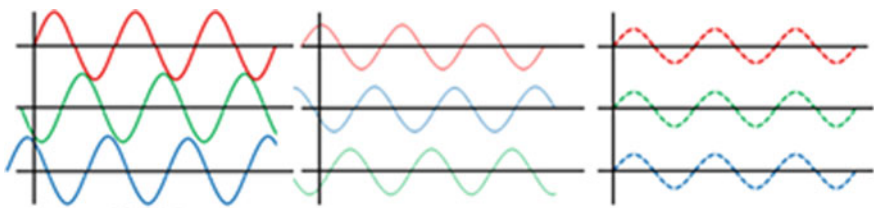
Fig. 3 Example of 3-phase positive (*left*), negative (*Center*), zero sequence component (*right*)
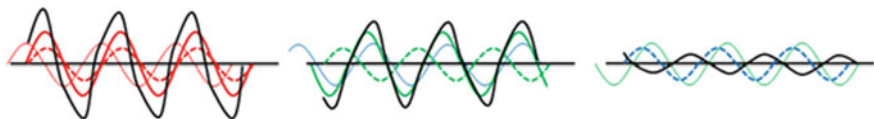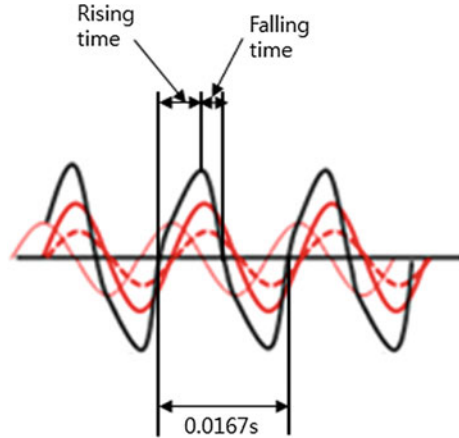

Fig. 4 Sum of positive, negative and zero sequence component

Fig. 5 Half-period frequency duty ratio analysis algorithm

$$\text{Duty Ratio}[\%] = \frac{Falling\ time}{Rising\ time + Falling\ time} * 100\%\qquad(16)$$

The measured frequency was expressed using Euler's theorem Eq. 5 and the duty ratio of the frequency in the formula above Eq. 16 was analyzed.

3   Experiment

3.1   *Experiment Condition*

In this experiment, 1[HP], 220[V], 4 poles, 60 Hz, a three-phase induction was used.

The numbers of rotor slot and stator are 44, 36.

In addition, the experiment was performed under the following conditions to consider changes in the load conditions.

– Operation of an induction motor by using inverter
– Adjustment of the motor operation speed by using dynamometer
– The initial speed—the rated speed 1690[rpm]
– The frequency of samplings, sampling rate and the No. of samplings: 1[s], 10,000[S/s], 10,000[s]
– Full load current and torque: 3.8[A]/2.2[A] and 0.43[kg m]

As an induction motor is operated by an inverter unlike a simulation, noise components exist as in the actual data in Fig. 6. For removal of noises, a filtering (IIR butterworth lowpass filter) was performed.

After a filtering was performed, current signals were obtained as in Fig. 7.
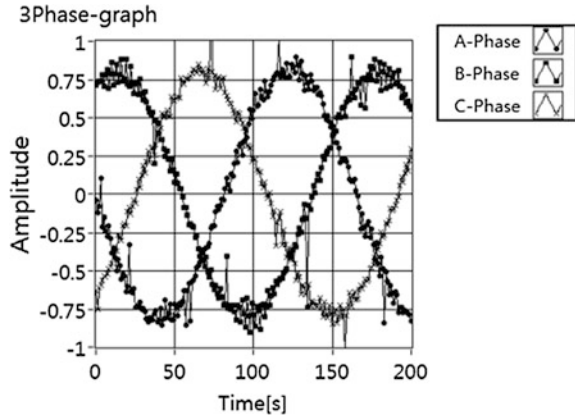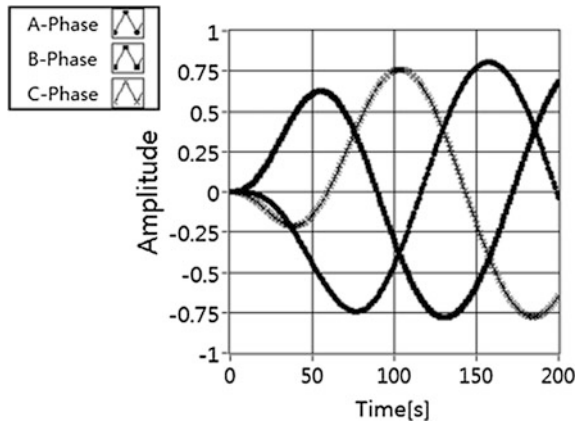
Fig. 6 Noise component of the stator current

Fig. 7 Current signal after IIR filtering

3.2 Measurement Error Range

At 60 Hz (power frequency), the amount of time for a period is approximately 0.0166–0.0167 [s] and that for a half period is 0.00833 [s]. The ideal rising time or falling time should be measured as 0.00416–0.00417 [s], half of the amount of time for the half period.

To measure this, a sampling rate and synchronization must be achieved; for this paper, data collection was performed: 10,000 at the speed of 10,000 [S/s]. With the sampling rate used for the paper, the amount of time for one period was measured as 166–167 [10^{-4} s]. The frequency for the power should be accurately expressed with decimal points—analogue data—but the errors that occur due to the measured sampling rate should be considered as follows (Table 1).

The maximum error range for a half-period is 83 [10^{-4} s] and if the falling time is 41 [10^{-4} s] as high as approximately 49.4 % can be set as the error range for the normal duty ratio when there is no particular threshold setting method of the fault diagnosis.

Table 1 Measurement error

	One period (166)	One period (167)
Half period	83	83 or 84
Rising time	41 or 42	42
Falling time	42 or 41	42

3.3 Frequency Distortion Phenomenon in Case of Stator Fault and Duty Ratio

An experiment was performed at 1690 rpm, the full-load state, to check the actual distortions of frequencies in steady state Fig. 8 (left) and when an 8-turn winding fault has occurred Fig. 8 (right); and its result is as follows.

The measurements were performed at the sampling rate of 10,000 [S/s], at frequency, the actual frequency of 60 Hz times 10,000. In Fig. 8, the x-axis represents time and $[10^{-4}$ s] represents the units.

As for the sine wave in Fig. 8, it can be seen that the amount of time for one period, approximately 167 $[10^{-4}$ s] from 116 to 283 $[10^{-4}$ s] is the same as that of one period at 60 Hz, 0.0167[s].

The result of a comparison of the rising times and falling times in the sections showing positive numbers is shown in Table 2.

The data in Table 2 are the result of the measurements performed at full-load and no-load, the measurements being performed for only one second per each state, from steady state to 8-turn winding. The Table shows duty ratios and the data are the average figures for the falling times and falling times in the frequencies—the sections of positive numbers.

As the fault becomes more serious, the amount of rising time started to increase while that of falling time decreased, and it can be seen that there is to be a change in duty ratio.
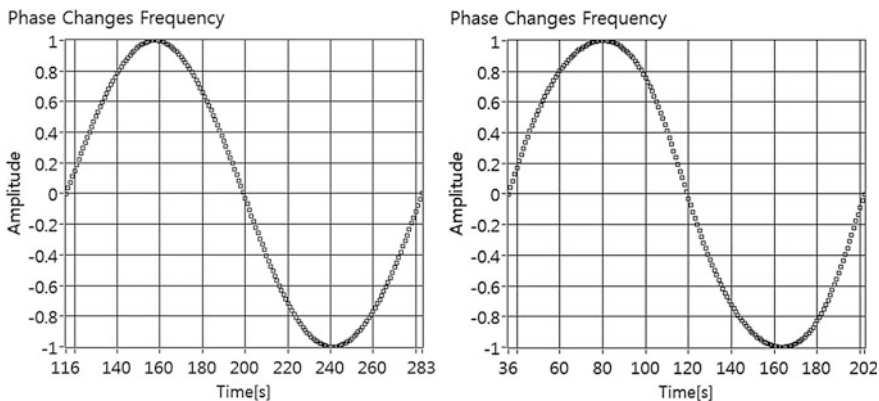

Fig. 8 Frequency of 1690 rpm by steady state (*left*), 8-turn windings (*right*)

Table 2 Average duty ratio at stator fault

	Full-load			No-load		
	Rising time	Falling time	Duty ratio (%)	Rising time	Falling time	Duty ratio (%)
Steady state	41.7736	41.6038	49.8982	41.7736	41.6226	49.9095
2-turn winding	41.9231	41.4615	49.7232	42.4118	40.9804	49.1418
4-turn winding	42.6	40.88	48.9698	43.8333	39.5208	47.4131
6-turn winding	43.2857	40.0816	48.0783	45.4091	38	45.5568
8-turn winding	44.3696	39.2391	46.9318	47.7179	35.7436	42.8265
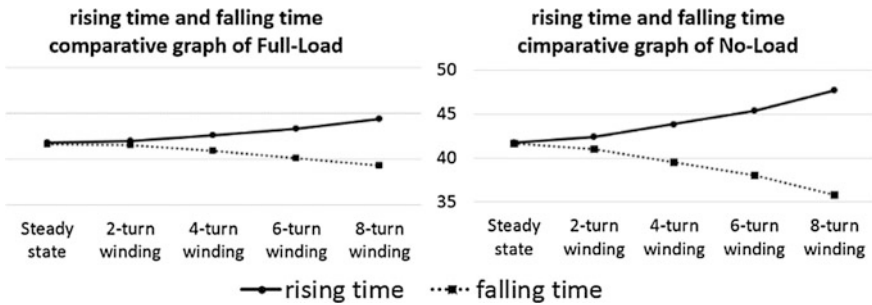
Fig. 9 Rising time and falling time comparative graph of full-load (*left*) and no-load (*right*)

Figure 9 (left) is a graph showing the result of a comparison of the rising times and falling times at full-load. Judging by each slope, it can be seen that there is no change in the overall frequency.

Figure 9 (right) is a graph showing the result of a comparison of the rising times and falling times at no-load. A pattern similar to that in Fig. 9 (left) can be seen. Therefore, it can be seen that the duty ratio more dramatically changes as the fault becomes more serious at no-load, rather than full-load.

3.4 Frequency Distortion Phenomenon in Case of Rotor Fault and Duty Ratio

A fault of one bar or two bars can occur following the occurrence of a fault of the rotor. In this case, it is recommended to perform a vibration test on the induction motor being used, as that moment is considered as the time when a fault can occur —the occurrence of approximately 36–42 dB in the existing sideband algorithm.

Table 3 Average Duty ratio at rotor fault

	Full-load			No-load		
	Rising time	Falling time	Duty ratio (%)	Rising time	Falling time	Duty ratio (%)
Steady state	41.7736	41.6226	49.9095	41.7736	41.6226	49.9095
1bar-broken	40.4808	43	51.5089	39.3396	44.0213	52.8081
2bar-broken	40.2453	43.1429	51.7374	39.5849	43.7872	52.5203
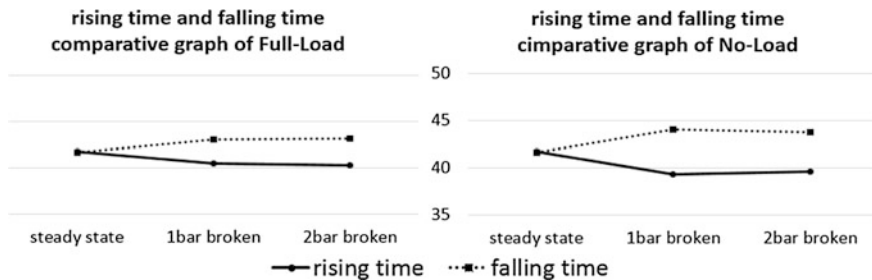
Fig. 10 rising time and falling time comparative graph of full-load (*left*) and no-load (*right*)

That is, in case of the occurrence of a 2-bar broken, that moment is considered as the time to test it for diagnosis and to repair.

Table 3 compare the characteristics of the rotors with the use of the suggested algorithm.

Table 3 compares the rising times, falling times, and duty ratios at full-load and no-load in case of the occurrence of a fault of rotor.

Figure 10 graphs that illustrate the rising times and falling times in the half-period frequencies in case of the occurrence of a fault of the rotor as in Table 3.

As you can see in Fig. 10 compared to Fig. 9, which illustrate stator faults, in case of stator fault, the ratios for rising time are higher than those for falling time; in case of rotor fault, the ratios for falling time are higher than those for rising time.

3.5   Experiment Result

For such experiment result, as you can see in Fig. 11 below, patterns above the ideal threshold value that considers the error range for the experiment environment for the paper are created:

Figure 11 is a graph showing the result of a comparison of duty ratios of average frequencies at full-load and no-load.

When considering measurement error range, it is possible to judge accurately whether a 4-turn winding fault or a more serious turn-to-turn short fault has
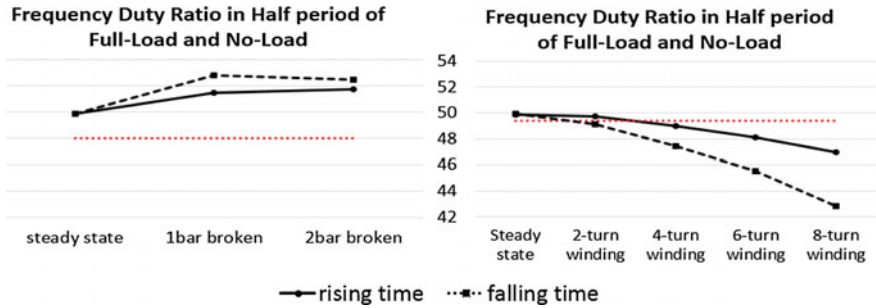
Fig. 11 Frequency duty ratio of full-load and no-load

occurred although the ideal duty ratio 49.4 % is considered as a reference threshold value (a 2-turn winding fault cannot be detected accurately in this case); and a 2-turn winding fault diagnosis can also be performed through the suggested method when being operated at no-load.

As the turn to turn short fault becomes more serious, the amount of rising time starts to increase while that of falling time decreases, and it is possible to predict a change in duty ratio.

Figure 12 compares all of the data on stator fault and rotor fault, which are electrical faults.

The result of the preceding experiment was expressed as the sum of successive average figures for frequencies measured for one second. For the measurements for Fig. 12, 10 rpm was increased every 30 s, from 1690 rpm (full-load) to 1790 rpm, and the Fig. 12 shows the successive average figures for the duty ratios of the half-period frequencies measured every 30 s.
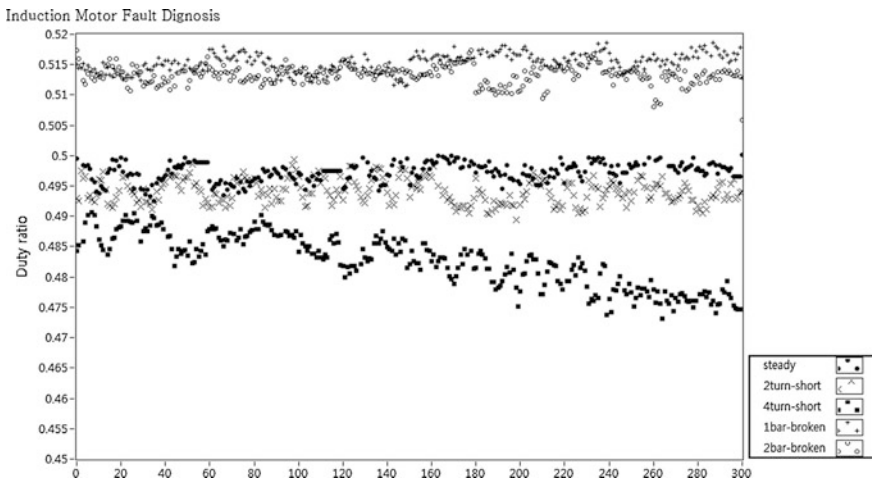
Fig. 12 Result of induction motor fault diagnosis in electrical fault

In Fig. 12, the black circle represents steady state; the 'X' mark 2 turn short fault; the quadrangle 4 turn short fault; the '+' mark 1-bar broken; and the white circle 2-bar broken.

As you can see in Fig. 12, when the cause of an electrical fault has occurred, stator fault and rotor fault (the paper's classification into these two types) can be linearly separated—the linear separation problem that the paper intended to solve.

As for stator fault, it can be seen that linear separation is impossible at the ideal threshold value that the paper suggested when a 2-turn short fault has occurred; but linear separation from steady state is possible at the value when a 4-turn short fault has occurred.

4   Conclusion

The paper suggested a new technique that identifies stator faults and rotor faults with the use of the duty ratios of the half-period frequencies created when a phase angle change occurs, induction motor faults being caused by an electrical fault.

The previous studies were performed only considering the possibility of stator fault or rotor fault; and the technique for detecting stator fault falsely identified a rotor fault as a stator fault in some cases, the technique causing serious accidents (or faults).

According to the findings of the paper, when a stator fault has occurred, the rising times of specific frequencies of the phase angle are longer, while when a rotor fault has occurred, the falling times of those are a little longer. With these characteristics, which appear because of a phase angle change, now it is possible to judge swiftly whether a fault has occurred.

In addition, as for stator fault, it could be seen that linear separation from steady state is perfectly possible in case of the occurrence of a 4-turn short fault, complete fault state.

It is judged that analysis studies on the characteristics through long-term operation will need to be performed in the future, to find out a method of identifying 2-turn short faults, the situation where the fault has not become more serious yet.

References

Arkan, M., Perovic, D. K., & Unsworth, P. (2001). Online stator fault diagnosis in induction motors. *IEE Proceedings: Electric Power Applications, 148*, 537–547.

Awadallah, M. A., Marcos, M. M., Gopalakrishnan, S., & Nehl, T. W. (2006). Detection of stator short circuits in VSI-fed brushless DC motors using wavelet transform. *IEEE Transactions on Energy Conversion, 21*(1), 1–8.

Bakhri, S., Ertugrul, N., & Soong, W. L. (2012, October). Negative sequence current compensation for stator shorted turn detection in induction motors. In *IECON 2012—38th Annual Conference on IEEE Industrial Electronics Society*.

Filippetti, F., Frauceschini, G., Tassoni, C., & Vas, P. (2006). Recent developments of induction motor drives fault diagnosis using AI techniques. *IEEE Transactions on Energy Conversion, 21*(1), 1–5.

Go, Y. J., Lee, B., Song, M. H., & Kim, K. M. (2014). A stator fault diagnosis of an induction motor based on the phase angle of park's vector approach. *Institute of Control, Robotics and Systems (in Korean), 20*(4), 408–413.

Kliman, G. B., Premerlani, W. J., Koegl, R. A., & Hoeweler, D. (2000). Sensitive, on-line turn-to-turn fault detection in AC motors. *Electric Machines and Power Systems, 28*, 915–927.

Kohler, J. L., Sottile, J., & Trutt, F. C. (2002). Condition monitoring of stator windings in induction motors: Part I—Experimental investigation of the effective negative-sequence impedance detector. *IEEE Transactions on Industry Applications, 38*, 1447–1453.

Lee, S. B., Tallam, R. M., & Habetler, T. G. (2003). A robust, on-line turnfault detection technique for induction machines based on monitoring the sequence component impedance matrix. *IEEE Transactions on Power Electronics, 18*, 865–872.

Song, M. H., Park, K. N., Han, D. G., & Yang, C. O. (2008). Auto-diagnosis for stator winding faults using distortion ratio of Park's vector pattern. *The Transactions of the KIEE (in Korean), 57*, 2.

Tallam, R. M., Habetler, T. G., & Harley, R. G. (2002). Transient model for induction machines with stator winding turn faults. *IEEE Transactions on Industry Applications, 38*, 632–637.

Tallam, R. M., Habetler, T. G., Harley, R. G., Gritter, O. J., & Burton, B. H. (2000). Neutral network based on-line stator winding turn fault detection for induction motors. In *Proceeding of IEEE IAS Annual Meeting, YoU*, pp. 375–380.

Mathematical Foundations and Software Simulation of Stress-Strain State of the Plate Container Ship

Anatoliy Nyrkov, Sergei Sokolov, Valery Maltsev and Sergei Chernyi

Abstract The article is devoted to construction of mathematical models for analysis of stress-strain state of the ship's plate. The model is based on the theory of plates pocket on the basis of mathematical models created a complex software, which together allow us to speak about building a handy tool both for research and for practical use, which allows to quickly calculate and evaluate the fatigue-stress state of the ship's deck. As a practical example taken container for which it was solve stress-strain state of its hatches, taking into account load containers on deck.

Keywords Marine · Fitting · Deformation · Simulates

1 Introduction

Structural mechanics of the ship, as an independent science, started at the beginning of the XX century. Based on previous knowledge of the theory of elasticity I. Bubnov were offered the first rules of allowable stresses for surface ships, developed strength evaluation methods and stability of the marine floors and backed plates (Mathai et al. 2013; Bojcov et al. 1982; Han et al. 2014; Goloskokov 2004). Most calculation methods have not changed significantly, but have been tested a number of experimental tests and structured in the form of formulas and tables of acceptable values. All of this documentation can be found on the website of the Russian Registry of Shipping.

Changing the shape and size of the hull or its parts caused by external forces or other influences (e.g., heating, cooling). Distinguish between General and local deformation. When the overall body is deformed as the hollow beam of variable

A. Nyrkov · S. Sokolov · V. Maltsev · S. Chernyi (✉)
Admiral Makarov State University of Maritime and Inland Shipping,
St. Petersburg, Russia
e-mail: sergiiblack@gmail.com

S. Chernyi
Kerch State Maritime Technological University, Kerch, Russia

cross section and may have a common longitudinal bending, transverse shear total and total torsion. General buckling is characterized by changes of curvature of the neutral axis of the hull in the vertical and horizontal planes of the ship, the total transverse shear—rotation normal to the cross section of the hull relative to the neutral axis, and the total torsion—angle of twist relative cross sections of the hull (Mathai et al. 2013; Bojcov et al. 1982; Han et al. 2014).

The peculiarity of the manufacture of ship hull structures, having large dimensions, is the use of a significant amount of manual fitting works at all stages of the technological process (in the manufacture of nodes, sections, during the construction of the hull on the slipway). Fitting work impact the quality and increase the manufacturing cost of structures.

Residual welding deformations and stresses can have an impact on technological and structural strength and therefore, further development and improvement of methods of calculation of welding deformations and stresses, it is important to assess the impact of "technological factor" strength of ship hull structures.

Also it is necessary to conduct studies of errors in the manufacture of hull constructions and the possibility of reducing the amount of fit during Assembly of the mounting housing connections based on probabilistic dependencies, volume fit from the values of the tolerances of the sizes of sections), the tolerances of the sections from the base plane and the tolerances on the displacement of the connected elements.

2 Modern Tools Used in the Design of Ships

Ship design is a complex process that must take into account the huge number of parameters such as stability, the influence of external forces, susceptibility to corrosion, stresses arising during operation, etc. For most applications engineering offices are increasingly using specialized software that allows you to lack the deep scientific knowledge, to make the necessary calculations.

Today a variety of such programs is sufficiently large: ANSYS, MatLab, Comsol, WinMachine, SolidWorks, etc. All of these products are not just designs and actually use the tool based on the production companies such as ABB, BMW, Boeing, Caterpillar, Daimler-Chrysler, Exxon, FIAT, Ford, BelAZ, General Electric, Lockheed Martin, Mitsubishi, Siemens, Shell, Volkswagen-Audi and others., and used in many of the leading industrial enterprises of Russia (Goloskokov 2004; Malyh 2012).[1] The benefits of using such software solutions are quite large. This is the lack of need for a mathematical analysis of the strength characteristics of objects and elements, a visual representation of the most vulnerable areas in terms of strength, obtaining numerical values of the necessary

---

[1]*The free encyclopedia, ANSYS*, url: https://ru.wikipedia.org/wiki/ANSYS.

parameters, there is no need for additional expensive and long-term experiments. But all of these advantages can give a rise of negative effects.

The computer power may be different, but in any case it is imperfect, resulting in a computational errors occur to estimate that is not possible, because the methods of calculation are hidden from the user. Lack of proper mathematical training specialist in charge of calculation does not allow for analysis of the result values (Malyh 2012).

3   Mathematical Model

There are always analytical methods to obtain the result of the formula calculation without errors or inaccuracies that may be explicitly represented and make the appropriate assessment, as an alternative to such software methods of calculation. Among the disadvantages of these methods for solving the problem is clearly seen the impossibility of creating a universal algorithm for calculation and, accordingly, the complexity of building software to solve this problem.

As an example of an alternative calculation method is invited to consider the finite difference method (grid method). The idea of the finite difference method is known for a long time, with the relevant works of Euler differential calculus. However, the practical application of this method was very limited because there are huge amount of manual calculations related to the dimension of the resulting system of algebraic equations, which solving required years. Nowadays, with the advent of modern high-speed computers, the situation has changed radically. This method has become convenient for practical use, and is one of the most effective in solving various problems of mathematical physics.

The object of research will be the ship plate, which simulates the flat of the ship. Traditional methods of structural mechanics of the ship in the calculation of stress-strain state is considered as an equivalent hull beam, i.e., finite stiffness beam. Mathematical laws, which describing the marine plate problems, will be different from the traditional ones (see Footnote 1) (Malyh 2012; Postnov et al. 1987).

In the study of the stress state of the plates we use a Cartesian coordinate system, combining with the median plane of the XOY plane of the plate. The theory of bending of thin plates based on Kirchhoff's hypothesis. Since the deformation of the plate w is larger displacements u, v, all movements considered to be small, and thus can be neglected nonlinear terms with respect to u and v, and replace the two members of the radical binomial. Deformation $\varepsilon_x, \varepsilon_y, \gamma_{xy}$ will have the following form:

$$
\begin{cases}
\varepsilon_x = \frac{\partial u}{\partial x} + \frac{1}{2}\left(\frac{\partial w}{\partial x}\right)^2 \\
\varepsilon_y = \frac{\partial v}{\partial y} + \frac{1}{2}\left(\frac{\partial w}{\partial y}\right)^2 \\
\gamma_{xy} = \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} + \frac{\partial w}{\partial x} \cdot \frac{\partial w}{\partial y}
\end{cases} .
\tag{1}
$$

For points of plate lying in a layer $z = $ const relation between displacements and deformations based on the hypothesis established direct normals. The deformations of the middle surface are related by compatibility:

$$\frac{\partial^2 \varepsilon_x}{\partial y^2} + \frac{\partial^2 \varepsilon_y}{\partial x^2} - \frac{\partial^2 \gamma_{xy}}{\partial x \partial y} = \left(\frac{\partial^2 w}{\partial x \partial y}\right)^2 \cdot - \frac{\partial^2 w}{\partial x^2} \cdot \frac{\partial^2 w}{\partial y^2}. \tag{2}$$

These dependencies are valid when the deformations are small. Consider a cross section perpendicular to the axis Ox and Oy. The state of stress can be characterized by the efforts of the plate (Fig. 1) per unit length of the corresponding section. All of these forces—the essence of the intensity of the forces applied to the surface of the median line after reducing her stress (see Footnote 1) (Mathai et al. 2013; Jasnickij 2005).[2]

On the Fig. 1 T_x, T_y—normal force, shows a relationship:

$$T_x = \int\limits_{-h/2}^{h/2} \sigma_x dz; \quad T_y = \int\limits_{-h/2}^{h/2} \sigma_y dz \, x + y = z. \tag{3}$$

$S = S_x = S_y = \int_{-h/2}^{h/2} \tau_{xy} dz$—shear force, defined in sections $x = $ const, $y = $ const the same formulas determined by the law of pairing shear stresses S_x. N_x, N_y—shearing forces:

$$N_x = \int\limits_{-h/2}^{h/2} \tau_{xz} dz; \quad N_x = \int\limits_{-h/2}^{h/2} \tau_{yz} dz. \tag{4}$$

In the sections $x = $ const and $y = $ const operating stress within the unit of length creates the following moments (Fig. 2):

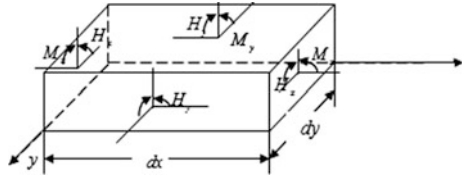M_x, M_y—bending moments, acting in sections $x = $ const и $y = $ const:

Fig. 1 Positive directions of forces in accordance with the rule of signs of stress
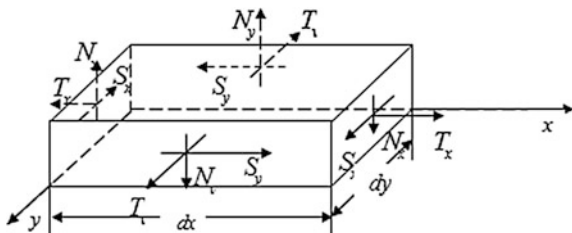
Fig. 2 Positive direction of
moments

$$M_x = \int\limits_{-h/2}^{h/2} \sigma_x \cdot z dz \quad M_y = \int\limits_{-h/2}^{h/2} \sigma_y \cdot z dz. \tag{5}$$

$H = H_x = H_y$—to que moment sin sections $x = $ const, $y = $ const:

$$H = \int\limits_{-h/2}^{h/2} \tau_{xy} \cdot z dz. \tag{6}$$

With the vanishing of the main vector and the main moment, we obtain the scalar equations of equilibrium:

$$\begin{cases} \frac{\partial T_x}{\partial x} + \frac{\partial S}{\partial y} = \frac{\partial}{\partial x}\left(N_x \frac{\partial w}{\partial x}\right) + \frac{\partial}{\partial y}\left(N_y \frac{\partial w}{\partial x}\right) + q\frac{\partial w}{\partial x} \\ \frac{\partial S}{\partial x} + \frac{\partial T_y}{\partial y} = \frac{\partial}{\partial x}\left(N_x \frac{\partial w}{\partial y}\right) + \frac{\partial}{\partial y}\left(N_y \frac{\partial w}{\partial y}\right) + q\frac{\partial w}{\partial y} \\ \frac{\partial N_x}{\partial x} + \frac{\partial N_y}{\partial y} = -\frac{\partial}{\partial x}\left(S \frac{\partial w}{\partial y}\right) - \frac{\partial}{\partial y}\left(S \frac{\partial w}{\partial x}\right) - \frac{\partial}{\partial x}\left(T_x \frac{\partial w}{\partial x}\right) - \frac{\partial}{\partial y}\left(T_y \frac{\partial w}{\partial y}\right) - q \end{cases}. \tag{7}$$

$$N_x = \frac{\partial M_x}{\partial x} + \frac{\partial H}{\partial y}; \quad N_y = \frac{\partial M_y}{\partial y} + \frac{\partial H}{\partial x}. \tag{8}$$

Using Airy stress function, allows to determine the stress from the formulas:

$$\sigma_x = \frac{T_x}{h} = \frac{\partial^2 \Phi}{\partial y^2}; \quad \sigma_y = \frac{T_y}{h} = \frac{\partial^2 \Phi}{\partial x^2}; \quad \tau_{xy} = \frac{S}{h} = -\frac{\partial^2 \Phi}{\partial x \partial y}. \tag{9}$$

The system of equations become equivalent to the single equation:

$$\frac{\partial^2 M_x}{\partial x^2} + 2\frac{\partial^2 H}{\partial x \partial y} + \frac{\partial^2 M_y}{\partial y^2} = -q - h\left(\frac{\partial^2 \Phi}{\partial y^2} \cdot \frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 \Phi}{\partial x^2} \cdot \frac{\partial^2 w}{\partial y^2} - 2\frac{\partial^2 \Phi}{\partial x \partial y} \cdot \frac{\partial^2 w}{\partial x \partial y}\right). \tag{10}$$

Thus, the behavior of the plate is described by two resolving equations: the equation of balance (10) and strain compatibility Eq. (2). The left side of the equilibrium equation can be expressed only through a deformation in the equation

of compatibility of deformation—only through function Airy. Using Hooke's law for isotropic material deformation compatibility Eq. (2) can be represented as follows:

$$\Delta\Delta\Phi = E\left[\left(\frac{\partial^2 w}{\partial x \partial y}\right)^2 - \frac{\partial^2 w}{\partial x^2} \cdot \frac{\partial^2 w}{\partial y^2}\right]. \tag{11}$$

Introducing in equilibrium Eq. (10) the expression for the bending and twisting moments, we can get a differential equation of the form:

$$D\Delta\Delta w = q + h\left(\frac{\partial^2 \Phi}{\partial y^2} \cdot \frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 \Phi}{\partial x^2} \cdot \frac{\partial^2 w}{\partial y^2} - 2\frac{\partial^2 \Phi}{\partial x \partial y} \cdot \frac{\partial^2 w}{\partial x \partial y}\right). \tag{12}$$

The two Eqs. (11) and (12) provide a resolution system of differential equations of the Karman's theory of plate:

$$\begin{cases} \Delta\Delta\Phi = E\left[\left(\frac{\partial^2 w}{\partial x \partial y}\right)^2 - \frac{\partial^2 w}{\partial x^2} \cdot \frac{\partial^2 w}{\partial y^2}\right] \\ D\Delta\Delta w = p(x, y) + h \cdot \left(\frac{\partial^2 \Phi}{\partial y^2} \cdot \frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 \Phi}{\partial x^2} \cdot \frac{\partial^2 w}{\partial y^2} - 2\frac{\partial^2 \Phi}{\partial x \partial y} \cdot \frac{\partial^2 w}{\partial x \partial y}\right) \end{cases}. \tag{13}$$

where $\Delta\Delta = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) = \frac{\partial^4}{\partial x^4} + 2\frac{\partial^4}{\partial x^2 \partial y^2} + \frac{\partial^4}{\partial y^4}$.

This system of equations isn't widely used for practical calculations, and its solutions brought to the numerical reference data are scarce because of the complexity of the calculations for solving systems with large number of unknowns (Mathai et al. 2013; Bojcov et al. 1982). In modern conditions of rapid development of information technology, such calculations can be trusted computing (see Footnote 2) (Postnov et al. 1987; Jasnickij 2005; Chernyi et al. 2015; Zhilenkov and Chernyi 2015; Chernyi 2016).

4 Building Software and Example of the Proposed Model

For example, implementation of the decision of the system use the task of loading a container with some of the cargo. As the ship type was taken "Artist Saryan". It has 6 holds various sizes (Fig. 3). Each hold closed hatch, which in turn is placed in the same load. The greatest interest for calculating hatches are the fourth and the third holds, as their dimensions are sufficiently large and the load is adequately high. These hatches consist of four identical elements 12.96 m long and 10.7 m wide. With the help of modern software will make evaluations and calculations (see Footnote 2) (Chernyi et al. 2015; Zhilenkov and Chernyi 2015; Chernyi and Zhilenkov 2015a, b; Chernyi 2015; Chernyi 2016).
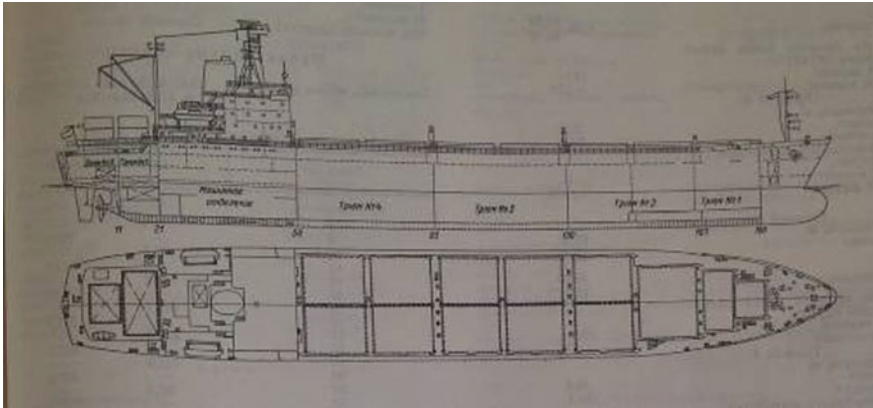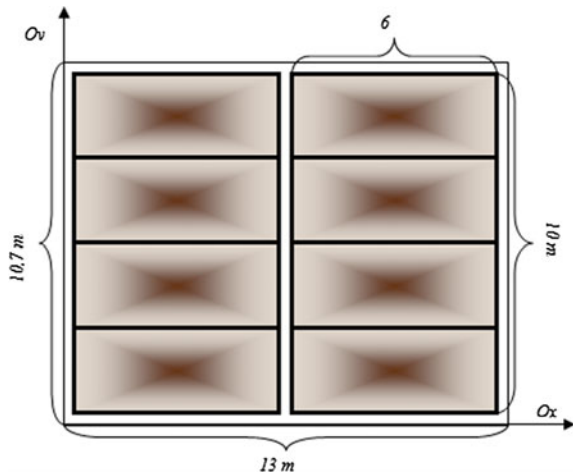
Fig. 3 Scheme of the ship such as "Artist Saryan"

Fig. 4 The layout containers

View on the scheme of arrangement of sixteen twenty-foot containers in two stacks of 8 each (Fig. 4).

In the simulation the problem of deformation plates of the main deck is supposed to free support, then the boundary conditions have no bending moments and deformations at the edges of the plate:

$$w|_\Gamma = 0$$

$$\frac{\partial^2 w}{\partial x^2} + v\frac{\partial^2 w}{\partial y^2}\bigg|_{\Gamma_1} = 0 \tag{14}$$

$$\frac{\partial^2 w}{\partial y^2} + v\frac{\partial^2 w}{\partial x^2}\bigg|_{\Gamma_2} = 0.$$

where Γ—edge of the plate, and $\Gamma = \Gamma_1 \cup \Gamma_2$, where Γ_1—edge of plate which parallel axe Oy, a Γ_2—Ox.

Boundary conditions for Airy function:

$$\frac{\partial^2 \Phi}{\partial y^2} \cdot \cos\left(\stackrel{\wedge}{n,x}\right) + \frac{\partial^2 \Phi}{\partial x \partial y} \cdot \cos\left(\stackrel{\wedge}{n,y}\right) = F_1$$

$$-\frac{\partial^2 \Phi}{\partial x \partial y} \cdot \cos\left(\stackrel{\wedge}{n,x}\right) - \frac{\partial^2 \Phi}{\partial x^2} \cdot \cos\left(\stackrel{\wedge}{n,y}\right) = F_2. \tag{15}$$

where F_1 и F_2—stresses on Γ.

In view of these conditions has been established software package, as a result of which the numerical values have been obtained strain and stresses, as well as corresponding graphs (Fig. 5).

The maximum deflection in this example was 0.00708907 m and the highest stresses is achieved at the corners of the plate. Accuracy of the method in this case will be equal to h^2, where h—step decomposition.

Fig. 5 The diagram of strain state
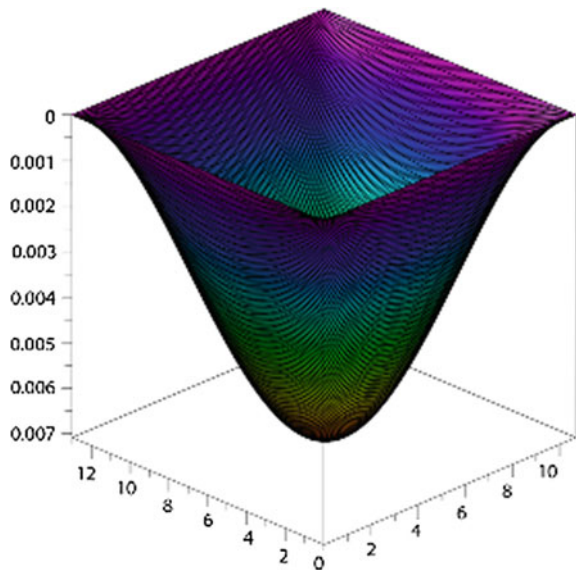
In generalizing risk assessments, it is recommended to analyze uncertainty and accuracy of the results. In most cases, a source of uncertainty in the case of the processes of deformation will be incomplete data on production technology, reliability and precise operation of the equipment as well as human errors. In order to interpret the results of the risk assessment properly, it is recommended to understand the nature of the uncertainties and their causes. Sources of uncertainty are recommended to be identified and evaluated and results to be presented. The authors carried out a qualitative and quantitative risk assessment. Qualitative risk analysis reveals the sources and causes of risk, stages and work in the course of which a risk may arise (finding potential risk areas, identifying activities associated with risks, finding practical benefits and negative effects of identified risks). Quantitative analysis provides numerical values of both individual risks and for the entire facility. It is also possible to calculate the potential damage and to develop a system of preventive measures aimed at minimizing the occurrence of risk situations. In the risk analysis, risk matrix and generic risk assessment algorithm were used.

5   Conclusion

As a result of this work have been formulated some mathematical basis for calculating the stress-strain state of the ship's deck plates covers a container, was created support software automates the process of calculation and visualization of results. This calculation was carried out without taking into account the cross-beam set for which the plate is mounted. Accounting for this fact affect the result and thereby improve the accuracy. Given all the design features of the deck, it is possible, on the basis of the proposed models to create a universal software to calculate the fatigue-stress state of the deck of a ship, and to choose the right function of the load can be seen loads of any type, not just the container. The data obtained will help to draw a conclusion not only on the current loading of the deck element, but also to assess the likelihood of damage in the future and the need to take appropriate action.

References

Bojcov, G. V., Palij, O. M., Postnov, V. A., & Chuvikovskij, V. S. (1982). *Handbook of structural mechanics of the ship in three volumes, Plates.* Theory of elasticity, plasticity and creep. Numerical methods, Shipbuilding. Vol. 2.

Chernyi, S. (2015). The implementation of technology of multi-user client-server applications for systems of decision making support. *Metallurgical and Mining Industry, 3*, 60–65.

Chernyi, S., & Zhilenkov, A. (2015a). Analysis of complex structures of marine systems with attraction methods of neural systems. *Metallurgical and Mining Industry, 1*, 37–44.

Chernyi, S., & Zhilenkov, A. (2015b). Modeling of complex structures for the ship's power complex using XILINX system. *Transport and Telecommunication, 16*(1), 73–82.

Chernyi, S., Zhilenkov, A., Sokolov, S., & Titov, I. (2015). Seif-contained drilling rig automatic control system efficiency improvement by means of assuring compatibility and integration methods development. *Metallurgical and Mining Industry, 7*(3), 66–73.

Chernyi, S. (2016). *Analysis of the energy reliability component for offshore drilling platforms within the Black Sea* (vol. 1 Issue 2, pp. 106–110). Neftyanoe Khozyaystvo - Oil Industry.

Goloskokov, D. P. (2004). Uravneniya matematicheskoy fiziki. Resheniye zadach v sistemeMaple [Equations of mathematical physics. The solution of tasks in the system Maple]. Saint-Petersburg: Piter, 538.

Han, F. L., Wang, C. H., Hu, A. K., & Liu, Y. C. (2014). Fatigue strength assessment analysis of large container ship. *Applied Mechanics and Materials, 602–605*, 385–389.

Jasnickij, L. N. (2005). For whom the tolls ANSYS, or why so many were falling aircraft missiles to explode, collapse of the building. 1(342).

Malyh, M. D. (2012). *Proceedings of the seminar Finite Element Method on the example of the first boundary value problem for the Poisson equation*. Physics Department of Moscow State University.

Mathai, A., John, G. P., & Jin, M. J. (2013). Ultimate torsional strength analysis of container ship. *International Journal of Engineering Science and Technology., 5*(3), 512–518.

Postnov, V. A., Rostovcev, D. M., Suslov, V. P., & Kochanov, J. P. (1987). Structural mechanics of the ship and the theory of elasticity: The Textbook for high schools: in 2 volumes. *Shipbuilding, 2*.

Zhilenkov, A., & Chernyi, S. (2015). Investigation performance of marine equipment with specialized information technology. *Procedia Engineering, 100*, 1247–1252.

Kalman Filtering for Precise Mass Flow Estimation on a Conveyor Belt Weigh System

Tauseef Rehman, Waleed Tahir and Wansoo Lim

Abstract Conveyor belt weigh systems are widely used worldwide in industry for mass flow estimation. This paper investigates the application of Kalman filtering for tachometer response correction and thus accurate flowrate measurement. The tachometer is a sensor with a pulse-train output with a frequency proportional to the conveyor belt rpm. Under harsh conditions, as are generally found in the industry, the pulse-width of the tachometer output is susceptible to noise, thus corrupting the conveyor belt's speed measurement. Thus, a Kalman filter has been employed for accurate estimation of the conveyor belt speed, and thus, the mass-flow estimate. To facilitate this investigation, the belt conveyor system (plant) and the Kalman filter were initially modeled and simulated in MATLAB Simulink. This was followed by development of the Kalman filter based mass flow estimator on a Zynq-7000 based Digilent Zedboard, and interfacing it with a Speedgoat Realtime Target Machine (RTM) on which the plant model ran in real-time. The results are presented at the end, demonstrating the effectiveness of the proposed estimation technique implemented on actual hardware.

Keywords Mass flow estimation · Conveyor weigh system · Belt scale · Robust estimation · Kalman filtering

1 Introduction

One of the most extensively used systems for weight measurement and mass flow estimation, are conveyor belt weigh systems. They are widely employed in the manufacturing, pharmaceutical, packaging, chemical, automotive, food and mining industries. In certain applications, especially in the mining and minerals industry

T. Rehman (✉) · W. Tahir · W. Lim
Embedded Systems Group, Burqstream Technologies, Islamabad, Pakistan
e-mail: tauseef@burqstream.com

Fig. 1 Conveyor belt system being used in a stacker at a mining site

(Fig. 1), accurate weight and mass flow estimation is an integral requirement for the conveyor belt systems. In mass production applications, where very large quantities (in the order of hundreds of tons) of material is to be transported using large scale conveyor belt systems, having an accurate weight and mass flow estimation system in place is essential. A higher accuracy of the weigh system benefits both the producer as well as the consumer. Profitability is increased for the producer as a result of higher manufacturing efficiency, while for the customer, quantity and package quality are assured.

The weight of the product is estimated while it is carried over a weigh-table using a transport system. For mass production systems, usually this transportation system is a conveyor belt. A load cell, placed under the weigh-table, measures the load of the product. Using this measurement from the load cell coupled with the belt speed of the conveyor, the mass flow rate of the product can be estimated (Vayrynen 2013). Historically, computationally intensive signal conditioning was not possible due to performance limitations of embedded processors. However, with the advent of low-cost high performance embedded processors, advanced digital processing techniques can be widely employed for signal conditioning and accurate mass flow estimation in conveyor belt systems.

1.1 Background

There are two main constraints faced in achieving a higher accuracy in flowrate measurement using conveyor-belt-based dynamic weighing systems. First is a superimposed noise on the load original signal from the load-cell, and the other is a noise that is superimposed on the useful tachometer signal. The noise in the load-cell measurements is predominantly caused due to mechanical vibrations and electrical noise and can be catered for by numerous filtering techniques as demonstrated by Halimic M., Robert L., and Jafaripana et al. In this research, our point of focus in this research is the noise corruption in the tachometer signal, whereas the load-cell measurement noise has been catered for using an effective moving average filter.

Ideally, the tachometer output for a conveyor belt moving with a constant velocity should be in the form of a pulse-train with a constant frequency. However, while dealing with practical systems, it is found that this actually is not the case. The tachometer pulse-train output in actual large scale conveyor systems is observed to have a time-varying frequency (Fig. 2).

This is caused predominantly due to mechanical factors e.g. non-alignment of the rotary shaft axis with the actual axis of rotation. While mechanical correction may be a trivial solution for the case of small scale conveyor systems, it becomes a tedious and time consuming task for large scale systems as in Fig. 1. Furthermore, it is observed that even after mechanical correction, such a problem is highly likely to arise again due to the high loads sustained by the large mechanical components during strenuous operation.

In order to perform the correction in a computer processor, one method may be to calculate the average pulse-rate in one complete frequency oscillation cycle (Fig. 2b) and use it in flowrate computation. However, averaging the pulse-rate
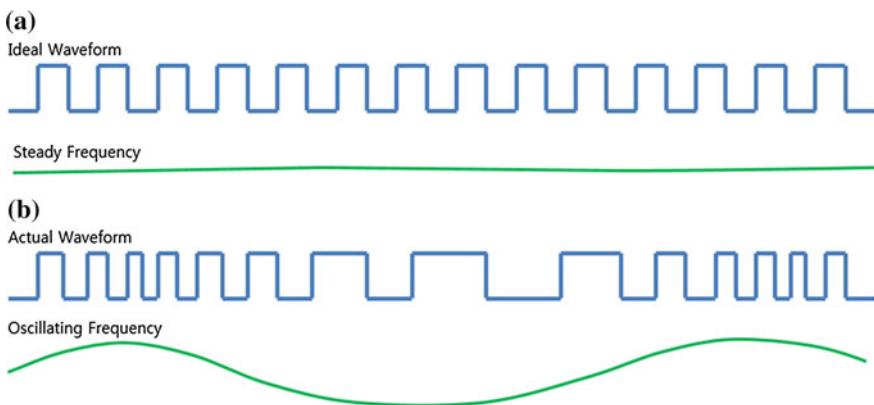

Fig. 2 **a** Ideal tachometer waveform for large scale conveyor systems (constant frequency). **b** Actually observed tachometer for large scale conveyor systems (time-varying frequency)

over such a relatively large time period adversely affects the instantaneous belt-speed estimate. Thus, a more robust and intelligent mechanism for filtering was desired. Kalman filter is one such tool which is employed widely for true value estimation from a set of noisy measurements. It uses estimated values from a mathematical model of the system coupled with actual measurements from the plant and fuses them to produce an accurate estimate with minimal variance. In the next section, we discuss the use of Kalman filter to determine the true instantaneous belt velocity from our measured value.

2 Filtering Approach

Kalman filtering is a robust estimation technique used widely for applications requiring accurate measurement. It works by intelligently fusing the actual measurements from a noisy sensor (Harb 2013), which in our case happens to be a tachometer, with the estimated measurement from a mathematical model of the system, which is the conveyor belt system in our case. The accurate estimate of the conveyor belt velocity at any time 'k' is given by the following equation (Catlin 1989):

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-). \tag{1}$$

where
$\hat{x}_k$ A posteriori (corrected) estimate of state variable
$\hat{x}_k^-$ A priori (predicted) estimate of state variable
z_k System output measurement
H Output matrix
K_k Kalman gain

The Kalman gain is computed recursively in every computation cycle such that it minimizes the error covariance (Louis et al. 2007). Suppose we have a linear system model for the conveyor belt system given by the following state space description:

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1} + w_{k-1}. \tag{2}$$

$$z_k = H\hat{x}_k + v_k. \tag{3}$$

where
$\hat{x}_k^-$ A priori (predicted) estimate of state variable
$\hat{x}_{k-1}$ A posteriori (corrected) estimate of state variable of previous step
A System matrix
B Input matrix
u_{k-1} Input of the system (forcing function)
z_k Measured output of the system (conveyor belt velocity)
w_{k-1} System Noise with probability distribution $N(0, Q)$
v_k Measurement Noise with probability distribution $N(0, R)$

It is evident from Eqs. (2) and (3) that neither the system model, nor the sensor measurement can be completely relied upon, for an accurate estimate of the system output, due to the inherent system noise, modelling inaccuracies, and measurement noise. The Kalman filter optimally filters this noise in a recursive two-step process. For every iteration, there is a 'prediction' step, which is followed by a 'correction' step (Chen 1994). Initially, in the prediction step, the system model is used to get an initial, or 'a-priori' estimate of the state matrix (Eq. 2). This is followed by the computation of the Kalman gain 'K_k' such that the a priori error covariance is minimized:

$$K_k = P_k^- H^T \left(H P_k^- H^T + R \right)^{-1}.$$ (4)

$$P_k^- = A P_{k-1} A^T + Q.$$ (5)

where
P_k^- A-priori error covariance
P_{k-1} A-posteriori error covariance of previous iteration
Q Process noise covariance
R Measurement noise covariance

Next, in the correction or update step, this Kalman gain is then used to apply a correction to the previously computed a priori state estimate to get an 'a-posteriori' estimate of the state matrix (Eq. 1). This Kalman filter cycle is depicted in Fig. 3.

The experimental setup and implementation of the filtering technique for mass flow estimation on the conveyor belt system has been highlighted in the next section.
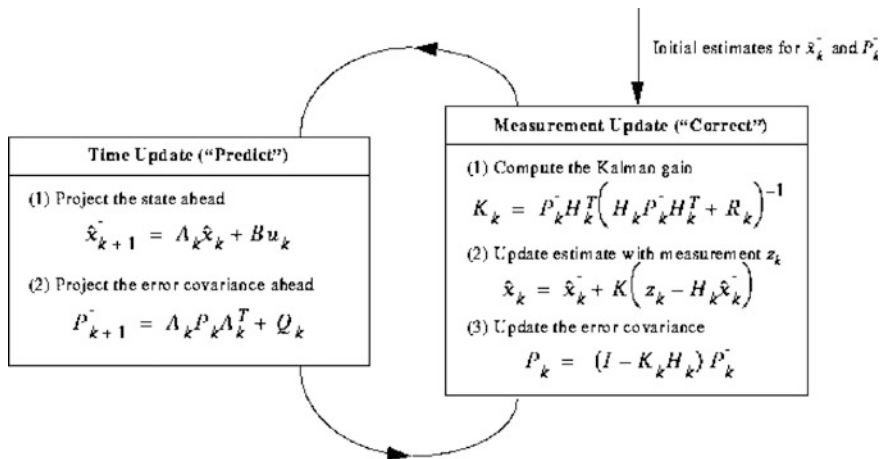
Fig. 3 Operational diagram of the Kalman filter depicting recursive transition between the prediction and correction update steps

3 Modeling

3.1 System Modeling

As highlighted in Sect. 2, the state space model of a physical system is necessarily required in order to apply a Kalman filter for estimation. In our particular case, the system is a conveyor belt system. The quantity to be estimated by the Kalman filter is the system output i.e. they conveyor belt speed. Modeling using Newtonian dynamical equations overlooks numerous seemingly negligible factors which effect the system e.g. friction. However, in actual systems, the effects of these small factors add up to have a significant effect, resulting in an inconsistency between the mathematical model and the actual system itself. Keeping this fact in mind, this research used such a modeling approach in which the mathematical model of the system closely mimicked the actual belt conveyor system, even whilst incorporating friction and other such effects.

Initially the system output was synthetically generated as a first order response to a step input forcing function. This corresponding input-output data was then used for estimation of an accurate discrete state space model for the belt conveyor system. It is noteworthy that as the actual input-output values from the system are being utilized for the modeling purpose, factors which are occasionally neglected in mathematical modeling, like frictional effects and system noise, are very much catered for in the resulting system model. Furthermore, it is highlighted that the system is considered to be linear and time invariant during this process. The system matrix 'A', input matrix 'B', and output matrix 'C' for our system are given as follows:

$$A = \begin{bmatrix} 0.975 & -1.953\mathrm{e}{-7} \\ 2.382\mathrm{e}{-7} & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} 256 \\ 0 \end{bmatrix}$$

$$H = \begin{bmatrix} 0 & 404.522 \end{bmatrix}$$

Thus, the state space representation for our system becomes:

$$\hat{x}_{k+1} = \begin{bmatrix} 0.975 & -1.953\mathrm{e}{-7} \\ 2.382\mathrm{e}{-7} & 0 \end{bmatrix} \hat{x}_k + \begin{bmatrix} 256 \\ 0 \end{bmatrix} u_k + w_k. \tag{6}$$

$$z_k = \begin{bmatrix} 0 & 404.522 \end{bmatrix} \hat{x}_k + v_k. \tag{7}$$

It can be seen from the state equations above that we have a second order state space model for the dynamic conveyor belt system. Since the mass flow estimator described in this paper was being designed for an actual industrial system, thermal

heat dissipation and continuous operation in a harsh environment, e.g. a mining site, and other such factors were carefully considered in selecting the order of the state space system. Increasing the order of the system by one point exponentially increases the computational load of the Kalman filter on the low-power embedded platform due to higher order matrix multiplications. After experimentation with systems of various orders to meet the hardware design specification of 50 % max continuous CPU utilization of the dual-core ARM Cortex-A9 present on the Zynq-7000 SoC, a second order state space system was finally selected for the conveyor belt model.

It is further highlighted that a separate 6th order system model was also computed for running on a Speedgoat real-time target machine (RTM), as will be explained in Sect. 4.1. The processing bandwidth on an RTM is significantly higher than that of the ZedBoard, thus a 6th order system model was used for more accurate simulation of the actual system.

3.2 Input Modeling

In addition to modeling the conveyor belt system, modeling the forcing function under which the actual system is expected to operate is also necessary. In our particular case, system input is the input to the DC motor of the conveyor belt. Furthermore, the targeted industrial systems are operated with a uniform belt velocity, as generally is the case with such systems. One example of this is the coal conveyors used in large scale power plants.

Thus, the input was modeled as two separate forcing functions. The first is a simple Heaviside unit step function defined as:

$$f_1(t) = \begin{cases} 0 & for.t < 0 \\ V_b/2 & for.t = 0 \\ V_b & for.t > 0 \end{cases}$$

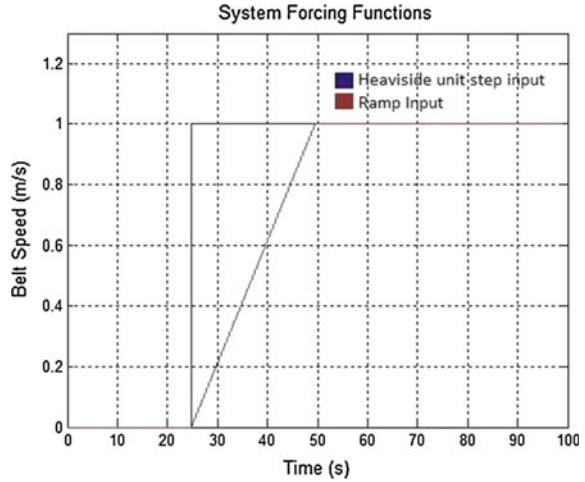The second forcing function for the conveyor belt system is a ramp function:

$$f_2(t) = \begin{cases} 0 & for.t < 0 \\ t & for.0 \leq t \leq T \\ V_b & for.t > T \end{cases}$$

where
V_b Steady state belt velocity
T DC motor time to achieve steady state velocity (Fig. 4)

Fig. 4 Conveyor belt speed modeled as heaviside unit step, and ramp input

4 Experimentation

4.1 Experimental Setup

After verification of the system and input forcing function models via simulation in MATLAB Simulink, experimentation was done on hardware. The experimental setup consisted of a Speedgoat performance real-time target machine (RTM) running a 6th order discrete system model at 10 kHz. An IO104 analog interface card was used with the RTM to output the simulated tachometer and loadcell signals. System noise was also modeled and added into the 6th order state space model of the system running on the RTM, which then closely mimicked an actual industrial conveyor belt system for the estimator.

A Digilent ZedBoard was interfaced with the RTM via Pmod AD5 for flowrate estimation based on the simulated tachometer and loadcell signals from the RTM (Fig. 5). The conveyor belt velocity was accurately estimated on the ZedBoard by the Kalman filter, while a moving average filter was employed for the filtering of the loadcell signal. The mass flowrate was then estimated at the required 10 Hz, whose results are presented in the next section.

4.2 Results

For the experimentation, the plant model ran on the RTM as a real-time simulation while the ZedBoard connected to the RTM via 2 channel A/D convertor served as the estimator. The Kalman filter running on the ZedBoard estimated the conveyor belt velocity with high accuracy, which was then fused with the filtered load-cell
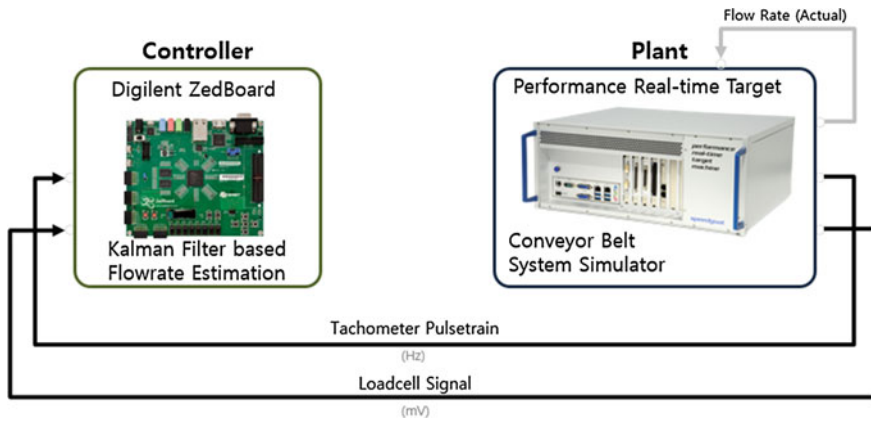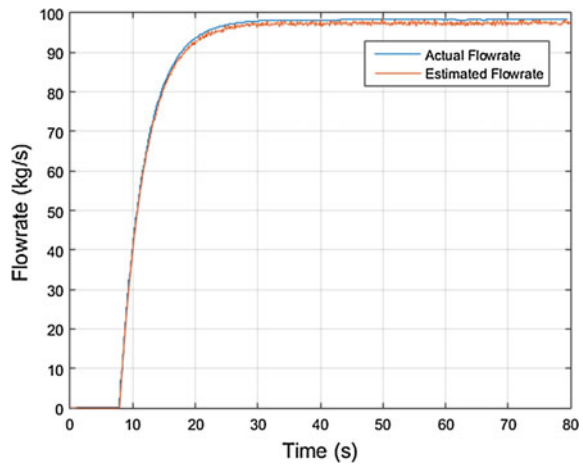
Fig. 5 Experimental setup

Fig. 6 Actual flowrate versus estimated flowrate

signal to get a precise estimate of the actual mass flow rate. The ZedBoard implementation was optimized for real-time processing, and the output flowrate was acquired at the required rate of 10 Hz. The results achieved have been demonstrated in graphs shown in Figs. 6 and 7.

It is evident from the results that Kalman filter can be employed for optimal filtering and precise mass flow estimation of dynamic belt conveyor systems. The proposed filtering technique can significantly enhance the accuracy of existing mass flow estimation systems without replacing existing tachometer sensors in old industrial setups. A linear time invariant approximation has been used for modeling the system, which can be replaced by a more complex, yet more accurate non-linear and time varying approximation. Adaptive filtering techniques may also be employed in order to further enhance the mass flow estimation accuracy.

Fig. 7 Actual mass transferred (tonnes) versus estimated mass transferred
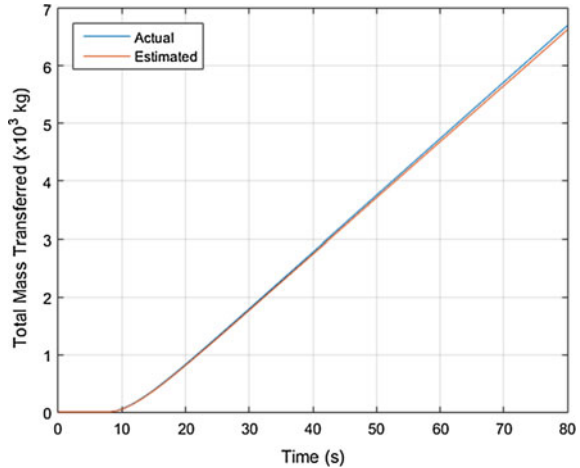
5 Conclusion

In conventional conveyor belt systems, the methods employed for mass flow estimation focus more on load-cell filtering and thus, are limited in accuracy due to inaccuracies in the tachometer measurement. In this study, we address this issue and propose a method for accurate estimation of conveyor belt velocity using the noisy tachometer signal. An accurate mathematical model of the system is derived and its utilization in a Kalman filter is demonstrated for accurate mass flow estimation of the belt velocity. The load-cell signal is filtered via a moving average filter, and combined with the accurate belt velocity estimate from the Kalman filter, precise mass flow estimation with less than 1 % error has been demonstrated.

References

Catlin, D.E. (1989). The discrete Kalman filter. In *Estimation, control, and the discrete Kalman filter* (pp. 133–163).

Chen, G. (1994). Robust stability analysis of Kalman filter under parametric and noise uncertainties. In *Approximate Kalman filtering* (pp. 179–189).

Halimic, M., & Balachandran, W. (1995). Kalman filter for dynamic weighing system in industrial electronics. In *Proceedings of the IEEE International Symposium on ISIE'95* (Vol. 2, pp. 10–14, 786–791).

Harb, A. M. (2013). *Enhancing the performance of dynamic weighing systems using Kalman filter*. http://library.iugaza.edu.ps/thesis/112429.pdf. Accessed October 15, 2015.

Jafaripana, M., et al. (2003). *Load cell response correction using analog adaptive techniques*. In *International Symposium on Circuits and Systems*.

Louis, F. L., et al. (2007). Continuous time Kalman filter. In *Optimal and robust estimation* (2nd ed., pp. 150–157).

Vayrynen, T. (2013, February 6). *Mass flow estimation in mineral processing applications* (Master of Science Thesis). Tampere University of Technology.

Part III
Automation and Control Engineering

Stiffness Analysis and Optimization for a Bio-inspired 3-DOF Hybrid Manipulator

Dan Zhang and Bin Wei

Abstract This paper mainly studies the maximum stiffness, minimum stiffness and global stiffness of a three-degree-of-freedom 4UPS-PU hybrid mechanism that can be used as the head section of a biologically inspired groundhog-like rescue robot. First of all, the kinematic and Jacobian matrix of the hybrid mechanism are derived; secondly, the maximum and minimum stiffness distribution trend of the 4UPS-PU manipulator are investigated and analyzed; and finally the global stiffness and its optimization are studied.

1 Introduction

Parallel manipulators typically consist of a moving platform that is connected to a fixed base by several limbs. The number of limbs is preferably but not necessarily equal to the number of degrees of freedom of the moving platform such that each limb is driven by one actuator and all actuators can be mounted on or near the fixed base. Parallel manipulators have been widely used during the past decade as compare to their counterpart serial manipulator because of the merits that parallel manipulators possess, such as high stiffness, low inertia, etc. Recently, biologically inspired manipulators have attracted many researchers' interests. Chinese Academy of Sciences has investigated the snake inspired serial manipulators, also many universities in the United States have manufactured the prototype of the creeper models and investigated other biologically inspired manipulators. Furthermore, stiffness analysis on manipulators has been the topic of numerous researchers over the past decade. For examples, Tsai and Joshi (2000, 2002) investigated the stiffness properties of 3-DOF position mechanisms for use in the hybrid kinematic

D. Zhang · B. Wei (✉)
University of Ontario Institute of Technology, Oshawa, ON, Canada
e-mail: Bin.Wei@uoit.ca

machines, Zhang (2000), Zhang and Gosselin (2001) have studied the stiffness and compliance of N-DOF parallel manipulators with being considered the lumped kinetostatic model accounting for joint and link compliance. Li and Xu (2008) have researched the stiffness of 3-PUU parallel robot, etc.

Stiffness is a measurement of the ability of a body to withstand the action of deformation because of action of external forces. The stiffness of a mechanism at a given point within the workspace can be evaluated by the stiffness matrix. The stiffness matrix relates the forces/torques that are applied at the end effector in Cartesian space to its Cartesian displacements. Stiffness can be a critical element in many scenarios due to the fact that it relates to accurate positioning and other dynamic performances.

Here, we first investigated the maximum and minimum stiffness of 4UPS-PU hybrid manipulator which has 3 degrees of freedom, i.e. translation along z axis and rotations about x axis and y axis, that can be used as the head section of the biologically inspired groundhog-like rescue robot (Zhang and Gao 2011); secondly, we investigated the global stiffness and its optimization in order to obtain the optimal stiffness in each direction.

2   Geometric Modeling of 4UPS-PU Mechanism

When the accident occurred at the coal mine, the rescue robot can obtain environmental detecting parameters and hazard information, and discover these survivors information and transferred it to the ground, and to help emergency rescue teams to have their searching and rescue planned, so as to minimize the loss of lives and property.

Aimed at detecting the scenario of unreachable accident spot due to tunnel block in a coalmine tragedy, and based on the bionics principle of groundhog and parallel robot mechanism, a groundhog-like detecting and rescuing robot system has been developed as shown in Fig. 1, which includes a robot drilling head which can rotate with solid stiffness, a multi-joint body which can wriggle forward in narrow space,

Fig. 1 Groundhog-like robot
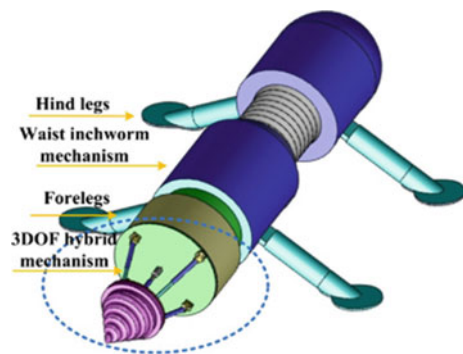
and digging legs which can push the coal backwards. This system is capable to crawl or drill through the collapsed objects to enter the accident spot, to detect the flammable gas density, and to send the measured signals back to the ground via wireless network. So far, based on our best knowledge, this kind of robot system has never been reported. This groundhog-like detecting and rescuing robot system can be used to detect gas density of accident environment and alert further dangers, as well as to communicate with the rescuing team via wireless network.

As shown in Fig. 2, the head section 4 UPS-PU mechanism has four same variable-length legs and one passive constraining leg that connect the fixed base to a moving platform. In this manipulator, each of the four identical legs consists of a universal joint, a moving link, an actuated prismatic joint, another moving link and a spherical joint attached to the platform. The middle chain that connects the base center to the platform center is a passive constraining leg, and it consists of a prismatic joint attached to the base, a moving link and a universal joint attached to the platform. The last leg is used to force the motion of the platform to three degrees of freedom. For the purpose of analysis, a fixed reference frame (XYZ) is attached to the base at point O, located at the center of the circumferential circle of the base. For each leg, another coordinate system $(uvw)_i$ is attached to the fixed base at B_i, such that u_i is in direction of b_i and at angle φ_i from X-axis as shown in Fig. 1. A moving frame (xyz) is attached to the moving platform at point O_e, located at the center of the circumferential circle of the moving platform. The angle between a_i and x-axis is defined φ'_i. Here we assume $\varphi'_i = \varphi_i$.

Fig. 2 Schematic representation of the 4UPS-PU mechanism
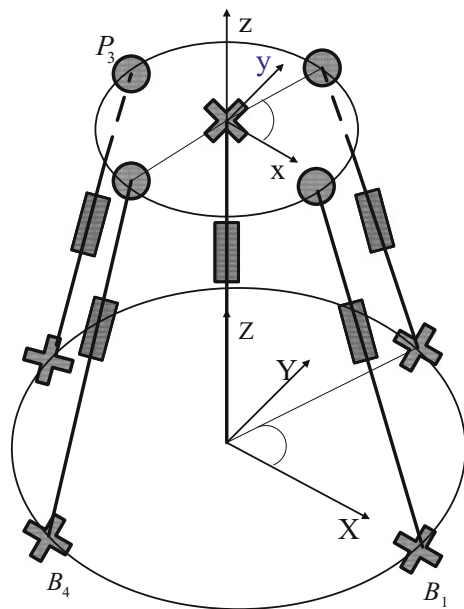
3   Inverse Kinematics

As shown in Fig. 2, all vectors are expressed in fixed reference frame as follows,

$$b_i = [\, r \cos \varphi_i \quad r \sin \varphi_i \quad 0\,]^T \tag{1}$$

$$a_i = R a_i' \tag{2}$$

Here,

$$a_i' = [\, c \cos \varphi_i' \quad c \sin \varphi_i' \quad 0\,]^T \tag{3}$$

which r and c are the radii of the base and moving platform. Moreover, R is the rotation matrix of the moving platform with respect to the base and can be written as (Cui et al. 2009),

$$R = \begin{bmatrix} \cos \psi & 0 & \sin \psi \\ \sin \varphi \sin \psi & \cos \varphi & -\sin \varphi \cos \psi \\ -\cos \varphi \sin \psi & \sin \varphi & \cos \varphi \cos \psi \end{bmatrix} \tag{4}$$

The position vector of point O_e with respect to origin of the fixed coordinate frame can be written as $[\, 0 \quad 0 \quad h\,]^T$.

The vector l_i can be written as,

$$l_i = p_e + a_i - b_i \tag{5}$$

The length of the ith leg is written as,

$$q_i^2 = l_i^T l_i \tag{6}$$

The Jacobian matrix can be determined by time differentiating equation (6) as following,

$$B\dot{q} = At \tag{7}$$

$$J = B^{-1}A \tag{8}$$

where A and B are defined for the manipulator as Zhang et al. (2009), Liu et al. (2008)

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix}, \quad B = \begin{bmatrix} q_1 & 0 & 0 & 0 \\ 0 & q_2 & 0 & 0 \\ 0 & 0 & q_3 & 0 \\ 0 & 0 & 0 & q_4 \end{bmatrix}$$

$$t = \begin{bmatrix} \dot{\varphi} \\ \dot{\psi} \\ \dot{h} \end{bmatrix}, \quad \dot{q} = \begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \\ \dot{q}_4 \end{bmatrix}$$

where the elements of A can be written as (Zhang et al. 2010; Wu et al. 2010)

$$
\begin{aligned}
a_{i1} &= -rc \sin \varphi_i \cos \varphi_i \cos \varphi \sin \psi + rc \sin \varphi \sin^2 \varphi_i \\
&\quad + hc \cos \varphi_i \sin \varphi \sin \psi + hc \sin \varphi_i \cos \varphi \\
a_{i2} &= rc \cos^2 \varphi_i \sin \psi - rc \sin \varphi_i \cos \varphi_i \sin \varphi \cos \psi \\
&\quad - hc \cos \varphi_i \cos \varphi \cos \psi \\
a_{i3} &= h - c \cos \varphi_i \sin \psi \cos \varphi + c \sin \varphi_i \sin \varphi \quad (i = 1, 2, 3, 4)
\end{aligned}
$$

4 Maximum and Minimum Stiffness

When a mechanism conducts a mission, the end-effector applies force to the environment. The reaction force will make the end-effector to be deflected away from its desired location. The deflection amount can be written as a function of applied force and the manipulator stiffness. So the stiffness of a manipulator has an impact on its accuracy of positioning. The stiffness of a mechanism is depending on several factors, includes the size and material of the links, mechanical transmission mechanisms, actuators, and control system, etc. In this paper, we assume the main source of the compliance come from the compliance of the actuators.

Let F denote end effector output force vector and Δx represent the end effector displacement vector. It can be shown that:

$$F = K \Delta x \tag{9}$$

where

$$K = J^T K_J J \tag{10}$$

is known as the stiffness matrix, and K_J is a 4×4 diagonal matrix in which each non-zero diagonal element, k_i, represents the stiffness constant of the ith joint actuator. Furthermore, if $k_1 = k_2 = k_3 = k_4 = k$, the above equation reduces to

$$K = k J^T J \tag{11}$$

We observe from Eq. (10) that K is a 3×3, symmetric and position dependent matrix. In addition, the deflection is relying on output force direction, in other words, the stiffness of a manipulator is direction dependent. Let λ_i be an eigenvalue

and f_i be the corresponding eigenvector of the stiffness matrix K at a designated platform location. Then λ_i is the stiffness in the corresponding eigenvector direction. It follows that if λ_{min} is the minimum eigenvalue and λ_{max} is the maximum eigenvalue, then the minimum stiffness takes place in the f_{min} direction and maximum stiffness takes place in the f_{max} direction.

The stiffness constant is taken to be 1000 N/m for a liner actuator. The maximum and minimum stiffness values at each point within the workspace of the 4UPS-PU manipulator have been computed. The maximum stiffness mapping at the elevation of z = 0.17 m as a case for the manipulator is shown in Fig. 3. The minimum stiffness mapping at the elevation of z = 0.17 as a case for the manipulator is shown in Fig. 4. Figure 3 shows that the highest stiffness value occurs when ψ and φ are equal to zero. Figure 4 shows that the lowest stiffness value occurs near the manipulator workspace boundary, this is because a manipulator becomes singular when it approaches the workspace boundary.

Fig. 3 Maximum stiffness of the 4UPS-PU manipulator
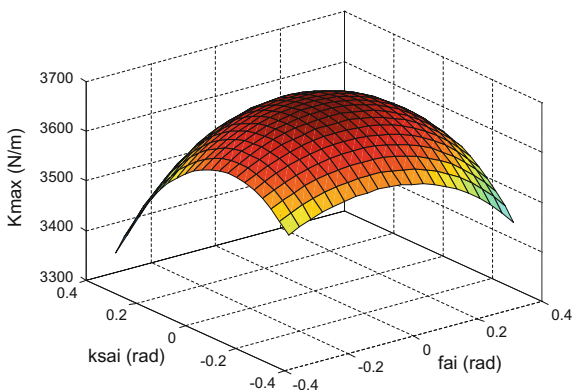
Fig. 4 Minimum stiffness of the 4UPS-PU manipulator
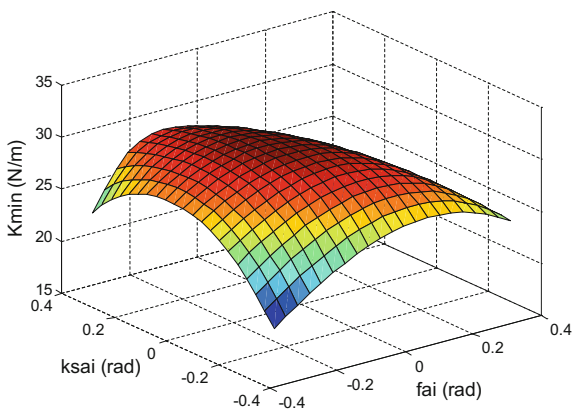
5 Stiffness Optimization of 4UPS-PU Manipulator

5.1 Global Stiffness of the Manipulator

As stated in Eq. (11), the stiffness of the manipulator is expressed by a 3×3 matrix. The diagonal elements of the matrix express the pure stiffness of the mechanism in each direction. In order to obtain the optimal stiffness in each direction, one can write an objective function (12), with stiffness element to maximize. The global stiffness K_{global} is defined as following,

$$K_{global} = \eta_1 k_{11} + \eta_2 k_{22} + \eta_3 k_{33} \tag{12}$$

where k_{ii} (i = 1, 2, 3) is the diagonal element of the stiffness matrix of the mechanism. η_i is the weight factor for each directional stiffness, which evaluates the priority of the stiffness in this direction. This would maximize the sum of the diagonal element which is global stiffness. Even though we couldn't get each diagonal element individually maximized, each stiffness can be optimized by distributing the weighting factors.

For the purpose of numerical analysis first, we assume r = 0.2 m, $\varphi = 0$, $\psi = 0$, $\eta_i = 1$, the design variables are c and h, their bound are $c \in [0.1, 0.5]$m and $h \in [0.1, 0.5]$m according to the practical requirements. Then we can obtain K_{global} stiffness distribution as shown in Fig. 5. From Fig. 5, one can see that the maximum K_{global} is around 4100 N/m. Similarly when we assume r = 0.2 m, $\varphi = 20°$, $\psi = 25°$, $\eta_i = 1$, the global stiffness K_{global} distribution is shown in Fig. 6. From the figure one can see that the maximum K_{global} is around 4000 N/m. When the design variables are φ and ψ, and assuming different values of r, c and h, one can also have the K_{global} distribution. Through different experiments we found that r, φ, ψ, c, h dramatically affect the K_{global}. Next, we need to determine the values of the c, h, r, φ and ψ simultaneously that lead to the maximum K_{global}.

Fig. 5 Kglobal distribution when r = 0.2 m, $\varphi = 0°, \psi = 0°$

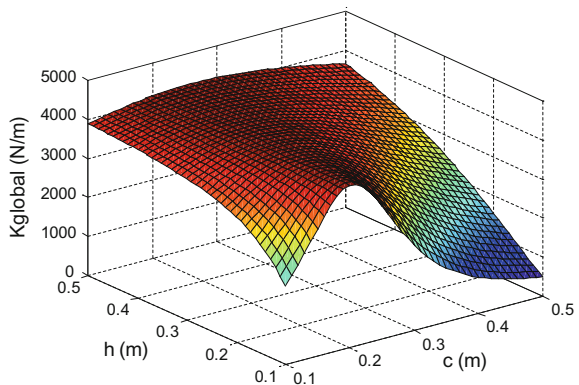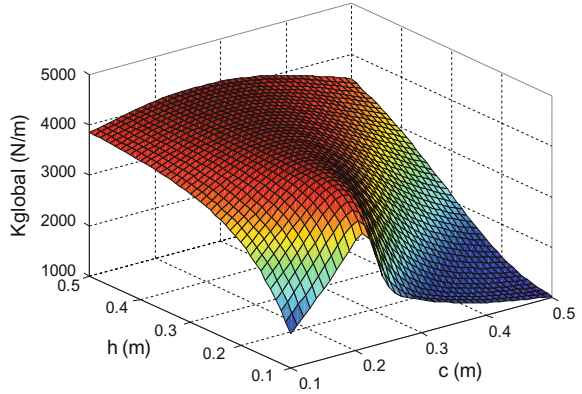**Fig. 6** Kglobal distribution
when r = 0.2 m,
$\varphi = 20°, \psi = 25°$

5.2 Optimization

Genetic algorithm (GA) is based on the natural selection and it repeatedly modified a population of individual solutions. At each step, the GA selects individuals at random from the current population to be parents and use them produce the children for the next generations (Hu 2008). Genetic algorithms have the advantages of good convergence and robustness properties (Zhang 2009).

The genetic algorithm is employed here to optimize the global stiffness K_{global} of the bio-inspired hybrid mechanism. Our objective function is: $K_{global} = \eta_1 k_{11} + \eta_2 k_{22} + \eta_3 k_{33}$, our purpose is to maximum the K_{global}. Note that the optimization functions in the GA minimize the objective function, in order to maximize the objective function, we need to minimize $-K_{global}$, because the point at which the minimum of $-K_{global}$ occurs is the same as the point at which the maximum of K_{global} occurs. The design variables are c, h, r, φ and ψ. Their bound are the following according to practical requirements,

$$c \in [0.1, 0.5]\text{m}, \quad h \in [0.1, 0.5]\text{m},$$

$$r \in [0.1, 0.5]\text{m}, \quad \varphi \in [-30°, 30°], \quad \psi \in [-30°, 30°]$$

Some genetic parameters and operators are set as:

Scaling function: Proportional;
Selection function: Roulette;
Crossover function: intermediate;
Crossover ratio: 1.0;
Mutation function: adaptive feasible;
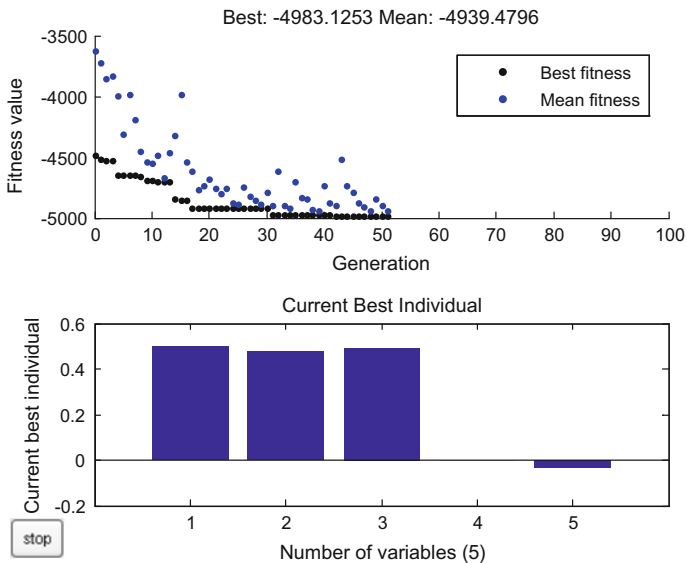Population size: 20;
Maximum number of generations: 100.

Fig. 7 The best fitness value and the best individuals of the global stiffness optimization

Run the optimization and we can have the following results:

The results of the global stiffness optimization are shown in Fig. 7. The points at the bottom of the plot represent the best fitness values and the points above them are the mean fitness values in each generation. The plot also displays the vector entries of the individual with the best fitness function value in each generation. The optimal parameters are obtained after 51 generations as follows,

$$[c, h, r, \varphi, \psi] = [0.496 \, \text{m}, 0.476 \, \text{m}, 0.491 \, \text{m}, 0.002 \, \text{rad}, -0.037 \, \text{rad}]$$

And the maximum K_{global} is 4983.1253 N/m

The results suggest that in order to make the global stiffness of the manipulator reach the maximum, the radii of the moving platform should be 0.496 m, the heave of the manipulator should be 0.476 m, the radii of the base should be 0.491 m, $\psi = 0.002 \, \text{rad}$ and $\psi = -0.037 \, \text{rad}$.

6 Conclusion

This paper mainly focuses on the maximum and minimum stiffness distribution trend of 4UPS-PU mechanism as well as the global stiffness and its optimization of the mechanism. First of all, the maximum and minimum stiffness distribution trend of the 4UPS-PU mechanism that can be used as the head section of the rescue robot which can be served in many dangerous ragged places are analyzed; secondly, in

order to obtain the optimal stiffness in each direction, we investigated the global stiffness K_{global} and optimized it through GA as to obtain the optimal stiffness in each direction.

References

Cui, G. H., Zhang, Y. W., Zhang, Y. H., et al. (2009). Configuration design and analysis of a new 3-SPS/S spatial rotation parallel manipulator. *Journal of Jilin University: Engineering and Technology Edition*, *39*(Sup.1), 200–205.

Hu, X. L. (2008). *Design and analysis of a three degrees of freedom parallel kinematic machine* (Dissertation for Master Degree). Canada: University of Ontario Institute of Technology.

Li, Y. M., & Xu, Q. S. (2008). Stiffness analysis for a 3-PUU parallel kinematic machine. In *Mechanism and machine theory* (pp. 186–200).

Liu, Y. B., Zhao, J., & Cai, H. G. (2008). Study on kinematics and dynamics of a novel 6-PRRS parallel robot. *Journal of Jilin University* (*Engineering and Technology Edition*), *38*, 1220–1224.

Tsai, L. W., & Joshi, S. (2000). Kinematics and optimization of a spatial 3-UPU parallel manipulator. *Journal of Mechanical Design, 122*, 439–446.

Tsai, L. W., & Joshi, S. (2002). Kinematic analysis of 3-DOF position mechanisms for use in hybrid kinematic machines. *Journal of Mechanical Design, 124*, 245–253.

Wu, X., Chen, M. L., & Zhang, Y. B. (2010). Analysis of position and singularity of a parallel mechanism with three degrees of freedom. *Transactions of the Chinese Society for Agricultural Machinery, 41*(2), 208–213.

Zhang, D. (2000). *Kinetostatic analysis and optimization of parallel and hybrid architecture for machine tools* (Ph.D. thesis). Canada: Laval University.

Zhang, D. (2009). *Parallel robotic machine tools*. Heidelberg: Springer.

Zhang, D., Bi, Z. M., & Li, B. Z. (2009). Design and kinetostatic analysis of a new parallel manipulator. *Robotics and Computer-Integrated Manufacturing, 25*, 782–791.

Zhang, D., & Gao, Z. (2011). Hybrid head mechanism of the groundhog-like mine rescue robot. *Robotics and Computer-Integrated Manufacturing, 27*, 460–470.

Zhang, D., & Gosselin, C. M. (2001). Kinetostatic modeling of N-DOF parallel mechanisms with a passive constraining leg and prismatic actuators. *Journal of Mechanical Design, 123*, 375–381.

Zhang, Y. W., Cui, G. H., Zhun, Z. S., et al. (2010). Singular loci analysis of a new 3-SPS-1-S spatial rotation parallel manipulator. *Transactions of the Chinese Society for Agricultural Machinery, 41*, 199–203.

Robust Gust Rejection on a Micro-air Vehicle Using Bio-inspired Sensing

William A. Dean, Badri N. Ranganathan, Ivan Penskiy, Sarah Bergbreiter and J. Sean Humbert

Abstract Growing demand for robust, low-computation sensing and control of micro-air vehicles motivates development of new technology. A MEMS wind flow sensor was previously developed at the University of Maryland, drawing inspiration from setae structures seen in biology. In this work this sensor is integrated onto a quadrotor platform and utilized for sensing a gust perturbation. Suitable signal processing methods were used to isolate the gust-related perturbation from other ambient fluctuations affecting the sensor. The problem of gust rejection is then formulated using standard tools from robust control theory and a controller is obtained using μ–synthesis. Theoretical analysis of the μ controller's performance improvement is carried out with simulations. A robustly stable controller was implemented on a quadrotor micro-air vehicle to improve lateral state regulation in the presence of a lateral gust stream. Flight testing revealed attenuation of the lateral velocity and perturbation from the projected path.

Keywords Micro-air vehicle · Bio-inspired sensing · Gust speed estimation · Robust control

1 Introduction

Micro-air vehicles are an increasingly popular platform for research, commercial, and military applications. Limitation in size, weight, and power for this type of technology makes autonomous flight difficult. This is even more of a problem when coupled with the increased agility seen in small-scale aircraft (Ol et al. 2008). As a result, gust perturbations to these vehicles can result in crashes. Thus, development of robust, light-weight, low-power sensing and control capabilities is vital.

W.A. Dean (✉) · B.N. Ranganathan · I. Penskiy · S. Bergbreiter · J.S. Humbert
Department of Aerospace Engineering, University of Maryland, College Park, MD, USA
e-mail: wdean@terpmail.umd.edu

Biology serves as a rich source of inspiration for engineering new methods of sensing and control. Flying insects are useful for this discussion as they demonstrate robustly stable flight while simultaneously addressing similar constraints on payload and computing power. A wind sensing mechanism is used by insects for navigation, and it is suggested that wind sensitive setae may be involved (Reynolds et al. 2010). It is found that flow sensing can also provide stability augmentation on an MAV (Keshavan and Humbert 2010). Disturbances, such as (primarily) wind gusts, are known to have extremely adverse effects on the performance of MAVs. Much effort has gone into developing control strategies to subjugate these effects using pre-defined gust models (Aouf et al. 2000), but the turbulent and unpredictable nature of these disturbances makes it difficult for vehicles to react in a timely fashion (Zarovy et al. 2010). Simulation efforts provide promising results, but often require numerous simplifications (Waslander and Wang 2009). Flow field estimation has also been used for disturbance rejection, which has incorporated effects such as blade flapping, blade drag, and non-conventional aircraft platforms (Yeo et al. 2015; Sydney et al. 2013). Unfortunately, these models may not stay accurate for all conditions an MAV might be subject to, and may be computationally intensive. Given a way to directly measure the gust disturbance, a vehicle's performance could be vastly improved. Previous efforts to develop similar sensors are discussed in Chen et al. (2005), Krijnen et al. (2006), and Engel et al. (2005). While prior work have developed sensors, no publications have been found that demonstrates the deployment of these sensors on any air vehicle.

A Kapton-based 'hair' sensor has been developed, at the University of Maryland, to measure wind flow (Ranganathan et al. 2015). This device was conceived from the setae structures seen in nature which are commonly found covering the bodies of flying insects. This work discusses the integration of these sensors to estimate wind gust speed on a quadrotor vehicle and use this signal as feedback to improve performance in the presence of a lateral wind gust disturbance. The hair sensor used is a robust, lightweight, free standing, directional sensor system suitable for MAVs. This work will demonstrate first-time use of this type of sensing as a means to improve gust rejection capabilities.

The main contribution of this paper is the integration of a hair-like gust sensor onto a quadrotor vehicle and development of appropriate signal processing methods and controller design framework for successful demonstration of gust rejection. The paper is organized as follows: Sect. 2 describes the hair sensor and associate signal conditioning that is used for gust sensing. Section 3 presents the estimation method used for gust speed determination. Section 4 describes the synthesis and analysis of a robust gust rejection controller. Section 5 describes the experimental flight results based on the chosen sensing and control schemes. Section 6 presents the conclusions from this work.

2 Bio-inspired Hair Sensor

The design and fabrication of the hair sensor used in this work was discussed in (Ranganathan et al. 2015), and are briefly summarized here for reference. Strain gauge based cantilever-paddle structures were made using gold as strain gauge material and Kapton as structural material. This sensor is robust to handling and provides a bi-directional response to air flow. The sensor is shown on the right in Fig. 1.

The sensing strain gauge patterns are located on the roots of the cantilevers, and connected in series. The output of the sensor varies with the strain ε induced at the gauge which is related to incident fluid pressure P as shown in $\varepsilon = \frac{6PAL}{Ewt^2}$, where E is Young's Modulus of Kapton, t is the thickness, w and L are the width and length of the cantilevers, A is effectively the area of the paddle exposed to flow. The strain gauge is connected in a Wheatstone network and the signal is amplified in two-stages using instrumental amplifier chips (INA-326) followed by low-pass filtering to reduce the amplified noise. Figure 2 shows the circuit diagram for the signal conditioning board.

R_S is the sensing resistance, R_R is the reference resistance, and R_{D1}, R_{D2}, R_5, and R_6 are potentiometers tuned to set appropriate reference voltages. R_f and C_f form the low-pass filter where the cut-off frequency is set to 10 Hz. A total gain of 2000 was applied over the two stages (first stage—50, second stage—40) which provides a suitable signal magnitude without causing saturation.

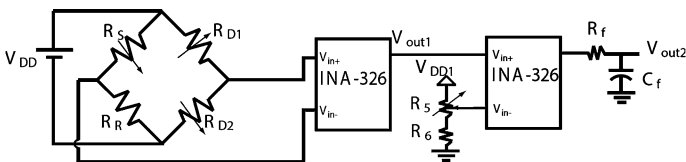**Fig. 1** Setae (*left*) and hair sensor (*right*) compared to Quarter (*center*)

**Fig. 2** Signal conditioning circuit (simplified)

3 Gust Speed Estimation

The output of the signal conditioning circuit has a fixed DC offset and air flow causes perturbation to the signal around this offset. This signal is filtered with a 2nd order Butterworth band-pass filter (frequency range: 0.87–8.47 rad/s). The high-pass part of this filter removes the non-zero offset and any other drift-like variations and the low-pass part removes the high frequency content inherent to the generated gusts.

Upon removal of the gust, an overshooting effect is seen in the output from this filter, which is compensated for by using half-wave rectification. This rectification could be applied for this work because the experiments used only unidirectional wind gusts. A threshold level of 0.1 V was used to set any output from the filter below it to zero.

The sensor was exposed to the generated gust at distances of 0.5, 0.75, 1.0, 1.25, and 1.5 m away from the nozzle, where the gust speeds were measured (independently using a handheld anemometer) as 9.61, 7.60, 5.81, 4.92, and 3.58 m/s, respectively. A static linear scheme was used for estimation of gust speed based on this data which was then used for feedback. This estimation process does not approximate with a high level of accuracy given the limited window of operation of the sensor and unsophisticated estimation scheme. However, it does offer a conservative estimate that will not drift from zero unless a significant gust affects the vehicle (and sensor). Figure 3 shows the estimation of the gust speeds mentioned earlier.

For our experiments, the expected gust magnitude stimulus is around 9 m/s to be able to perturb the quadrotor in a noticeable fashion.

Fig. 3 Gust speed estimation
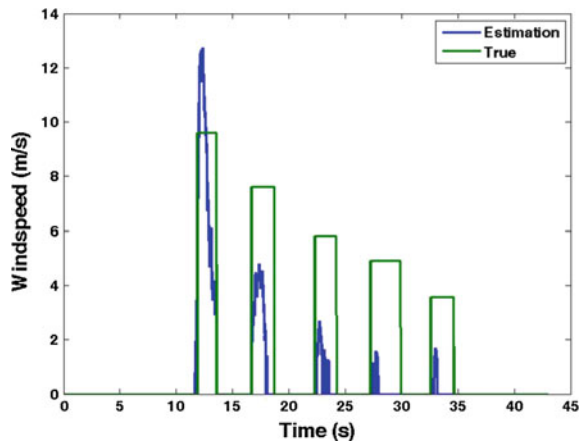
4 Robust Controller Design

The hair sensor was used to estimate the gusts over only a small portion of the vehicle while the entire vehicle passed the gust disturbance in flight testing. In an attempt to compensate for the unmeasured disturbance affecting the vehicle and the possibly inaccurate gust speed estimation, a robustly stable μ-controller was designed using gust speed as feedback. A robustly stable controller provides (or maintains) system stability for a range of plant uncertainty. A robust controller was used in this work as a means to provide a framework that incorporates system uncertainty in addition to improving gust rejection performance. The methodology behind the controller synthesis and performance analysis was referenced from Skogestad and Postlethwaite (2007). Figure 4 shows the block diagram used for this controller design.

This system is modeled with additive uncertainty, input disturbance, and performance measured at the output. Additive uncertainty and input disturbance were chosen from a variety of other possible combinations to allow the highest possible bandwidth in the performance weighting function. The state labels and block names are listed in Table 1.

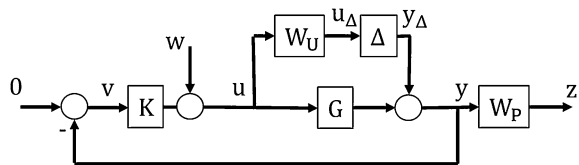Fig. 4 System model used for disturbance rejection controller design

Table 1 Block diagram labels and physical meanings

Name	General meaning	Physical value
v	Signal available to controller	Negative of the gust estimate
u	Control inputs	Lateral control input
u_Δ	Input to the uncertainty block	Weighted control input
y_Δ	Output from uncertainty block	Uncertainty added to the output
y	Output from the plant	Lateral states
w	Exogenous input	Lateral gust
z	Regulated output	Weighted states
K	Controller	Gust rejection controller
G	Plant	Augmented lateral dynamics
W_U	Uncertainty weighting function	
Δ	Uncertainty block	
W_P	Performance weighting function	

4.1 Controller Synthesis

The plant G incorporates lateral dynamics for the states v (lateral velocity), p (roll rate), Φ (roll angle), and V_g (gust velocity). The state-space equations of motion are written in Eq. 1.

$$\begin{bmatrix} \dot{x} \\ \dot{x}_d \end{bmatrix} = \begin{bmatrix} A_L & G_d C_d \\ 0 & A_d \end{bmatrix} \begin{bmatrix} x \\ x_d \end{bmatrix} + B_L u + B_d d, \quad y = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ x_d \end{bmatrix} \qquad (1)$$

where $\mathbf{x} = \begin{bmatrix} v \\ p \\ \Phi \end{bmatrix}$. $y = x_d = v_g, u = u_{Roll}$.

The lateral dynamics state space model (A_L, B_L) is augmented with a first-order Gauss-Markov model for the gust dynamics (Gadewadikar et al. 2009). The shaping filter for the gust is then represented in Eq. 2.

$$\dot{x}_d = \frac{-1}{\tau} x_d + \rho d \qquad (2)$$

Here $\tau = 3.2$ s is the correlation time of the wind, $\rho = 1$ is a weighting factor, and d (also seen in Eq. 1) is a white noise process that drives the gust dynamics. Thus, $(A_d, B_d, C_d, D_d) = (-0.3125, 1, 1, 0)$. The vehicle's lateral dynamics were populated with first-order stability derivatives, which were calculated using system identification techniques (Klein and Morelli 2006). It should be noted that these dynamics were identified with the stabilizing avionics active and with additional hardware (see Sect. 5) added to the vehicle.

The disturbance's effect on the dynamics was defined by $\mathbf{G_d}$. This matrix (3×1) was set equal to the negative of the first column of $\mathbf{A_L}$, which can be interpreted as a gust moving in the positive y-axis (of the vehicle's body frame) and is equivalent to the vehicle moving in the negative y-direction (Nelson 1998). The LTI system that was used for controller design is given in Eq. 3.

$$\begin{bmatrix} \dot{x} \\ \dot{x}_d \end{bmatrix} = \begin{bmatrix} -0.3380 & -0.2812 & 9.81 & 0.3380 \\ 1.0824 & -0.9652 & -33.584 & -1.0824 \\ 0.0175 & 1 & -1.041 & -0.0175 \\ 0 & 0 & 0 & -0.3125 \end{bmatrix} \begin{bmatrix} x \\ x_d \end{bmatrix} + \begin{bmatrix} -0.0644 \\ 11.3946 \\ 0.4348 \\ 0 \end{bmatrix} u + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} d,$$
$$y = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ x_d \end{bmatrix} \qquad (3)$$

The above system (in Eq. 3) will represent the LTI matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ for the rest of this paper. At this point it might be noticed that the plant $G(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} = 0$, since the only observed state is not dependent on the vehicle's states and is not affected by control inputs. Thus, an observer must be constructed so that full state feedback can be assumed for controller synthesis. A Luenberger

observer (Barmish and Galimidi 1986) was used for this purpose. The state estimate dynamics are given in Eq. 4.

$$\dot{\hat{x}} = \mathbf{A}\hat{x} + \mathbf{B}u + \mathbf{B_d}d + \mathbf{L}(y - \hat{y}), \ \hat{y} = \mathbf{C}\hat{x}. \tag{4}$$

L is a static gain matrix which defines the observer dynamics and is a (4×1) matrix of ones. A qualitative interpretation of this observer is that it acts as an estimator of the state perturbations due only to effects from the disturbance.

The performance weight $\mathbf{W_P}$ can be set to enforce a specific level of nominal performance with the trade-off of greater control inputs and pushing the limits of system stability. $\mathbf{W_U}$ can be thought of as a weight which normalizes the uncertainty perturbation from the nominal plant $(\mathbf{G_N})$ to be less than 1, where the identified plant is $\mathbf{G} = \mathbf{G_N} + \mathbf{W_U}\boldsymbol{\Delta}$ (for additive uncertainty). Robust stability is guaranteed if $\|\mathbf{M}\|_\infty < 1$ where for this problem, $\mathbf{M} = \mathbf{W_U}\mathbf{K}(\mathbf{I} + \mathbf{GK})^{-1} = \mathbf{W_U}\mathbf{KS}$. **S** is the (output) sensitivity transfer function and **I** is the identity matrix. We chose $\mathbf{W_U} = 0.3\frac{0.5S+1}{0.03125S+1}$ and $\mathbf{W_P} = \frac{0.2S+5}{S+1.5}\mathbf{I_{4x4}}$ to best represent our problem and performance requirements. We design a controller by using the technique of μ—synthesis, which minimizes the structured singular value for the system and thereby allows to achieve desired performance in the presence of a range of uncertainty. For the system under consideration, we obtain the linear fractional transformation matrix $\mathbf{N} = \begin{bmatrix} -\mathbf{W_U}\mathbf{KS} & -\mathbf{W_U}\mathbf{S_I} \\ \mathbf{W_P}\mathbf{S} & \mathbf{W_P}\mathbf{SG} \end{bmatrix}$. The controller, **K**, that minimizes μ(**N**) is found using DK-iteration. The full description of this methodology is outside the scope of this paper and can be referenced in Skogestad and Postlethwaite (2007). Combining the synthesized controller with the simple observer yields the following controller, which was implemented in flight testing.

$$
\begin{aligned}
K_{final} &= \mathrm{K}(s\mathrm{I} - (\mathrm{A} - \mathrm{BK} - \mathrm{LC}))^{-1}\mathrm{L} \\
&= \frac{-2.311s^7 + 253.4s^6 + 243.4s^5 + 3846s^4 + 8029s^3 + 25180s^2 + 15130s + 2984}{\substack{s^9 + 30.34s^8 + 456.1s^7 + 4357s^6 + 27560s^5 + 113500s^4 \\ + 293700s^3 + 436900s^2 + 278600s + 27960}}
\end{aligned}
\tag{5}
$$

4.2 Theoretical Analysis

The first test to perform is the robust stability check $\left(\|\mathbf{M}\|_\infty < 1\right)$. Note that the upper left component of **N** is **M** and that it is equivalent to compute to the structured singular value of that portion of **N** with an unstructured perturbation. This result is shown in Fig. 5 and reveals that the system is robustly stable.

We analyze the effects of the controller on the system to see how much improvement it offers. The performance weight was designed to reduce **SG**, which represents the effect of the input disturbance at the output. Without feedback, **G** relates the input disturbance to the output. These are compared in Fig. 6.

Fig. 5 Maximum singular value of **M**
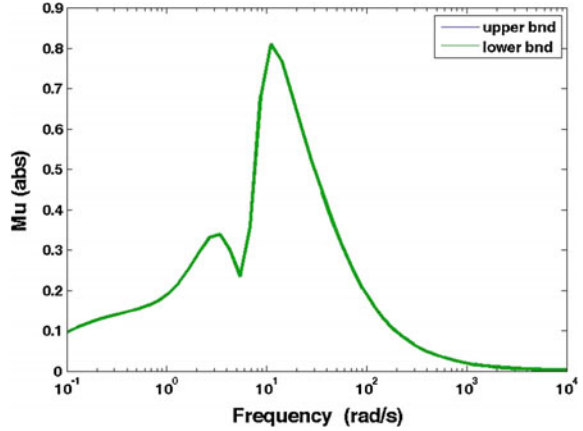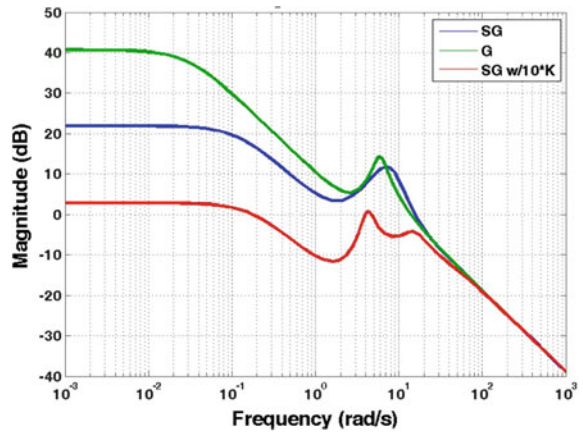

Fig. 6 Effect of input disturbance on output

There is clear improvement for frequencies up to 6.9 rad/s, where **SG** is slightly larger than **G**, after which the singular values fall off together. It was also seen that the controller can be multiplied by a static gain to improve performance further. A gain of 10 is chosen to show how **SG** is reduced by such a change. In the hardware implementation, the controller gain will be increased until expected sensor deflection results in a control input of about 66 % of the maximum possible control input.

Before implementing the controller in hardware, it was used in simulation to show that this type of controller works. A step function was input to the gust shaping filter (d in Eqs. 1, 2, and 3) for one second. The state of interest is the lateral velocity of the vehicle. Assuming a constant speed forward motion is the desired flight mode, reduction of the lateral velocity will reduce the perturbation from the undisturbed path. Roll rate and roll angle are not used for this analysis

Fig. 7 Simulated lateral
velocity comparison
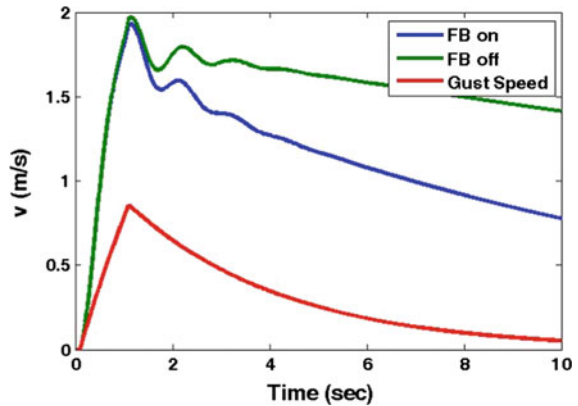
because they are heavily affected by the control input. Figure 7 displays the effects
of the gust on the vehicle with the gust sensing feedback on versus off.

It is seen that the effect on the lateral velocity is improved, but doesn't do so until
after the gust has already started declining in amplitude. Controllers which can
improve this response further are currently being worked upon.

5   Experimental Validation

The experimental procedure required a quadrotor vehicle to hover at a fixed
position with its heading pointed along a path perpendicular to the gust stream. The
vehicle was then given a forward motion impulse (from a human pilot), then given
no human input so that it could move in a straight line until being perturbed by the
gust. A leaf blower was used to generate a discrete gust stream. The vehicle was
pushed laterally by the gust, giving it a non-zero lateral velocity. A DJI Flame
Wheel 330 was controlled by an ArduPilot Mega avionics package, which also
implemented the filtering scheme mentioned previously. A Vicon Tracker system
was used to record true rigid body motion. Figure 8 shows the vehicle with all
associated hardware mounted on it.

Ten passes were used for comparing feedback ON vs feedback OFF perfor-
mance. These passes were chosen such that the mean path for both cases were as
close to each other as possible until the vehicle displayed signs that they were
affected by the gust. This guaranteed approximately equal distances from the gust
source, and thus equal levels of applied disturbance. The lateral velocity compar-
ison is shown in Fig. 9. In addition, the Cartesian trajectory comparison is shown in
Fig. 10.

Reduction in the lateral velocity is seen in Fig. 9. The root-mean-square of the
lateral velocity is reduced by 19.7 %. The Cartesian trajectory is shown as another
means to display improvement in disturbance rejection performance. Figure 10

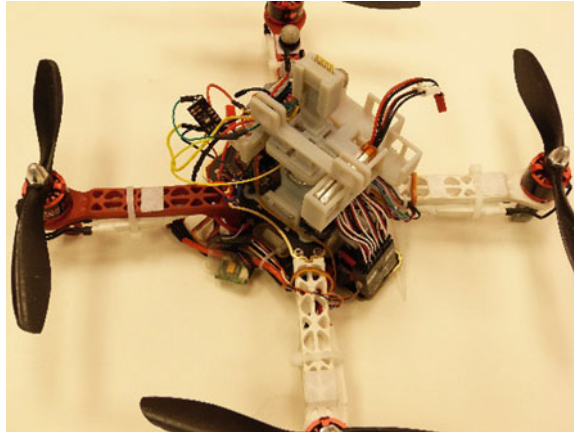**Fig. 8** Quadcopter with hair
sensor

Fig. 9 Experimental lateral
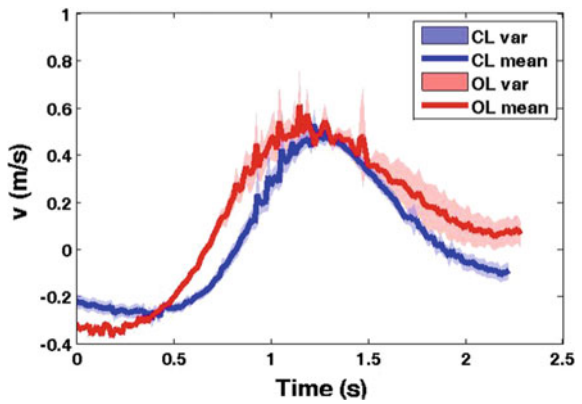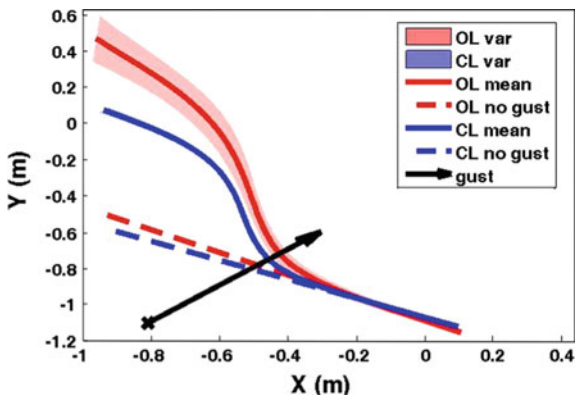velocity comparison

Fig. 10 Experimental
Cartesian trajectory
comparison

reveals a 17.1 % reduction in mean Cartesian displacement from the projected undisturbed trajectory, when using the hair sensor feedback. Note that this value would increase if the vehicle had more room to continue on its path, but pilot control commands took control to prevent the vehicle from crashing.

These numbers have potential to improve, but are still indicative of the applicability of these sensors. Thus, the goal of first-time implementation of these unique sensors for gust rejection is achieved.

6   Summary and Conclusion

Bio-inspired hair sensors were integrated for the first time on a micro-air vehicle and wind gust disturbance rejection was shown to be feasible. The hair sensor outputs were able to estimate gust speed by using a digital band-pass filter and half-wave rectifier. Linear static estimation was used for the gust speed estimation and found to perform adequately.

A gust rejection control system was developed for the quadrotor platform using techniques from robust control. The controller was designed to regulate lateral vehicle states in the presence of an input disturbance, while utilizing lateral gust speed feedback. Analysis of this scheme indicated that this controller would reduce the effect of the disturbance on the vehicle, both theoretically and in simulation. Experimental flight testing served as the last piece of validation and revealed a 19.7 % reduction in lateral velocity and a 17.1 % reduction in mean Cartesian displacement. The results serve as evidence of the ability to implement these unique sensors for gust rejection, while gust rejection performance could be improved in future work. With an improved sensor scheme and control design, the performance can be further enhanced.

References

Aout, N., Boulet, B., & Botez, R. (2000). H-2 and H-inf optimal gust load alleviation for a flexible aircraft. In *American Control Conference, 2000. Proceedings of the 2000*, 3, pp. 1872–1876. IEEE.

Barmish, B. R., & Galimidi, A. R. (1986). Robustness of luenberger observers: Linear systems stabilized via non-linear control. *Automatica, 22*(4), 413–423.

Chen, J., Engel, J., Chen, N., & Liu, C. (2005). A monolithic integrated array of out-of-plane hot-wire flow sensors and demonstration of boundary-layer flow imaging. In *18th IEEE International Conference on Micro Electro Mechanical Systems, Miami, Florida, USA*, pp. 299–302.

Engel, J., Chen, J., Chen, N., Pandya, S., & Liu, C. (2005). Development and characterization of an artificial hair cell based on polyurethane elastomer and force sensitive resistors. In *Sensors, 2005 IEEE*, p. 4. IEEE.

Gadewadikar, J., Lewis, F. L., Subbarao, K., Peng, K., & Chen, B. M. (2009). H-infinity static output-feedback control for rotorcraft. *Journal of Intelligent and Robotic Systems, 54*(4), 629–646.

Keshavan, J., & Humbert, J. S. (2010). Mav stability augmentation using weighted outputs from distributed hair sensor arrays. In *American Control Conference (ACC), 2010*, pp. 4445–4450. IEEE.

Klein, V., & Morelli, E. A. (2006). *Aircraft system identification: Theory and practice*. Reston, VA, USA: American Institute of Aeronautics and Astronautics.

Krijnen, G. J., Dijkstra, M., van Baar, J. J., Shankar, S. S., Kuipers, W. J., de Boer, R. J., et al. (2006). Mems based hair flow-sensors as model systems for acoustic perception studies. *Nanotechnology, 17*(4), S84.

Nelson, R. C. (1998). *Flight stability and automatic control* (Vol. 2). WCB/McGraw Hill.

Ol, M., Parker, G., Abate, G., & Evers, J. (2008). Flight controls and performance challenges for mavs in complex environments. *AIAA Paper*, p. 6508.

Ranganathan, B. N., Penskiy, I., Dean, W., Bergbreiter, S., & Humbert, S. (2015). Bio-inspired wind frame state sensing and estimation for mav applications. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Hamburg, Germany, September 2015.

Reynolds, A. M., Reynolds, D. R., Smith, A. D., & Chapman, J. W. (2010). A single wind-mediated mechanism explains high-altitude non-goal oriented-headings and layering of nocturnally migrating insects. *Proceedings of the Royal Society B: Biological Sciences, 277* (1682), 765–772.

Skogestad, S., & Postlethwaite, I. (2007). *Multivariable feedback control: Analysis and design* (Vol. 2). New York: Wiley.

Summary of properties for kapton polyimide. (2015). http://www.dupont.com/content/dam/ assests/products-and-services/membranes-films/assets/DEC-Kapton-summary-of-properties. pdf

Sydney, N., Smyth, B., Paley, D., et al. (2013). Dynamic control of autonomous quadrotor flight in an estimated wind field. In *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*, pp. 3609–3616. IEEE.

Waslander, S. L., & Wang, C. (2009). Wind disturbance estimation and rejection for quadrotor position control. In AIAA Infotech@ Aerospace Conferenceand AIAA *Unmanned, Unlimited Conference*, Seattle, WA.

Yeo, D., Shrestha, E., Paley, D. A., & Atkins, E. (2015). An empirical modelof rotorcraft uav downwash model for disturbance localization and avoidance.

Zarovy, S., Costello, M., Mehta, A., Gremillion, G., Miller, D., Ranganathan, B., et al. (2010). Experimental study of gust effectson micro air vehicles. In *AIAA Atmospheric Flight Mechanics Conference*.

Development of Guidance, Navigation and Control System Using FPGA Technology for an UAV Tricopter

Arturo Cadena, Ronald Ponguillo and Daniel Ochoa

Abstract This paper presents the development of a Guidance, Navigation and Control (GNC) system for a rotorcraft unmanned aerial vehicle (UAV). The airframe and its associated mathematical model are explained. The vehicle electronics system is based on a FPGA development board with its peripherals. The GNC routines were coded using hardware description language VHDL with 32 bits floating point arithmetic and CORDIC algorithms. An Inertial Navigation System (INS) complemented by GPS, implemented by an Extended Kalman Filter is included. A hybrid approach using PID and Fuzzy controllers is proposed for GNC system design. System identification and calibration was done using a test stand that includes a gimbal mechanism where the vehicle is fixed and a LabVIEW application for control parameters tuning and data visualization purposes. Results derived from flights tests using the proposed system support our approach for FPGA based GNC system for the UAV tricopter.

Keywords Unmanned aerial vehicle · FPGA · PID · Fuzzy logic · Inertial navigation system

1 Introduction

In recent years, small rotor craft UAVs such as Tricopters, Quadcopters and Hexacopters have been used in an increasing number of applications: aerial surveillance, firefighting, traffic control, precise agriculture, etc. Small UAVs have

A. Cadena (✉) · R. Ponguillo · D. Ochoa
Facultad de Ingeniería en Electricidad y Computación, Escuela Superior Politécnica del Litoral, Km. 30.5 Vía Perimetral, Guayaquil, Ecuador
e-mail: acadena@espol.edu.ec

R. Ponguillo
e-mail: rponguil@espol.edu.ec

D. Ochoa
e-mail: dochoa@espol.edu.ec

simple structures, basic flight capabilities and are intrinsically unstable systems. External motion perturbations generate linear and angular accelerations that can cause a collision if are not compensated within few milliseconds. For autonomous flight, in which a pre-defined path must be followed, these systems combine guidance and navigation operations into a single unit, hence the name of Guidance, Navigation and Control (GNC) system.

A GNC system typically consists of a guidance law, a controller and navigation algorithms. GNC controllers for small UAVs take as input data from low-cost global positioning devices (GPS) and Micro-electromechanical inertial sensors (MEMS). Despite its low-accuracy a GPS device provides long term flight stability, however they are useless in indoor environments. In such case data from inertial sensors is used instead. The noise of low-cost GPS and MEMS devices degrade the performance of classic controllers. Usually, to reduce the effect of noise the Extended Kalman Filter (EKF) is employed to estimate the true position and orientation of the UAV as a preprocessing step.

Several techniques have been proposed to improve the design of GNC controllers for small UAVs: PID (Arducopter 2015), neural based algorithms (Badia et al. 2007), genetic algorithms (Artale et al. 2014), nonlinear adaptive control (Cai et al. 2007; Barks 2012; Salazar et al. 2006), etc. Although these techniques have proven adequate in terms of stabilization accuracy, their practical use is constrained by the computational resources of the UAV's on-board computer. Ideally, a controller should process input data and compute the controller's output in near real time with the smallest possible memory usage. Several computing technologies have been used in UAVs: microprocessors, ASIC and DSPs. Recently, FPGA has become a viable alternative for control system implementation due to its system design flexibility and low-power consumption. FPGAs can execute computing tasks sequentially, concurrently or a combination of both approaches. Hence, complex GNC routines can be implemented readily. Moreover, FPGA facilitates the integration of additional modules for specific tasks such as in flight data acquisition.

The use of FPGA in small UAVs is not new in François et al. (2009) a genetic algorithm for UAV planning was proposed. However, to the best of our knowledge this is the first attempt to implement a stabilization sub-system in this platform. This paper describes the implementation of a GNC system for a tricopter that runs on a low-cost FGPA board. The GNC controller uses a PID-Fuzzy controller that exploits concurrent programming techniques. To improve computation performance each VHDL code module uses 32 bits floating point arithmetic and high-speed floating-point CORDIC algorithm. The system was simulated and then tested in a controlled environment.

The paper is organized as follows: First, the methodology used is presented. The mathematical model of the airframe is described. Section 4 describes the electronic components. The implementation of a modular GNC using VHDL modules and experimental results are discussed in Sects. 5 and 6. Finally, conclusions and future works are given. Our paper has several contributions: a set of criteria for building a

low-cost UAV tricopter, a PID-Fuzzy controller implemented on a dedicated FPGA and statistics about computational resources used in FPGA based implementation of a GNC system.

2 Methodology

The tricopter used in this work has three rotors and one servo. Based on its kinematic and dynamic model the GNC system was designed. The vehicle has two operation modes, remotely controlled and fully autonomous. For PID-Fuzzy controller tuning and calibration of INS a test stand is used. This setup allows the UAVs to move freely while flight dynamics data is collected. The state vectors from the controller are sent to a computer via radio link and processed to obtain the controller's optimal parameters. These parameters are then stored in the FPGA's memory and used during flight. Figure 1 depicts the elements our experimental setup.

3 Mechanics

The developed tricopter was designed to carry out surveillance missions. The tricopter has a small acoustic emission and a low visual silhouette, basic requirements for surveillance missions (Lin et al. 2015). Our tricopter has a T shape and it is made of structural aluminum. The vehicle has three rotors with 10×4.5 mm propellers and 42 cm distance between each other. The tricopter employs brushless motors D2836/8 with a peak power of 336 W. The aft rotor has tilting capability to
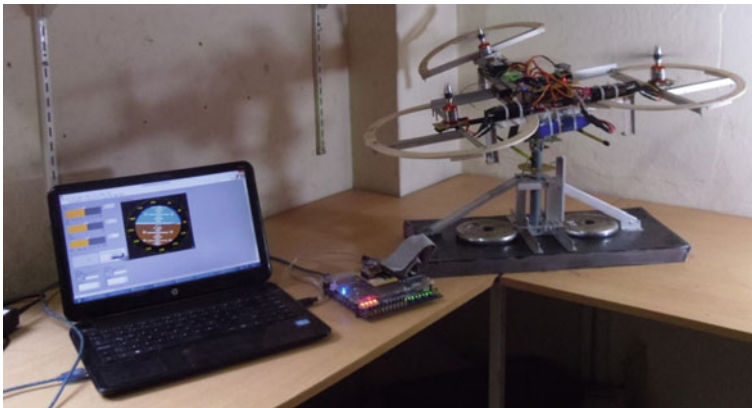

Fig. 1 Experimental setup for GNC system design

change the tricopter's heading, this rotor is mounted on a servomotor controlled base. The FPGA development board and its peripherals are located at the geometrical center of the airframe. Other hardware components were placed accordingly to keep the center of gravity near to the barycenter the frame. Excessive vibration negatively affects the performance of the INS. Hence, the Inertial Measurement Unit (IMU) was mounted in the airframe on a vibration insulator. Finite Elements Analysis (FEA) was carried out to determine the stress and structural vibration (Inman 2007) of the frame. Once all components were mounted the total vehicle's weight was 1150 g.

The orientation of an UAV can be described by Euler angles. The orientation of the airframe with respect to a reference frame is defined by the angles pitch(θ), roll (φ), and yaw(ψ) along the y, x, and z axis respectively. The test stand has a central shaft along the z axis, on the top of this shaft is located a gimbal where the vehicle is attached with a set of screws, near to the vehicle's center of gravity. This configuration allows three rotational degrees of freedom: pitch, roll and yaw. The central shaft also can displace up and down along the z axis giving the translational degree of freedom of heave. The differential equations of the 6 degrees of freedom describing the translational motion, rotational motion, and kinematics are given below (Sai and Hla 2015):

$$\dot{u} = rv - qw - g \sin \theta + F_X/m \tag{1}$$

$$\dot{v} = -ru - pw - g \sin \varnothing \cos \theta + F_y/m \tag{2}$$

$$\dot{w} = qu - pv - g \cos \varnothing \cos \theta + F_z/m \tag{3}$$

$$\dot{p} = (I_{yy} - I_{zz})qr/I_{xx} + L/I_{xx} \tag{4}$$

$$\dot{q} = (I_{zz} - I_{xx})pr/I_{yy} + M/I_{yy} \tag{5}$$

$$\dot{r} = (I_{xx} - I_{yy})pq/I_{zz} + N/I_{zz} \tag{6}$$

$$\dot{\varnothing} = p + q \sin \varnothing \tan \theta + r \cos \varnothing \tan \theta \tag{7}$$

$$\dot{\theta} = q \cos \varnothing - r \sin \varnothing \tag{8}$$

$$\dot{\Psi} = (q \sin \varnothing + r \cos \varnothing) \sec \theta \tag{9}$$

where (Fx, Fy, Fz) are external forces, (L, M, N) are external moments acting on the center of gravity, (u, v, w) are translational velocities, (p, q, r) are rotational velocities and (Ixx, Iyy, Izz) are the inertial moments of the vehicle. These equations are used to design our PID-Fuzzy control approach as explained in Sect. 5.

4 Electronics

The tripcoter's electronic components are presented in Fig. 2. The main element is the onboard computer a FPGA development board (TERASIC DE0 nano). The FPGA reads data from a set of kinematic sensors and radio frequency (RF) communication devices to generate guidance commands codified as PWM signals to the ESC (Electronic Speed Controller) and the aft servo motor. The brushless motors are driven by the ESC with 50 Hz PWM signals, each ESC can drive currents up to 20 A. The servo is controlled by a 50 Hz PWM signal. The sensors and communication devices are connected to the FPGA development board by an IDE port through a daughter card, the port include power supply pins with 5 and 3.3 V.

The sensor suite includes an IMU (Inertial Measurement Unit) CH6 Robotics UM6 composed by 3 axis accelerometer and gyroscope, a 12 channels GPS Parallax and a LV Ultrasonic Rangefinder Maxsonar to measure the distance to the ground. The communication devices are a XBee Pro 900 MHz radio module and a six channel 2.4 GHz Radio Control receiver as a backup for manual control. The same XBee RF device is used to send data to the PC during the calibration phase. Additionally, there is a serial port in the FPGA development board for GNC's algorithms debugging. The power supply is a Li-Po battery of 11.1 V, 2500 mAh, 25 C. A step-down voltage regulator feeds with 5.2 V the FPGA development board. According to our tests this configuration allows for up to 10 min of flight.
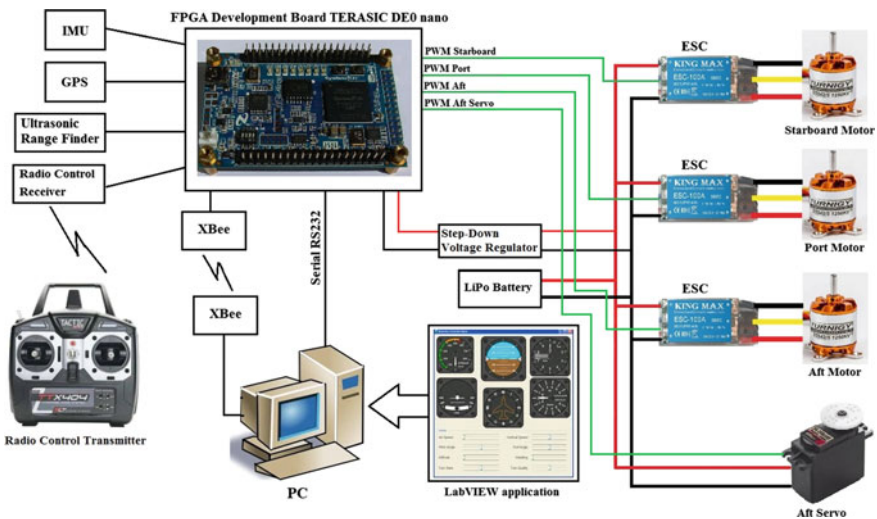

Fig. 2 Tricopter electronics components layout

5   Guidance, Navigation and Control System

All functions of the vehicle are implemented on VHDL modules that run on the
FPGA development board at 50 MHz. The VHDL modules were divided in two
layers:

- A low layer that implements communication protocols for sensors and com-
  munications devices, RAM (Random Access memory) managing and generation
  of PWM signals.
- An upper layer that implements the Guidance, Navigation and Control System.
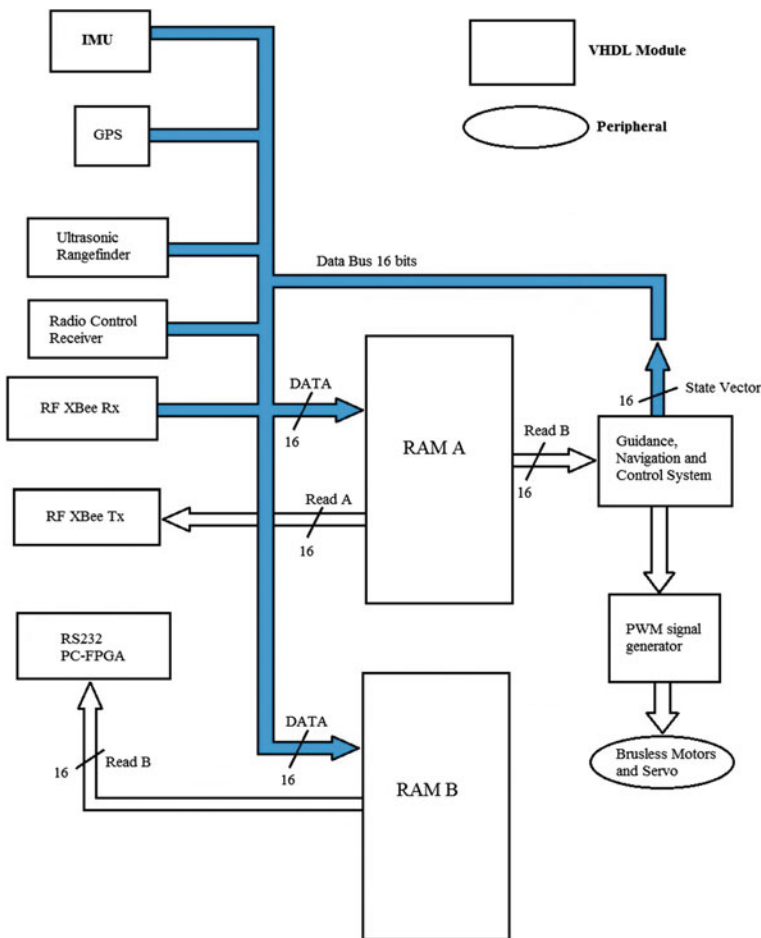
  The VHDL modules layout is depicted in Fig. 3.

Fig. 3 VHDL hardware layout

Each sensor has an associated communication protocol, the IMU uses SPI, the GPS uses NMEA 0183 and the ultrasonic range finder uses RS232 with TTL levels. The communication protocols were implemented in independent VHDL modules. These VHDL modules send its lectures to a pair of RAM called RAM A and RAM B every 20 ms through a data bus of 16 bits and have alarm flags in case of sensor malfunction. The RAM has 64 words of 16 bits and they are implemented with FPGA M9K Embedded Memory Blocks. The write and read processes are driven by a VHDL module. RAM A is accessed by the upper layer and a VHDL module that manages RF communications. The RAM B is accessed by a VHDL module that controls serial communications between the FPGA development board and a PC.

The communication routines with the RF XBee devices are splitted in two modules for transmission and reception tasks respectively. The "RF XBee Tx" module reads the content of the RAM A after the update process and generates a data frame with a checksum. This data frame is processed by the LabVIEW application, running on a PC that logs flight parameters. The "RF XBee Rx" module receives a data frame from a PC. The data frame is encrypted and contains commands from the operator such as update flight trajectory or return to base. The transmission time of the data frames between the RF modules is less than 17 ms. The "Radio Control Receiver" module reads the PWM signals from the Receiver allowing to the operator a manual flight operation with line of sight. The "PWM signal generator" module generates 50 Hz PWM control signal for the ESC and aft servo with 8 bits of resolution, with a duty cycle that in the [1500 1900] µs range.

The upper layer is formed by a single VHDL module that contains several sub-modules that implement logic and mathematical operations for GNC's tasks. The mathematical operations are based on IEEE 754 32 bit floating point arithmetic (Grover and Soni 2014). The trigonometric functions are implemented by high speed floating-point CORDIC algorithm (Singh and Bhad 2013). The implementation of the operations addition, subtraction, multiplication, division and CORDIC using VHDL code take advantage of the FPGA design flexibility to decide what process will be concurrent or sequential. All algorithms were simulated with ModelSim to avoid overflows and singularity points. Implementation was done using logic operators and states machines to achieve processing time below 20 ms. This ensures real time capability for UAV applications and reduce the number FPGA logic elements.

The vehicle uses a classical GNC architecture that is shown in Fig. 4. The function of guidance system is to determine the reference coordinates where the tricopter has to go. We used a fairly simple algorithm called "Line of Sight" (Caccia 2004) that generates the reference yaw angle.

$$\psi_{ref} = \arctan 2(y_{ref} - y_{est}, x_{ref} - x_{est}) \tag{10}$$

The mission manager provides the list of coordinates of a flight path. The INS estimates position, velocity and attitude of the vehicle with respect the reference frame. Since the INS uses low cost IMU input data presents errors such as bias, scale factors, random walk noise and temperature internal compensation.
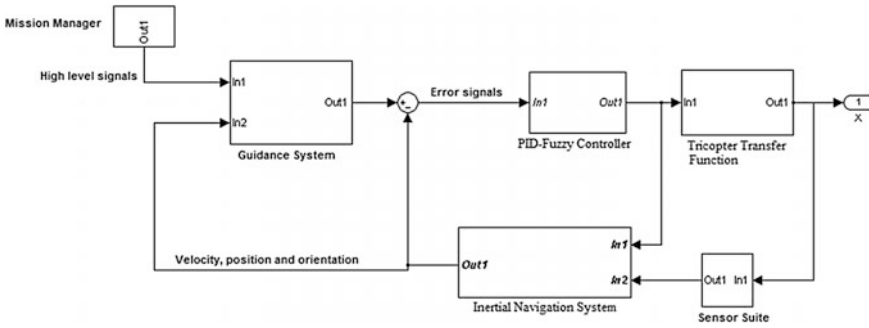
Fig. 4 Guidance, navigation and control system architecture

The Baseline of the Navigation System is the EKF that fuses data captured from the IMU, GPS and range finder sensors. The EKF was constructed using the discrete Kalman filter equations described in (Welch and Bishop 2001). The EKF requires a dynamical model of the vehicle and the noise process in order to minimize the errors of the estimated state vector. The estimated state vector by the EKF is send to the PC.

The vehicle's automatic control is based on PID and a Fuzzy controller, this architecture allows a faster and precise response. The automatic control drives four vehicle's degree of freedom: the Euler angles and the altitude. For this application the gain of the integral controller is zero. The architecture of the proposed controller is presented in Fig. 5.

The PID controllers are implemented by the discrete equation type A (Szafranski and Czyba 2011)

$$
\begin{aligned}
u[k] = u[k-1] + \frac{K_d}{T_s}(e[k] - 2e[k-1] + e[k-2]) \\
+ K_p(e[k] - e[k-1]) + K_i T_s e[k]
\end{aligned} \tag{11}
$$

where k is the sample number, $e(k)$ error signal, $u(k)$ controller output, Ts sampling period, Kp proportional gain, Ki integral gain and Kd derivative gain. This equation was implemented with VHDL code. For Fuzzy controller implementation there are many membership function (Vuong et al. 2006), the most commonly used for UAVs are the triangular, trapezoidal and gaussian membership functions. We chose the triangular and trapezoidal functions. These functions are suitable for VHDL implementation using IF–THEN sentences. The proposed membership rules to describe the flight dynamics are:

- Too negative: m_neg
- Negative: neg
- Zero: zero
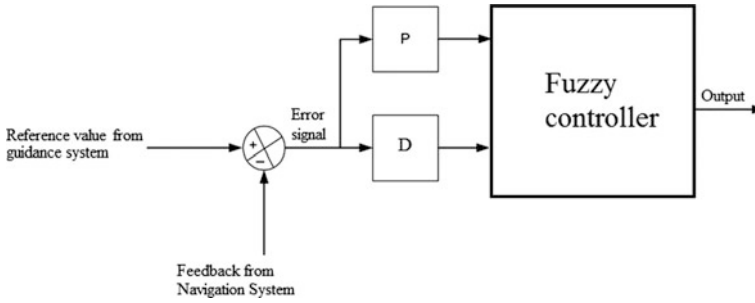- Positive: pos
- Too positive: m_pos

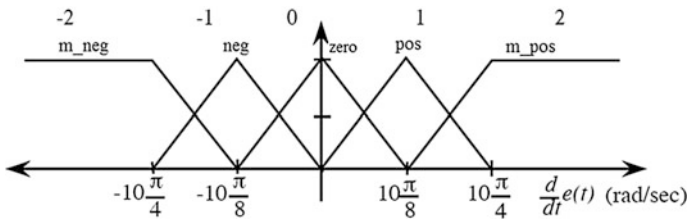Fig. 5   Automatic control system architecture

Fig. 6   Membership function for $\frac{d}{dt}e(t)$

   Figure 6 shows the membership function for $\frac{d}{dt}e(t)$. We used the centre of gravity as the defuzzification method. The GNC system was simulated using MATLAB/Simulink and then tested on the test stand. The PID-Fuzzy controller was optimized to fulfill three objectives: (1) stability generally measured in the frequency domain, (2) steady-state accuracy and (3) transient response, rise time, overshoot and settling time (Rys et al. 2011).

6   Results

The total time to compute an update state vector by the VHDL hardware was less than 1 ms. Hardware Resource Consumption of the DE0 nano FPGA development board was 60 % of FPGA's logic elements, 2 PLL, 5 embedded multipliers, 2176 bits of embedded memory and 24 pins of the IDE port. The GNC tests are divided in three stages. The first stage checks correct performance of the VHDL modules of the lower layer and simulation of the GNC algorithms using MATLAB/Simulink. Every communication protocol for sensors and RF devices was verified after the implementation of the corresponding VHDL module using applications that runs on a PC like Hyperterminal and LabVIEW. In the second stage a calibration process was carried out using the test stand for system identification, development of the

PID-Fuzzy controllers and INS. The vehicle was rigidly adjusted to the gimbal by a set of screws and then the vehicle executed predetermined maneuvers. The information obtained from sensors was used to generate the numeric constants for tuning the controller. The setup for calibration process using the test stand is shown in Fig. 1. The data of the test flight was captured from the onboard sensors through the XBee RF data link. The step response of the PID-Fuzzy controllers was obtained during the indoor test flight and is shown in Fig. 7. The step response fulfills the controller design requirements. During the indoor test flight was found the optimal calibration values for the EKF to obtain a performance of accumulative navigation error of less than 10 % since last GPS position update after 1 h.

Finally, outdoor test flight with winds between 5 and 8 knots was carried out. Previously were defined two flight trajectories by the waypoints (Suzuki et al. 2012). The first trajectory is a simple square, the second trajectory is a lemniscate,
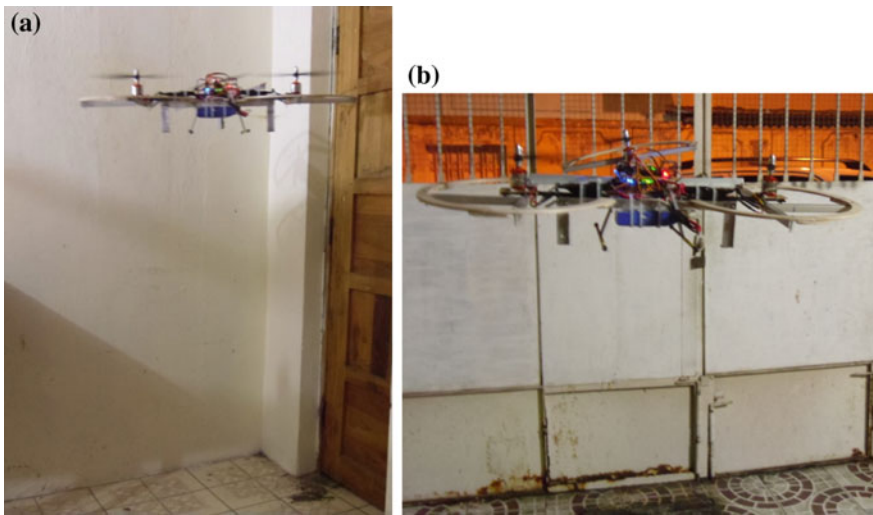

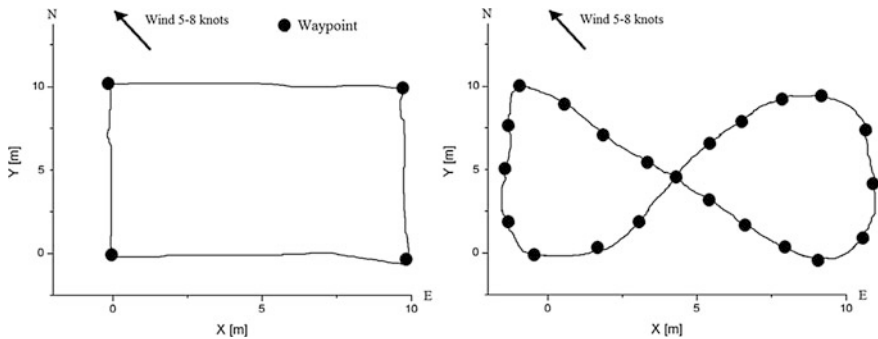**Fig. 7 a** Indoor flight test. **b** Outdoor flight test

Fig. 8 Flight test results: outdoor way point and followed path

path commonly used for surveillance purposes. The coordinates were downloaded to the RAM of the FPGA. The vehicle followed the waypoints successfully compensating external perturbation by winds. Figure 8 shows the paths followed by the vehicle during the test. After the PID-Fuzzy tuning process for the roll and pitch channels, the resulting values of the Kd gain are a bit higher than expected. Usually a high value of Kd generates a fast response with an overshoot (Srikanth et al. 2013) but is compensated by the fuzzy logic controller, this characteristic is useful in flight with heavy wind condition. The vibration affects the performance of the INS, during the tests was necessary to relocate the IMU along the central body of the airframe to obtain a better performance (Fig. 9).
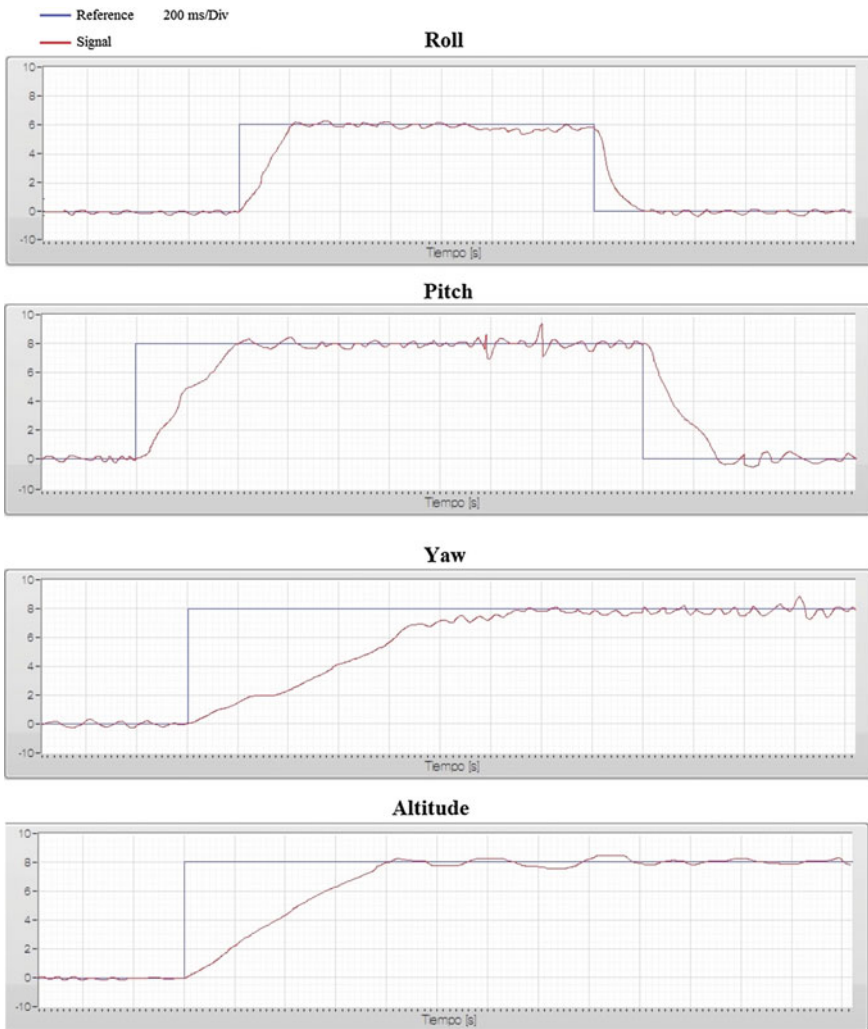

Fig. 9 Flight test results: controller step response

7 Conclusions and Future Works

This paper presented a FPGA-based real time GNC implemented by VHDL modules. The FPGA accomplished the required computational resources demanded by the GNC algorithms using 32 bits floating point arithmetic and CORDIC algorithms. The GNC development process was showed. The tricopter was successfully controlled by the GNC based on EKF and PID-Fuzzy algorithms. The stability control and waypoint navigation using the INS were successfully verified by the flight tests. The PID-Fuzzy controller showed good resistance against wind obtaining a steady flight. The future work will be to improve the GNC with a self-tuning PID-fuzzy controller based on neural networks or genetic algorithms.

References

Arducopter manual on-line http://copter.ardupilot.com/wiki/tuning/, visited 2015.

Artale, V., Milazzo, C., Calogero, O., & Ricciardello, A. (2014). Genetic algorithm applied to the stabilization control of a hexarotor. In Proceedings of the International Conference on Numerical Analysis and Applied Mathematics 2014 (ICNAAM-2014). doi:10.1063/1.4912983

Badia, S., Pyk, P., & Verschure, P. (2007). A fly-locust based neuronal control system applied to an unmanned aerial vehicle: The invertebrate neuronal principles for course stabilization, altitude control and collision avoidance. The International Journal of Robotics Research July 2007, 26, 759–772.

Barsk, K. (2012). Model predictive control of a tricopter. Master thesis, Linkopings Univer-sitet, Linkoping, Sweden.

Caccia, M. (2004) Preliminary sea trials of SESAMO: An autonomous surface vessel for the study of the air-sea interface. CNR-ISSIA Sez. Di Genova, Tech. Rep. Rob-04-SESAMO pt.

Cai, G., Chen, B., Peng, K., Dong, M., & Lee, T. (2007). Design and implementation of a nonlinear flight control law for the yaw channel of a UAV helicopter. In 2007 46th IEEE Conference on Decision and Control (pp. 1963–1968), December 12–14, 2007.

François, C., Mohamed, T., Gilles, L., & Giovanni, F. (2009). FPGA implementation of genetic algorithm for UAV real-time path planning. Journal of Intelligent and Robotic System 2009, 54, 3. doi:10.1007/s10846-008-9276-8

Grover, N., & Soni, M. (2014). Design of FPGA based 32-bit Floating Point Arithmetic Unit and verification of its VHDL code using MATLAB. International Journal of Information Engineering and Electronic Business, 1, 1–14.

Inman, D. (2007). Engineering vibration (3rd Ed.). London: Pearson.

Lin, E., Tan, B., Goh, J., Mok, N., Vengadashalapathy, L., Yoong, T., & Srigrarom, S. (2015). Development of UGS tilt-rotor surveillance tricopter UAV. Journal of Applied Science and Engineering, 18(2), 177–186.

Rys, A., Czyba, R., & Szafranski, G. (2011). Practical aspects of trirotor MAV development. In Proceedings of the International Micro Air Vehicles Conference 2011 (Summer Ed., pp. 64–69), Harde, The Netherlands.

Sai, K., & Hla, M. (2015). Modeling and analysis of tri-copter (VTOL) aircraft. International Journal of Emerging Engineering Research and Technology, 3(6), 54–62.

Salazar, S., Kendoul, F., Lozano, R., & Fantoni, I. (2006). Real-time control of a small-scale helicopter having three rotors. In International Conference on Intelligent Robots and Systems (pp. 2924–2929), Beijing, China.

Singh, A., & Bhad, D. (2013). Analysis of simple CORDIC algorithm using MATLAB. *International Journal of Scientific & Engineering Research*, *4*(6).

Srikanth, G., Tarun, A., Sai, K., Suraj, C., Ramesh, G., & Veena, D. (2013). Design of multicopter test bench international. *Journal of Modeling and Optimization*, *3*(3).

Suzuki, S., Ishii, T., Okada, N., Iizuka, K., & Kawamura, T. (2012). Autonomous navigation, guidance and control of small electric helicopter. *INTECH International Journal of Advanced Robotic Systems*.

Szafranski, G., & Czyba, R. (2011). Different approaches of PID control UAV type quadrotor. In Proceedings of the International Micro Air Vehicles conference 2011 (Summer Ed.).

Vuong, P., Madni, A., & Vuong, B. (2006). VHDL implementation for a fuzzy logic controller. World Automation Congress (WAC), Budapest, Hungary.

Welch, G., & Bishop, G. (2001). An introduction to the Kalman filter. In *Computer graphics. Annual conference on computer graphics & interactive techniques*. Addison-Wesley, Los Angeles, CA, USA: ACM Press. SIGGRAPH 2001 course pack edition, course No. 8.

Fault Recoverability Analysis
via Cross-Gramian

Hamid Reza Shaker

Abstract Engineering systems are vulnerable to different kinds of faults. Faults may compromise safety, cause sub-optimal operation and decline in performance if not preventing the whole system from functioning. Fault tolerant control (FTC) methods ensure that the system performance maintains within an acceptable level at the occurrence of the faults. These techniques cannot be successful if the necessary redundancy does not exist in the system. Fault recoverability which is also known as control reconfigurability is a mathematical measure which quantifies the level of redundancy in connection with feedback control. Fault recoverability provides important and useful information which could be used in analysis and design. However, computing fault recoverability is numerically expensive. In this paper, a new approach for computation of fault recoverability for bilinear systems is proposed. This approach uses cross-gramian and reduces the computations significantly. The contribution of this paper is twofold. Firstly the concept of cross-gramian is extended to support discrete-time bilinear systems and an iterative algorithm for cross-gramian computation is proposed. Secondly a cross-gramian based approach for computation of fault recoverability is proposed which reduces the computational burden significantly. The proposed results are used for an electro-hydraulic drive to reveal the redundant actuating capabilities in the system.

Keywords Discrete-time bilinear systems · Cross-gramian · Fault recoverability · Fault-tolerant control · Electro-hydraulic system

H.R. Shaker (✉)
Center for Energy Informatics, The Maersk Mc-Kinney Moller Institute,
University of Southern Denmark, Odense, Denmark
e-mail: hrsh@mmmi.sdu.dk

1 Introduction

Engineering systems are vulnerable to different kinds of faults. Typical faults are sensor faults, actuator faults, equipment fouling, and many others (Harris et al. 1999). The faults may compromise safety, cause sub-optimal operation and decline in performance if not leading to overall system failure (Harris et al. 1999; Shaker and Tahavori 2011).

Fault tolerant control (FTC) methods ensure that the system performance remains in an acceptable level at the occurrence of the faults. These methods therefore reduce the costs significantly, ensure safety and guarantee an acceptable performance. The existence of the sufficient redundancy in systems and processes is an important prerequisite for the success of such methods.

The level to which a system can bear faults through feedback control is an inherent characteristic property of a system.

A system without sufficient control redundancy cannot be made to tolerate faults no matter which feedback control strategy is employed.

There are different kinds of redundancy which have been introduced in literature which are the results of developments in fault detection and diagnostics. The notion of redundancy which has been introduced in Wu et al. (2000) evidently reveals the inherent capability of a system to tolerate faults through feedback control. In Wu et al. (2000) fault recoverability which is also known as control reconfigurability is introduced. Fault recoverability is a mathematical measure which quantifies the level of redundancy in connection with feedback control. This measure has been later extended to support bilinear systems in Shaker (2013), Shaker and Tahavori (2011). The concept of the second order modes has been use to define a fault recoverability measure. In Patton (1997), Shaker (2013), Shaker and Tahavori (2011) the observability and the controllability gramians of the system are obtained first and are used to compute the fault recoverability. However, this process is numerically expensive since for obtaining each gramian one (generalized) Lyapunov equation has to be solved.

In this paper a new method for computing fault recoverability is proposed. This method supports bilinear systems and therefore can easily support linear system. This approach uses cross-gramian and reduces the computations significantly. The contribution of this paper is twofold. First the concept of cross-gramian is extended to support discrete-time bilinear systems and an iterative algorithm for cross-gramian computation is derived. Second a cross-gramian based approach for computation of fault recoverability is proposed. In previously proposed approaches in Patton (1997), Shaker (2013), Shaker and Tahavori (2011), two (generalized) Lyapunov equations are needed to be solved for each parameter values. However, in the method which is proposed in this paper only one (generalized) Sylvester equation has to be solved for each parameter value. Therefore the computational burden reduces significantly.

The results in this paper support bilinear systems. Bilinear systems are important class of nonlinear systems and they have already been used successfully in a lot of applications (Deutscher 2005; Mohler 1991; Svoronos et al. 1980).

The paper is organized as follows. In the next section, we review the concept of gramians for discrete-time bilinear systems. In Sect. 3 we extend the concept of cross-gramians to support discrete-time bilinear systems. An iterative method for computation of the cross-gramian is also presented in this section. Section 4 presents how cross-gramian can be used for computing the fault recoverability. In Sect. 5 the proposed fault recoverability measure is used for the bilinear models of a hydraulic drive. The last section concludes the paper.

The notation used in this paper is as follows: M^* denotes the transpose of matrix if $M \in \mathbb{R}^{n \times m}$ and complex conjugate transpose if $M \in \mathbb{C}^{n \times m}$. $\lambda_i(M)$ stands for the eigenvalues of matrix M.

2 Review on Controllability and Observability Gramian for Discrete-Time Bilinear Systems

The controllability and the observability gramians are well-known matrices, which are widely used to check the controllability and the observability of the dynamical systems. These gramians are solutions to (generalized) Lyapunov equations and are very popular for model order reduction and control configuration selection (Antoulas 2005; Gugercin and Antoulas 2004; Shaker and Komareji 2012, 2013, 2014a, b). The controllability and observability gramians show how difficult a system is to control and to observe.

For a bilinear dynamical system which is described by:

$$\Pi: \begin{cases} x(k+1) = Ax(k) + Nx(k)u(k) + Bu(k), \\ y(k) = Cx(k), \end{cases} \tag{1}$$

where $x(t) \in \mathbb{R}^n, u(t) \in \mathbb{R}, y(t) \in \mathbb{R}^m$, the controllability gramian P is defined as (d'Alessandro et al. 1974; Dorissen 1989; Zhang et al. 2003):

$$P := \sum_{i=1}^{\infty} \sum_{k_i=0}^{\infty} \cdots \sum_{k_1=0}^{\infty} P_i P_i^*, \tag{2}$$

where

$$\begin{aligned} P_1(k_1) &= A^{k1} B, \\ P_i(k_1, \ldots, k_i) &= A^{k_i} [N P_{i-1}], \end{aligned} \tag{3}$$

and the observability gramian Q is defined as (d'Alessandro et al. 1974; Dorissen 1989; Zhang et al. 2003):

$$Q := \sum_{i=1}^{\infty} \sum_{k_i=0}^{\infty} \cdots \sum_{k_1=0}^{\infty} Q_i^* Q_i, \tag{4}$$

where

$$
\begin{aligned}
Q_1(k_1) &= CA^{k1}, \\
Q_i(k_1, \ldots, k_i) &= [Q_{i-1}N]A^{k_i}.
\end{aligned}
\tag{5}
$$

The gramians are given by the solutions of the generalized Lyapunov equations (Zhang et al. 2003):

$$
\begin{aligned}
APA^* - P + NPN^* + BB^* &= 0, \\
A^*QA - Q + N^*QN + C^*C &= 0.
\end{aligned}
\tag{6}
$$

The generalized Lyapunov equations can be solved iteratively. The controllability gramian P can be obtained by Zhang et al. (2003):

$$P = \lim_{i \to \infty} \widehat{P}_i \tag{7}$$

where

$$
\begin{aligned}
A\widehat{P}_1 A^* - \widehat{P}_1 + BB^* &= 0, \\
A\widehat{P}_i A^* - \widehat{P}_i + N\widehat{P}_i N^* + BB^* &= 0, \quad i = 2, 3, \ldots.
\end{aligned}
\tag{8}
$$

In other words, all what we need to do is to run (8) and to compute $\widehat{P}_i$'s and terminate the computation when $\widehat{P}_i$ converges to a constant matrix. Dually the observability gramian can be computed in the similar way.

As in the linear case the Hankel singular values (HSV), i.e., the diagonal entries of the balanced P and Q, are the square roots of the eigenvalues of the gramians product, i.e. $\sqrt{\lambda_i(PQ)}$. The Hankel singular values show how difficult a system is to control and to observe. More details are discussed in Antoulas (2005), Zhang et al. (2003).

3   Cross-Gramian for Discrete-Time Bilinear Systems

Cross-gramian which is well-known for linear system contains both controllability and observability information. For bilinear dynamical systems in the form (1), we define cross-gramian as:

$$W_c := \sum_{i=1}^{\infty} \sum_{k_i=0}^{\infty} \cdots \sum_{k_1=0}^{\infty} P_i Q_i, \tag{9}$$

where P_i and Q_i are defined in (3) and (5).

It can be shown that the cross-gramian is the solution to the following Sylvester equation:

$$AW_{co}A - W_{co} + NW_{co}N + BC = 0 \tag{10}$$

The generalized Sylvester Eq. (10) can be solved iteratively. The cross-gramian W_{co} can be obtained by:

$$W_{co} = \lim_{i \to \infty} \widehat{W}_{coi} \tag{11}$$

where

$$
\begin{aligned}
A\widehat{W}_{co1}A - \widehat{W}_{co1} + BC &= 0, \\
A\widehat{W}_{coi}A - \widehat{W}_{coi} + N\widehat{W}_{coi}N + BC &= 0, \quad i = 2, 3, \ldots.
\end{aligned}
\tag{12}
$$

In other words, all what we need to do is to run (12) and to compute $\widehat{W}_{coi}$'s and terminate the computation when $\widehat{W}_{coi}$ converges to a constant matrix.

An interesting property of the cross-gramian is that the absolute value of the eigenvalues of the generalized cross-gramian equal Hankel singular values.

4 Fault Recoverability for Discrete-Time Bilinear Systems

For a given bilinear model, the smallest Hankel singular value of the bilinear system is used to define the fault recoverability. This measure reveals if there exists adequate redundancy in the process. The smallest Hankel singular value changes when a fault happens. In order to see if the level of redundancy is adequate for a particular fault, the amount of decrease in the smallest Hankel singular value needs to be observed for that particular fault. In such a way, we can check if the FTC methods can be effective for a particular model and a fault scenario.

Analogous to Wu et al. (2000), Shaker (2013), Shaker and Tahavori (2011), it is assumed that the faults can be presented in the model as parameters. This is a popular assumption in this context and in practice most of the faults can be represented as parameters in the model.

Consider a parameter varying bilinear model defined over the parameter set Ψ:

$$\Pi_\Psi: \begin{cases} x(k+1) = A(\xi)x(k) + N(\xi)x(k)u(k) + B(\xi)u(k), \\ y(k) = C(\xi)x(k), \end{cases} \tag{13}$$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}$, $y(t) \in \mathbb{R}^m$, $\xi \in \Psi$.

Assume without loss of generality that for $\xi = 0$ the model is fault free (healthy).

The cross-gramian is given by the solution of the generalized Sylvester equation over Ψ:

$$A(\xi)W_{co}(\xi)A(\xi) - W_{co}(\xi) + N(\xi)W_{co}(\xi)N(\xi) + B(\xi)C(\xi) = 0 \tag{14}$$

The fault recoverability ρ_S over the parameter set $S \subset \Psi$ for a bilinear model (13) is defined:

$$\rho_S := \min_{\xi \in S \subset \Psi} |\lambda_i(\mathbf{W}_{co})| \tag{15}$$

This measure not only quantifies the potentiality of a bilinear process to tolerate faults/changes, but also assists in making decisions for improvement of the bilinear system's redundancy configuration, such as sensor/actuator placement.

5   Fault Recoverability for Hydraulic Drive System

In this section, the fault recoverability is computed for the bilinear models of an electro-hydraulic drive.

A typical hydraulic system is a highly nonlinear system. Because of severe nonlinear behavior, the linear models are usually not sufficiently accurate to describe them and consequently the controllers which are designed based on the linear models of the hydraulic systems quite often do not end up with satisfying results in practice. To remedy this more complex models and consequently more complex methods are required. Bilinear system are a class of nonlinear systems with a well-established methods developed for them. Bilinear systems fit well for a lot of mechatronic hydraulic systems (Guo and Schwarz 1989; Guo et al. 1994; Schwarz et al. 1988; Schwarz and Ingenbleek 1994).

In the following the fault recoverability is computed for the bilinear models of an electro-hydraulic drive system. This electro-hydraulic drive was studied in Schwarz and Ingenbleek (1994). The bilinear model is a result of bilinearization methods which have been used to approximate this highly nonlinear system as several bilinear models (Schwarz and Ingenbleek 1994). Depending on the bilinear approximation method, different bilinear models for a particular process may be found or selectively constructed. In Schwarz and Ingenbleek (1994), three different forth order SISO bilinear models are presented which approximate hydraulic nonlinear drive system.

The bilinear models of the hydraulic drive system in Schwarz and Ingenbleek (1994), are SISO and continuous-time:

$$\begin{cases} \dot{x}(t) = A_c x(t) + N_c x(t) u(t) + B_c u(t), \\ y(t) = C_c x(t), \end{cases} \tag{16}$$

where $x(t) \in \mathbb{R}^4$, $u(t) \in \mathbb{R}$, $y(t) \in \mathbb{R}$.

Euler's forward discretization method is used to discretize this model and the sampling time $T_s = 0.1$ is used. The resulting discrete-time models are in the form:

$$\begin{cases} x(k+1) = Ax(k) + Nx(k)u(k) + Bu(k), \\ y(k) = Cx(k). \end{cases} \tag{17}$$

Let the fault which is needed to be studied be an actuator fault. The bilinear model in the form (13) in this case would be:

$$\begin{cases} x(k+1) = Ax(k) + Nx(k)u(k) + B(\xi)u(k), \\ y(k) = Cx(k), \end{cases} \tag{18}$$

where

$$B(\xi) = B + \begin{bmatrix} b_1\xi_1 & b_2\xi_2 & \dots & b_m\xi_m \end{bmatrix} \tag{19}$$

and b_i is the ith column of B. ξ_i's are actuator effectiveness which range from -1 to 0. Matrices A, N, C are constant and $S := \{\xi : -1 \le \xi_i \le 0\}$.

The fault recoverability of the bilinear models of the hydraulic drive in Schwarz and Ingenbleek (1994) for the actuator fault scenario are shown in Figs. 1, 2 and 3.

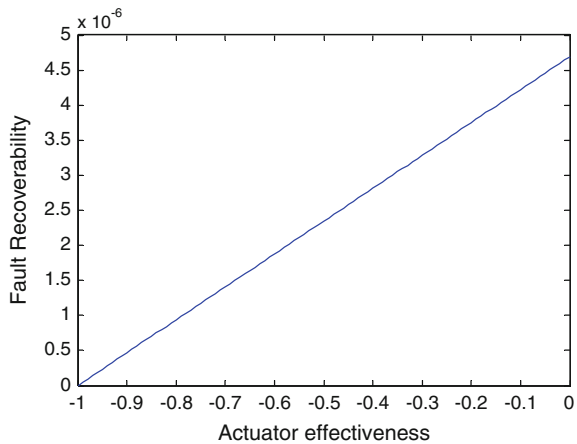Fig. 1 The fault recoverability for the first bilinear model of Schwarz and Ingenbleek (1994)

Fig. 2 The fault
recoverability for the second
bilinear model of Schwarz
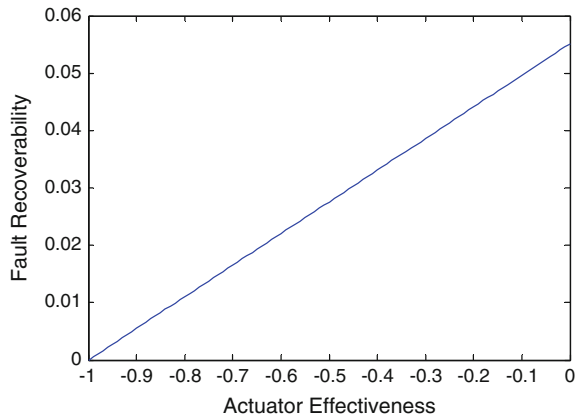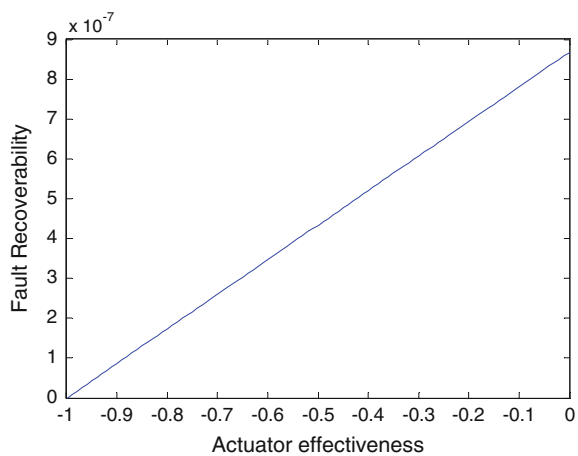and Ingenbleek (1994)

Fig. 3 The fault
recoverability for the third
bilinear model of Schwarz
and Ingenbleek (1994)

These figures show how the pointwise fault recoverability ρ_S varies in terms of control effectiveness ξ over S. As expected, fault recoverability is zero, when the worst-case scenario of a complete loss of effectiveness in the actuator happens. More details on how to use the fault recoverability measure for control and design purposes can be found in Wu et al. (2000). It should be noted that in the proposed approach for fault recoverability computation for each value of the parameter only one Sylvester equation in form of (10) has to be solved. However, in the previously proposed methods for each value of the parameters two Lyapunov equations in form of (6) have to be solved. Therefore the proposed method reduces the computational burden significantly.

6 Conclusion

The information about the level of the redundancy is important for FTC since existing a sufficient level of redundancy in the system is an important prerequisite for the success of FTC methods. Fault recoverability which is also known as control reconfigurability quantifies the level of redundancy in connection with feedback control. Computing fault recoverability is numerically expensive. In this paper, a new approach for computation of fault recoverability for bilinear systems has been proposed. Firstly the concept of cross-gramian has been extended to support discrete-time bilinear system. Then it has been used to compute fault recoverability. This approach only requires solving one Sylvester equation instead of two Lyapunov equations for each parameter and therefore reduces the computations significantly. The proposed results have been used for an electro-hydraulic drive to reveal the redundant actuating capabilities in the system.

References

Antoulas, A. C. (2005). *Approximation of large-scale dynamical systems* (Vol. 6). Siam.

d'Alessandro, P., Isidori, A., & Ruberti, A. (1974). Realization and structure theory of bilinear dynamical systems. *SIAM Journal on Control, 12*(3), 517–535.

Deutscher, J. (2005). Nonlinear model simplification using L 2-optimal bilinearization. *Mathematical and Computer Modelling of Dynamical Systems, 11*(1), 1–19.

Dorissen, H. (1989). Canonical forms for bilinear systems. *Systems & Control Letters, 13*(2), 153–160.

Gugercin, S., & Antoulas, A. C. (2004). A survey of model reduction by balanced truncation and some new results. *International Journal of Control, 77*(8), 748–766.

Guo, L., Schöne, A., & Ding, X. (1994). Control of hydraulic rotary multi-motor systems based on bilinearization. *Automatica, 30*(9), 1445–1453.

Guo, L., & Schwarz, H. (1989). *A control scheme for bilinear systems and application to a secondary controlled hydraulic rotary drive.* Paper Presented at the Proceedings of the 28th IEEE Conference on Decision and Control, p. 989.

Harris, T., Seppala, C., & Desborough, L. (1999). A review of performance monitoring and assessment techniques for univariate and multivariate control systems. *Journal of Process Control, 9*(1), 1–17.

Mohler, R. R. (1991). *Nonlinear systems, volume II: Applications to bilinear control: Englewood-cliffs.* New Jersey: Prentice-Hall.

Patton, R. J. (1997). *Fault-tolerant control systems: The 1997 situation.* Paper presented at the IFAC symposium on fault detection supervision and safety for technical processes.

Schwarz, H., Dorissen, H., & Guo, L. (1988). Bilinearization of nonlinear systems. *Systems Analysis and Simulation, 46*, 89–96.

Schwarz, H., & Ingenbleek, R. (1994). Observing the state of hydraulic drives via bilinear approximated models. *Control Engineering Practice, 2*(1), 61–68.

Shaker, H. R. (2013). Control reconfigurability of bilinear systems. *Journal of Mechanical Science and Technology, 27*(4), 1117–1123.

Shaker, H. R., & Komareji, M. (2012). Control configuration selection for multivariable nonlinear systems. *Industrial and Engineering Chemistry Research, 51*(25), 8583–8587.

Shaker, H. R., & Tahavori, M. (2011). *Control reconfigurability of bilinear hydraulic drive systems.* Paper presented at the International Conference on Fluid Power and Mechatronics (FPM).

Shaker, H. R., & Tahavori, M. (2013). Frequency-interval control configuration selection for multivariable bilinear systems. *Journal of Process Control, 23*(6), 894–904.

Shaker, H. R., & Tahavori, M. (2014a). Frequency-interval model reduction of bilinear systems. *Automatic Control, IEEE Transactions on, 59*(7), 1948–1953.

Shaker, H. R., & Tahavori, M. (2014b). Time-interval model reduction of bilinear systems. *International Journal of Control, 87*(8), 1487–1495.

Svoronos, S., Stephanopoulos, G., & Aris, R. (1980). Bilinear approximation of general non-linear dynamic systems with linear inputs. *International Journal of Control, 31*(1), 109–126.

Wu, N. E., Zhou, K., & Salomon, G. (2000). Control reconfigurability of linear time-invariant systems. *Automatica, 36*(11), 1767–1771.

Zhang, L., Lam, J., Huang, B., & Yang, G.-H. (2003). On gramians and balanced truncation of discrete-time bilinear systems. *International Journal of Control, 76*(4), 414–427.

Implementation of RFID-Based Car Ignition System (CIS) in Kazakhstan

Nurbek Saparkhojayev, Askar Kurymbayev and Azret Akhmetov

Abstract Car Ignition System (CIS) based on Radio frequency identification technology (RFID) will allow complete ignition access via RFID card, controlling the start system of the car and securing cars from autostealings. The proposed system is based on UHF RFID readers, Arduino microcontrollers, and identification cards containing RFID-transponders that are able to electronically store data that can be read/written even without the physical contact with the help of radio medium. This system is an innovative system, which describes the benefits of applying RFID-technology in the car security process. This paper presents the experiments conducted to set up RFID based CIS in Kazakhstan.

Keywords RFID · Arduino · Car ignition system · Smart system

1 Introduction

Radio frequency identification (RFID) is a rapidly growing technology that has high potential to make great economic impacts on many fields of industry. While RFID is a relatively old technology, more recent advancements in chip manufacturing technology and new applied security algorithms are making RFID practical for new applications and settings, particularly consumer item level tagging. These advancements have the potential to revolutionize mainly supply-chain management, inventory control, and logistics. At its most basic form, RFID systems consist of small transponders, also known as tags, attached to physical objects.

There are many types of RFID systems, which used in different applications and settings. These systems differ according to their power sources, operating frequencies, and functionalities. The properties and regulatory restrictions of a particular RFID system will determine its manufacturing costs, physical specifications,

N. Saparkhojayev (✉) · A. Kurymbayev · A. Akhmetov
"IT in Education" Scientific Lab, Almaty Management University,
227 Rozybakiev str, Almaty, Kazakhstan
e-mail: nursp81@gmail.com

and performance. Some of the most familiar RFID applications are item-level tagging with electronic product codes, proximity cards for physical access control, and contact-less payment systems. It is historically well known fact that the roots of radio frequency identification technology can be traced back to time when World War II was taken place (Violino 2005). Equipment monitoring and tracking system based on RFID technology is a complex, integrated system that presents an effective solution of maintaining items especially for large-scale environment (Sharma et al. 2014). This paper proposes the different components of RFID technology and focuses on its main core competencies, scalability and security. It will then provide detailed description of RFID-based CIS in Kazakhstan. As the technology grows very precipitately, RFID has received considerable attention worldwide and widely used in controlling and tracking objects ranging from human identification to product identification. Previous research done by many researchers and scientists have successfully showed that RFID has been increasingly expanded in various fields such as retail supply chain, animal identification, metro pass cards and e-passports, asset tracking, postal and courier services, construction industry, education, medical etc. Developments in RFID technology continue to relinquish capacities of the memory, wider reading ranges, and faster processing. It is not possible that the technology will ultimately replace barcode but RFID will continue to grow in its established niches where barcode or other optical and wireless technologies are not effective, such as in the chemical container industry and livestock industry. RFID enables tracking and controlling of items over distances that range from about a centimeter to hundreds of meters. RFID can track any items starting from inventories, mobile handy equipment; moreover, it can track people in real time as the tagged item travels around the organization. According to Domdouzis et al. (2007) tracking of mobile equipment can include wheelchairs, infusion pumps, and blood supplies. RFID systems have been widely used in many different application areas, among of which the most popular are followings: product tracking via manufacturing and assembly, inventory control, parking lot access and its control, container tracking, ID badges and access control, equipment tracking in hospitals and medical centers, etc. Compared to many different automatic identification technologies, and especially, compared to optical barcode systems, RFID-technology has important benefits, like the following: tag data can be read automatically beyond the line of sight, through certain materials, and from a range of several meters (Weis et al. 2003).

2 RFID System Components

RFID is a technology that uses radio waves to transfer data from an electronic tag—called an RFID tag or label, which is attached to an object—through a reader for the purpose of identifying and tracking the object (Fig. 1).
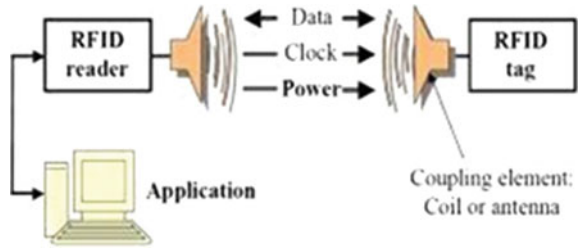
A conventional RFID system is made up of four components mainly,

1. RFID Reader;
2. RFID Antenna;
3. RFID transponder (or tag) electronically programmed with unique data;
4. Host computer with appropriate application software.

RFID systems consists of an antenna and a transceiver, which read the radio frequency and transfer the information to a processing device, and a transponder, or tag, which is an integrated circuit containing the RF circuitry and information to be transmitted. The system requires the usage of RF active readers able to communicate with the main station using three different communication mediums: serial cable, LAN cable, and WLAN antenna.

The proposed system is based on a MIFARE RFID-tag, specifically, the MIFARE MF1ICS50 typed RFID-tag. This type of tag was developed by the company NXP to be used in a contactless smart card according to ISO/IEC 14443 Type A. The MIFARE MF1ICS50 IC is used in such applications as public transportation ticketing, which major cities of the world have adopted as their e-ticketing solution. The MF1ICS50 chip consists of a 1 K-byte EEPROM, a RF-Interface and a Digital Control Unit. Energy and data are transferred via an antenna comprising a coil with a few turns directly connected to the MF1ICS50 (NXP).

For reading cards, device based on Arduino was developed, and this device acts as RFID-reader, and it enables the contact-free reading and writing of operations and works on a 13.56 MHz frequency (Fig. 2).
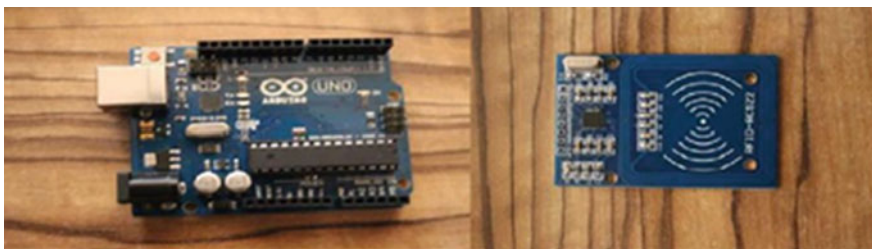

Fig. 2 The components of device: *left* Arduino UNO microcontroller, *right* RFID-reader

3 The Development Process of CIS

The automobile market was a testing ground for RFID technology for quite a while. The first to come up with an idea to embed radio chips were Germans (Volkswagen) as a solution for monitoring pressure gauge valves and as a part of anti-theft alarm system (SCDigest Editorial Staff 2009). Their idea was picked up by Japanese (Nissan, Toyota) and was used as a basis for keyless ignition system (Swedberg 2013; Drickhamer 2005). This system is a supplement to the regular alarm, which allows you to use individual card-key (which based on RFID card) to the ignition of the engine of the car, which means that another person has not access to the car. The most important point of using this system is that the possibility of having car ignition turn to the position "start" without ignition key and a card-key. The idea of this system is that one unique card can be used for several operations, such as: access to input/output building, ignition of the car, lock/unlock the car and many other operations. There are many similar system existing on the market, but we have researched two popular systems on Kazakhstani market that have many customers, and these systems are Pandora DXL 3910 and Starline B64 Dialog CAN respectively.

Pandora DXL 3910 system of Russian company Alarm Trade that differs from virtually all that can be seen in the market of auto accessories. The owner of a vehicle equipped by anti-theft system Pandora DXL 3910 does not carry with keychain. Basic operations are made using labels (so called simple key chains, usually with 2 buttons are not equipped with the display device, and transmitting the central unit only 2 basic commands—to enable or disable protection regime) and full control of the car alarm via the smartphone. Owner can use this voice for the GSM interface or install the mobile application that extends this system. Pandora Info program exists in versions for mobile devices on the IOS platform and Android. Flexible settings of system makes it uniquely adaptable to many requirements of the owner. Pandora DXL 3910 is perfectly integrated into the standard system of any modern car. However, we found main disadvantage of Pandora DXL 3910, as usual, has become one of its advantages—flexible programming. The complexity and abundance of options confuses even experienced owners of automotive electronics (AlarmTrade).

Describing the Starline system, needs to mention main points of this system. Implemented for its intended purpose, StarLine B64 Dialog CAN can serve as the control center service functions—interior lighting, temperature control, display of open door and many others. On the device there are two key chains—one with an LCD display, the other is more compact, but without indication. Both carry out two-way communication with a central unit at a distance of 2 km. The relatively low price was due to the absence in the immobilizer unit and GPS. According to customers of different kinds of security systems with additional services, lack of StarLine B64 Dialog CAN is the lack of function of the engine startup. In addition, base system is not provided for nonvolatile mode (StarLine).
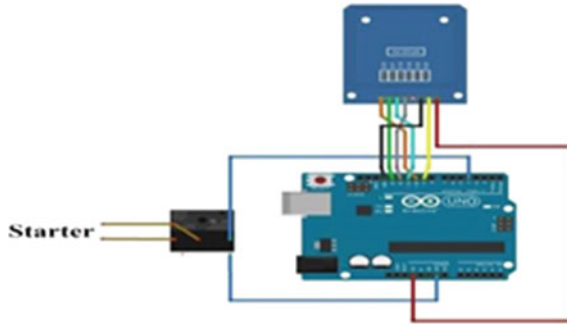
Fig. 3 Components of the system and their inter-connections

Considering all these facts, we decide to build our own ignition system that needs to be cheap, effective in terms of engine startup, and the main point on which we have put our main attention is that we will be using the University's individual card distributed by Administration of University. These cards permit to enter/leave university, access to the university library, as well to determine the student's card in public places. Retaining all the data and functions of the card, additional operation named "access to a car" has been developed. So that, using the university's card, it becomes possible to turn on a car's ignition. This system is capable of identifying the owner/driver and it is not an easy way to copy the unique key, which means system protects from theft (Fig. 3).

There are two different approaches of implementing such system:

1. One card-key for one car, only one person has an access to the car's ignition.
2. Seven key card for access to the car (e.g. case of large companies in the state of multiple drivers, to have access to the same vehicle).

The programming of the device is implemented on Arduino programming language based on C/C++. The microcontroller on the board is programmed by using the language Arduino (based on language Wiring) and the IDE Arduino (based on medium Processing).

The pseudo-code for reading cards is as follows:

if (checkTag(tagNumber)){// Request for comparing the value that is stored in memory

Serial.print("Start");// If value is the same with the value in memory digitalWrite (10,HIGH);// turn on "start"

The workflow of the algorithm is shown in Fig. 4.

The algorithm of reading cards described as follows:

Generally, the algorithm of the reading cards works as follows: As it may be seen, any card will be firstly passed thru RFID reader, which checks ID and then in case of successful operation, sends signal to Arduino. After this step, Arduino sends
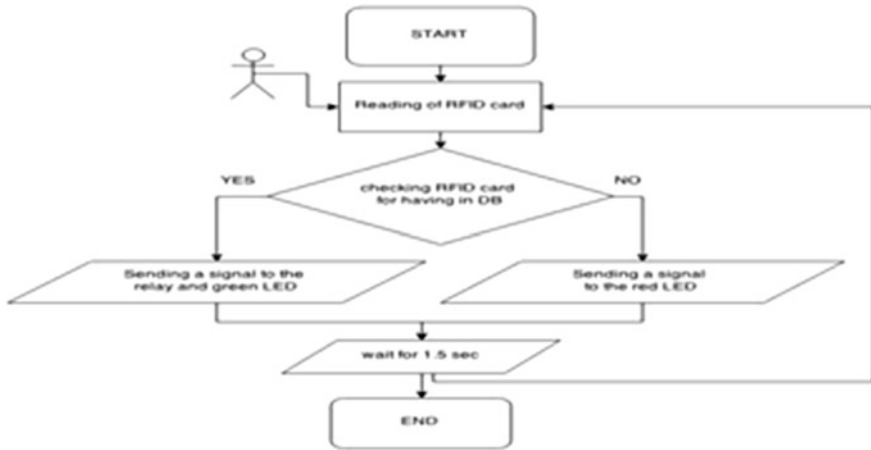
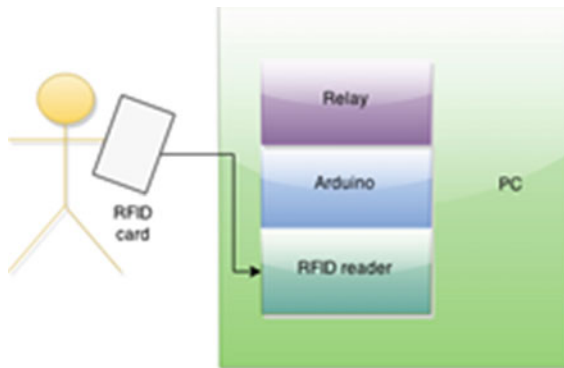**Fig. 4** Workflow of the CIS

Fig. 5 The algorithm of reading cards

signal to relay, and then car will be ready for driving. Connection to PC is needed in checking phase, on which system authorizes every registered user, that has ability to start car from their RFID cards (Fig. 5).

In Fig. 6, there are two pictures, one in left—shows how the cables and all sensors, and Microcontroller Arduino is connected with the car, and the second one in right shows how the car is started to work with a help of card. As it may be seen, the card sends signal to the relay, and it has been checked there, and in case of having access to start and drive the car, car will automatically start working.
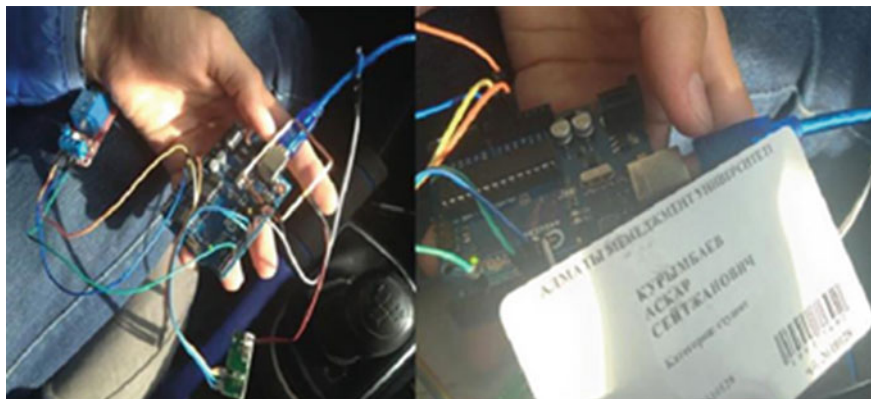
Fig. 6 *Left one* the connection of wires, sensors and Microcontroller with car; *right one* the start of car by swiping the card

4 Conclusion

In this research paper, the description of the research conducted in Almaty Management University based on RFID-technology was presented. This research showed the efficiency of applying this system for securing the car from possible steals by providing only authorized owners of cars to turn on ignition system of the car by applying his/her ID-card. We plan to extend this research by adding some functionalities, and by making appropriate steps in developing more specific architecture for CIS. Meanwhile, we would like to combine all our previous research works (Saparkhojayev and Shaiken 2013; Saparkhojayev and Guvercin 2012; Saparkhojayev 2015; Saparkhojayev et al. 2014; Saparkhojayev and Kurymbayev 2015) and build one unique system based on RFID-technology for University. The cards that have been employed for this specific system are RFID-cards, and the algorithm used has shown stable and reliable results; moreover, the algorithm discussed in research paper (Saparkhojayev and Shaiken 2013) can be used for this system to protect important private data. There was research done in Dhanalakshmi and Mamatha (2009), which showed how to build and implement Library Management system based on RFID-technology. At the same time, other cards should be checked and be replaced, because cards, that were used, for this research seemed to be secure less, and new cards should have enough memory size so that we can keep more data inside of them. We plan to use GPS and GSM technologies in educational system, and the work that was done in Patil and Walli (2013) is an impulse for this project implementation.

References

Alarm Trade official website, http://www.alarmtrade.ru

Dhanalakshmi, M., & Mamatha, U. (2009). RFID based library management system. In *Proceedings of ASCNT*, pp. 227–234.

Domdouzis, K., Kumar, B., & Anumba, C. (2007). Radio frequency identification (RFID) applications: A brief introduction. *Advanced Engineering Informatics, 21*, 350–355.

Drickhamer, D. (2005). Toyota South Africa to use RFID for vehicle tracking. *MH&L—Material Handling & Logistics*.

NXP official website, http://www.nxp.com

Patil, S. B., & Walli, R. M. (2013). Design and development of fully automatic AT89C52 based low cost embedded system for rail tracking. *International Journal of Electronic Communication and Soft Computing Science and Engineering, 1*(1), 9–14.

Saparkhojayev, N. (2015). RFID-based staff control system (SCS) in Kazakhstan. *Journal of Physics: Conference Series, 622*(1), 012050.

Saparkhojayev, N., Dauitbayeva, A., Nurtayev, A., & Baimenshina, G. (2014). NFC-enabled access control and management system. In *2014 International Conference on Web and Open Access to Learning (ICWOAL)*, pp. 1–4.

Saparkhojayev, N., & Guvercin, S. (2012). Attendance control system based on RFID-technology. *IJCSI International Journal of Computer Science Issues, 9*(3), 1694–0814.

Saparkhojayev, N. & Kurymbayev, A. (2015). Implementation of RFID-based computer access system (CAS) for Kazakhstani University. In *Proceedings 2015 International Conference on Information Processing and Control Engineering (ICIPCE 2014)*, Russia.

Saparkhojayev, N., & Shaiken, O. (2013). An implementation of MD5 hash algorithm for RFID tags. *International Journal of Computer Applications, 72*.

SCDigest Editorial Staff. (2009). Volkswagen Says it will be First Automaker to Extensively Use RFID in Daily Supply Chain Operations. *Supply Chain Digest*.

Sharma, S., Shimi, S. L., & Chatterji, S. (2014). Radio frequency identification (RFID) based employee monitoring system (EMS). *International Journal of Current Engineering and Technology, 4*(5), 3441–3444.

StarLine official website, http://www.starline-auto.ru/catalog/avtosignalizacii_s_dvuhstoronnej_svyazyu/starline_b64.html

Swedberg, C. (2013). Nissan Europe uses RFID-enabled social media to drive consumer interest. *RFID Journal*.

Violino, B. (2005). The history of RFID technology. *RFID Journal*.

Weis, S. A., Sarma, S. E. Rivest, R. L., & Engels, D. W. (2003). Security and privacy aspects of low-cost radio frequency identification systems. In *Security in pervasive computing*, pp. 201–212.

Design and Development of a Self-adaptive, Reconfigurable and Low-Cost Robotic Arm

Kemal Oltun Evliyaoğlu and Meltem Elitaş

Abstract This study presents the design, development and control of a low-cost, self-adaptive robotic arm with the advantages of being modular and reconfigurable to perform a variety of tasks in different applications such as education, medicine and assistance for daily living activities. Particularly for educational purposes, the robot arms can be differently assembled to fulfill various tasks and its mechanical and control scenarios can be studied in the courses.

Keywords Robotic arm · Cost-efficient · Low cost · Reconfigurable · Self-adaptive · Manipulator · Robot control · Easy control with multiple DOF

1 Introduction

Today robotic technologies are rapidly growing and evolving in different aspects of our lives ranging from manufacturing to mimicking human behavior and decision-making. Therefore, the claim of *robotic technologies are influencing our evolution* might be the case. The more they involve in our lives, the more requirements we demand from these technologies such as being cost efficient, more operational, cleverer, autonomous, user-friendly, practical, precise, safe, accurate, etc. Starting 1920s, *robotic* as a word took place in the plays and stories, however today we widely hear this word in manufacturing, earth and space exploration, laboratory research, weaponry, surgery, safety, personal assistance, etc. (Alqasemi

K.O. Evliyaoğlu (✉) · M. Elitaş
Mechatronics Program, Faculty of Engineering and Natural Sciences,
Sabanci University, Üniversite St. No: 27, 34956 Tuzla/İstanbul, Turkey
e-mail: koevliyaoglu@sabanciuniv.edu

M. Elitaş
e-mail: melitas@sabanciuniv.edu

2006; Korayem et al. 2015; Deegan 2008; Kolluru 2000). Since they have a broadly diverse application area, they differ in construction regarding the design, material, fabrication process, control algorithm, and manufacturing cost. Therefore, multi-purpose designs, simple manufacturing, replacement of its links and joints for reconfiguration, low cost are essential criteria particularly for the robotic technologies that has been part of our daily lives.

Traditionally, most of the robotic arms have been developed to fulfill the demanded tasks. For each application, the most efficient and specific mechanical structure designed and constructed to perform this specific task in a most efficient manner. For instance, for a simple weight lifting operation while a 2-DOF (Degrees of Freedom) robotic arm is required, for a surgical robotic arm at least 4-limb anthropomorphic is necessary (Paik et al. 2012). Thus, different kinds of mechanical arms are being designed and produced for different applications based on task-oriented design strategies. Although these strategies are effective, it limits the possible marketing demand of the robotic arms. Recently, to overcome this limitation, self-adaptive, reconfigurable, low-cost robotic arms are being engineered (Zhang et al. 2009; Yu and Nagpal 2010; Sharma and Noel 2012) and become multi-purpose oriented.

Another leading cause that strongly influences the robotic technologies is their cost of material and processes of production, operation, maintenance, and recycling. In recent decade, 3D (three-dimensional) printing technology has obtained an important role for the manufacturing of the robots. Its simple usage, practical and fast operation and cheap material handling provide an opportunity to fabricate the CAD (Computer Aided Designs) designs for many robotic research laboratories. Consequently, researchers manufactured robots (Sik Choi et al. 2011; He et al. 2014; Johnson et al. 2015), robotic grippers (Tlegenov et al. 2014) and robot hands (Mukhtar et al. 2015; Kappassov et al. 2013) for a variety of applications such as gloves for recovery of stroke affected people (Barbulescu et al. 2015; Rogers 2009; Richardson et al. 2003), posture and steering control (Umedachi and Trimmer 2014), soft skin for human-robot interaction (Kim et al. 2015). In addition to research, 3D-printing techniques significantly served for the robotic education (Ziaeefard et al. 2015; Zhang et al. 2013; Pearce 2012). Free, open-source 3D-printable hardware for laboratories has become a powerful way of building laboratory equipment and experimental setups for the students.

In this study, we focused on the design of a low-cost, 3D-printed, reconfigurable robotic arm that can be easily modified by replacing its link sizes and joint types for the optimal mechanical structures to increase its flexibility without changing its design and manufacturing process. Thanks to its reconfigurable functional units and architecture, it can achieve the demanded diversity of different applications. Therefore, it has a great potential to find a broader market with its low cost. Likewise, practical assembly of its structure owing to compactness of its parts makes the product easier to store, transport, and replace for long-term usage. It is important to acknowledge that Ashraf Elfasakhany and his co-worker's studies shed lighting ideas for low-cost and multifunctional production of our robotic arm

(Elfasakhany et al. 2011). We intend to propose this robotic arm for educational usage. Due to its "Lego-like" structure students might practice their robotic course changing the sizes of the links, adding and removing the links, attaching cameras and sensors, calculating kinematics and developing various control algorithms. Thus, this gadget might be used as an educational toolbox for motion control and robotics orientated laboratories and projects.

In this study, the design and development of reconfigurable, low-cost, 3D-printed robotic arm is presented. Its mechanical design, customization for educational usage, forward and inverse kinematics, torque calculations and possible control strategies have been discussed in the following sections.

2 Methods

2.1 Mechanical Design

We designed our self-adaptive, reconfigurable, low-cost robotic arm using Solidworks. Our design consists of individual joints formed by pairs of parallel plates connected by two standard-sized, long or short beams, as shown in Fig. 1. Short beam joints have the actuators attached at the ends of the arms, and linked to the long beam joints via the shaft of these actuators. The joints attached together with pins and cables that are placed at the ends of each link. Straightforwardly adding or removing the links and joints efficiently changes the architecture of the robotic arm (less than a minute). Since the torque capabilities of each actuator are very important, we designed the panels in such a way that various types of actuators in terms of size and power can be supported. In the case of high torque necessity, two parallel motors can be mounted at the joints.

Besides, each panel has its own wiring on the surfaces of the plates, which has open ends as demonstrated in Fig. 2. When the links attach each other, these open ends are connected to each other and forming a specific electrical circuit. The circuit arrangement produces an electrical signal into the program that generates an output to confirm that a new link is attached to the robot.

Using Ohm's law (Eq. 1), we can calculate the voltage drop in the circuit and determine the length of the attached link. Each link has the same wiring material; hence, the voltage drop (V α R) is correlated with the length of the link (R α L) (Eqs. 1 and 2). Where V is voltage, I is current, R is resistance, ρ is resistivity, L is length, and A is the cross-sectional area.

$$V = I * R \qquad (1)$$
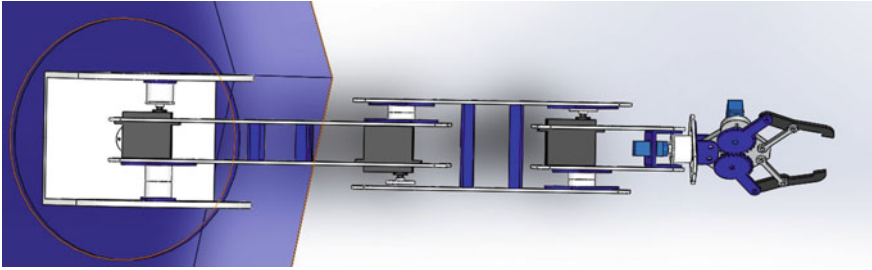
$$R = \frac{\rho L}{A} \qquad (2)$$

Fig. 1 Mechanical design of the robotic arm, *top view*. The *black square boxes* illustrate the motors at each joint while two *light blue rectangles* show the actuators for the gripper
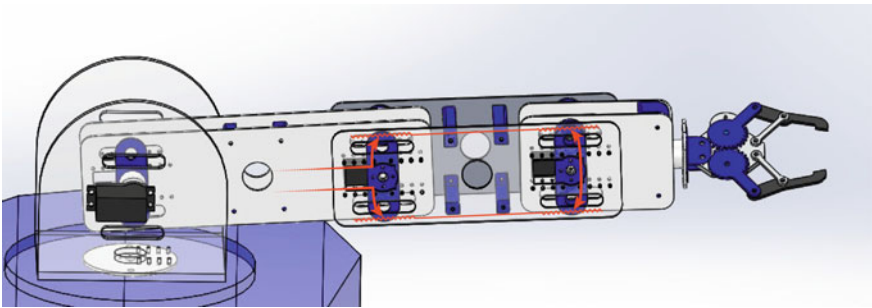

Fig. 2 Side view of the mechanical design with wiring illustration (*red lines*) to determine automatically length of the attached link

2.2 Customization

Attachment, detachment or changing the length of the links provides customization of the robotic arm. Customization process could be performed very easily by replacement of the screws at the joints, as represented in Fig. 3.

Relocating the screws at different confined spaces adjusts the length of the links. Length changes of the links can be measured by voltage drop in the circuit on the robotic arm and the program that generates kinematic equations for the robotic arm will be updated for its new link length. Figure 4 shows adjustment of the length of the links using the screws.

In our system, the kinematics and control algorithms can be calculated automatically to provide information and show the calculation errors when it is used for educational purposes. Figure 5 shows different configuration of the robotic arm when the new links are attached.
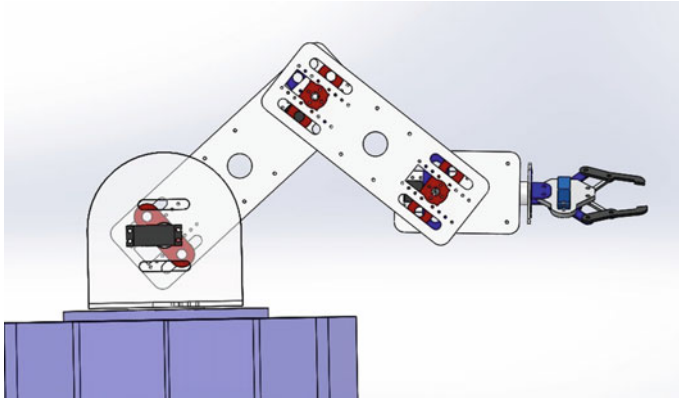
Fig. 3 Customization of the robotic arm via replacement of screws (*red*) at the joints. The length of the links can be adjusted using these *red* screws

2.3   Torque Calculations

This chapter introduces the torque calculations for the different configuration of the robotic manipulator as shown on Fig. 6. The torque values are determined based on the worst-case scenario that the L1-link is parallel to the ground. Where the arm is parallel to the Force = m * g (W1)

$$T_1 = L_1 Z_1 + \frac{1}{2} L_1 W_1 \tag{3}$$

When the second link is added, the wire is connects the open-ended circuit and it becomes continuous for the second link, and the system recognizes the attached link. The torque values can be derived using the following equations (Eq. 4).
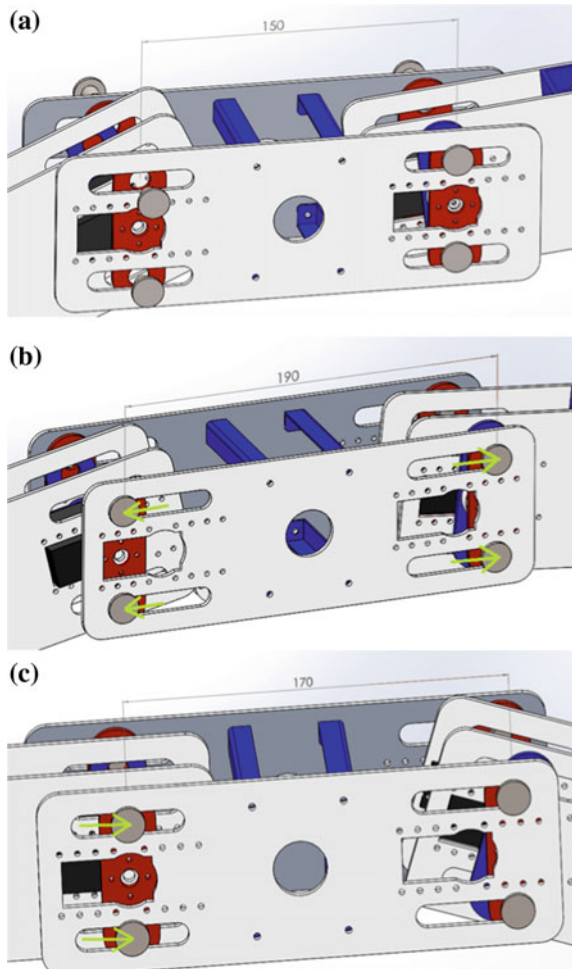
$$T_2 = (L_1 + L_2)Z_1 + \left(\frac{1}{2} L_1 + L_2\right) W_1 + L_2 Z_2 + \frac{1}{2} L_2 W_2$$

$$T_2 = T_1 + L_2 Z_1 + L_2 W_1 + L_2 Z_2 + \frac{1}{2} L_2 W_2 \tag{4}$$

$$T_2 = T_1 + L_2(Z_1 + W_1) + L_2\left(Z_2 + \frac{1}{2} W_2\right)$$

When, the torque values for 3 DOF and 4 DOF manipulators are calculated, the following equations are computed (Eq. 5).

Fig. 4 Three different
examples for adjusting the
length of the links by shifting
the joint connections. The
lengths are given in mm
(**a** 150 mm, **b** 190 mm,
c 170 mm)

$$T_3 = Z_1(L_1 + L_2 + L_3) + W_1 \left(\frac{1}{2}L_1 + L_2 + L_3 \right) + Z_2(L_2 + L_3)$$

$$+ W_2 \left(\frac{1}{2}L_2 + L_3 \right) + Z_3 L_3 + W_3 \frac{L_3}{2}$$

$$T_3 = T_2 + L_3(Z_1 + W_1) + L_3 \left(Z_3 + \frac{1}{2}W_3 \right)$$

$$T_4 = T_3 + L_4(Z_1 + W_1) + L_4 \left(Z_4 + \frac{1}{2}W_4 \right)$$

(5)
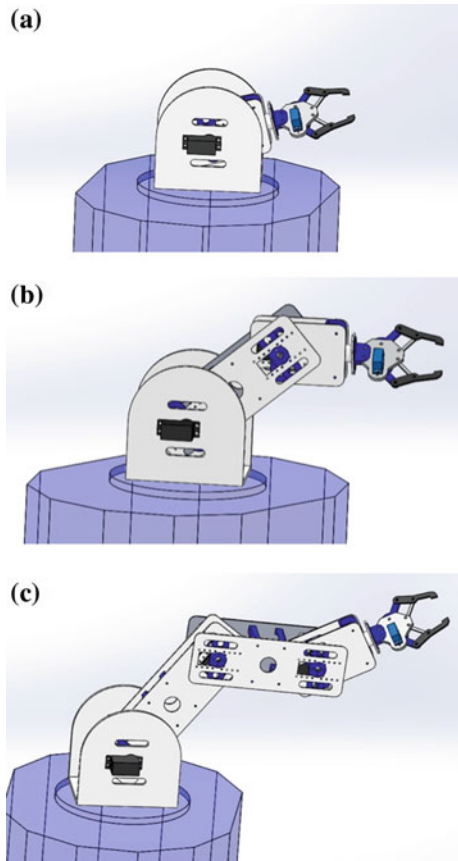
These formulas can be generalized as follows.

$$for\ i \in N\ \&\ i > 1,$$
$$T_i = L_i(Z_1 + W_1) + T_{i-1} + L_i\left(Z_i + \tfrac{1}{2}W_i\right) \tag{6}$$

Since the moment of inertia highly varies owing to the length and number of the links, in order to consider the noise (due to non-ideal actuators and joints), the calculated moment of inertia values are multiplied with a safety factor of 2.

2.4 Forward and Inverse Kinematics Calculations

The forward kinematics calculations of the 2-link assembly are presented in Eq. 7.

Fig. 6 Torque calculations for the different configurations. L is length, W is the weight, Z is the weight of the actuator
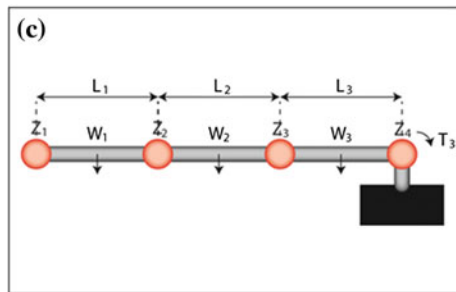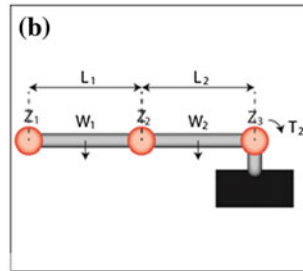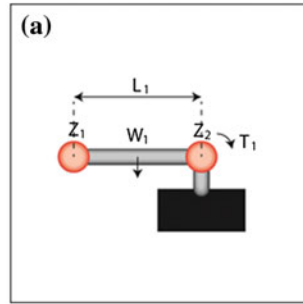
$$x = L_1 \cos(\theta_1) + L_2 \cos(\theta_1 + \theta_2)$$
$$y = L_1 \sin(\theta_1) + L_2 \sin(\theta_1 + \theta_2)$$
$$x^2 + y^2 = L_1^2 + L_2^2 + 2L_1L_2 \cos(\theta_2)$$
$$\cos(\theta_2) = \frac{x^2 + y^2 - L_1^2 - L_2^2}{2L_1L_2}$$
$$x = L_1 \cos(\theta_1) + L_2(\cos(\theta_1)\cos(\theta_2) - \sin(\theta_1)\sin(\theta_2))$$
$$x = \cos(\theta_1)(L_1 + L_2\cos(\theta_2)) - \sin(\theta_1)(L_2\sin(\theta_2)) \quad (7)$$
$$y = \cos(\theta_1)(L_2\sin(\theta_2)) + \sin(\theta_1)(L_1 + L_2\cos(\theta_2))$$
$$\cos(\theta_1) = \frac{x + \sin(\theta_1)L_2\sin(\theta_2)}{L_1 + L_2\cos(\theta_2)}$$
$$\sin(\theta_1) = \frac{(L_1 + L_2\cos(\theta_2))y - L_2\sin(\theta_2)x}{L_1^2 + L_2^2 + 2L_1L_1\cos(\theta_2)}$$

In order to achieve the specified positions of the end-effector of the 2-DOF robotic arm, the inverse kinematics are calculated as follows (Eq. 8).

$$\theta_2 = 180° - \cos^{-1}\left(\frac{L_1^2 + L_2^2 - x^2 - z^2}{2L_1L_2}\right)$$

$$\theta_1 = \tan^{-1}\frac{z}{x} + \cos^{-1}\left(\frac{L_1^2 - L_2^2 + x^2 + z^2}{2L_1\sqrt{x^2 + z^2}}\right) \quad\quad (8)$$

$$\theta_0 = \tan^{-1}\frac{y}{x}$$

Next, control algorithm will be created using the inverse kinematics. In order to control the 1-DOF robotic arm, the θ_1-angle will be determined as illustrated in Fig. 7. This configuration can be achieved when the arm is attached to a light

Fig. 7 Illustration of the end effector to the shoulder (*1 link*) assembly on the x-y plane
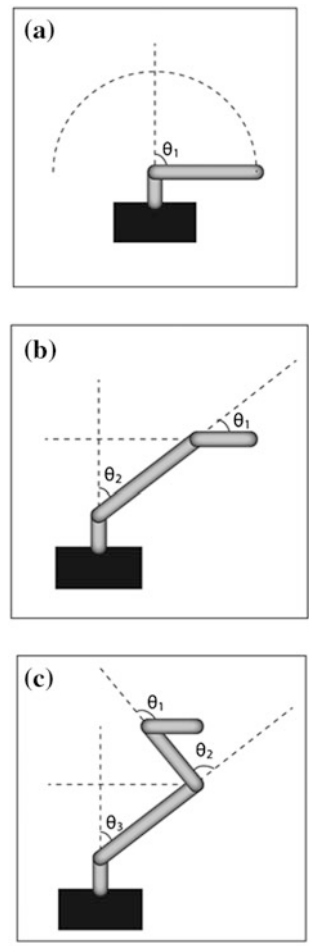
mobile base like an RC (Radio Control) car. In order to control the 2-DOF and 3-DOF manipulators, similar procedure will be applied using the inverse kinematics of the robotic manipulator.

3 Conclusions and Further Work

In this study we presented the design, basic kinematic and torque calculations for a reconfigurable, 3D-printed, low-cost robotic arm in order to be used in educational purposes. In last decade, 3D-printed, reconfigurable, self-adaptive, low-cost robotic arms have significantly contributed to robotic research and education, and it will continue in the next years. As mentioned above, these technologies do not only provide low-cost solutions, they are manufactured in a short time, their operation is simple and they are user-friendly for different applications ranging from surgical operations to daily services.

Further work for this application will be manufacturing of the robotic manipulator using the 3D-printer. Afterwards, we will develop different control algorithms for a variety of different applications. The first application will be its usage in the research laboratories to teach motion control and robotics courses. Therefore, an open-source control algorithm will be developed in addition to autonomous link-attachment detection. It will automatically calculate the forward, inverse kinematics and generated torques. Thus, students could generate different control algorithms, test them on the robot and compare their results with the automatically calculated values.

Last but not least, we will equip this low-cost robotic arm with sensors and implement robust control algorithms to make it smart, precise, more user-friendly and safe to be able to use in operation rooms of hospitals and to be able to provide eldercare at home.

References

Alqasemi, R., Edwards, K., & Dubey, R. (2006). Design, construction and control of a 7 DoF wheelchair-mounted robotic arm. In *International Conference on Intelligent Robots and Systems IEEE/RSJ*, pp. 19–24.

Barbulescu, M., Musat, A., & Popescu, D. (2015). 3D Printed Robotic Glove Useful for Recovery of. In *2015 20th International Conference on Control Systems and Computer Science (CSCS)*, pp. 833–837.

Deegan, P., Grupen, R., Hanson, A., Horrell, E., & Ou, S. (2008). Mobile manipulators for assisted living in residential settings. *Autonomous Robots, 24*, 179–192.

Elfasakhany, A., Yanez, E., & Baylon, K. (2011). Design and development of a competitive low-cost robot arm with four degrees of freedom. *Modern Mechanical Engineering, 1*, 47–55.

He, Y., Guo, S., & Shi, L. (2014). 3D printing technology-based an amphibious spherical underwater robot. In *Proceedings of 2014 IEEE International Conference on Mechatronics and Automation*. IEEE, (s. 1382–1387). Tianjin.

Johnson, B., Cole, B., & Cappelleri, D. (2015). 3D Printed surgical manipulator for minimally invasive lumbar discectomy surgery. In *ASME Additive Manufacturing + 3D Printing Conference & Expo (AM3D)*. Boston, MA USA.

Kappassov, Z., Khassanov, Y., Saudabayev, A., & Shintemirov, A. (2013). Semi-anthropomorphic 3D printed multigrasp hand for industrial and service robots. In *2013 IEEE International Conference on Mechatronics and Automation (ICMA)*, pp. 1697–1702.

Kim, J., Alspach, A., & Yamane, K. (2015). 3D Printed soft skin for safe human-robot interaction. IEEE/RSJ IROS.

Kolluru, R., Valavanis, K., Smith, S., & Tsourveloudis, N. (2000). Design fundamentals of a reconfigurable robotic gripper system. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, 30*(2), 181–187.

Korayem, M., Esfeden, R. A., & Nekoo, S. R. (2015). Path planning algorithm in wheeled mobile manipulators based on motion of arms. *Journal of Mechanical Science and Technology*, 29, 1753–1763.

Mukhtar, M., Akyurek, E., Kalganova, T., & Lesne, N. (2015). Control of 3D printed ambidextrous robot hand actuated by pneumatic artificial muscles. In *SAI Intelligent Systems Conference (IntelliSys)*, pp. 290–300.

Paik, J. K., Shin, B., Bang, Y. -B., & Shim, Y. -B. (2012). Development of an anthropomorphic robotic arm and hand for interactive humanoids. *Journal of Bionic Engineering, 9*, 133–142.

Pearce, J. M. (2012). Building research equipment with free, open-source hardware. *Science, 337*, 1303–1304.

Richardson, R., Brown, M., Bhakta, B., & Levesley, M. (2003). Design and control of a three degree of freedom pneumatic physiotherapy robot. *Robotica, 21*, 589–604.

Rogers, J. R. (2009). Low-cost teleoperable robotic arm. *Mechatronics, 19*, 774–779.

Sharma, A., & Noel, M. M. (2012). Design of a low-cost five-finger anthropomorphic robotic arm with nine degrees of freedom. *Robotics and Computer-Integrated Manufacturing, 28*, 551–558.

Sik Choi, H., Na, W., & Kang, D. (2011). A humanoid robot capable of carrying heavy objects. *Robotica, 29*, 667–681.

Tlegenov, Y., Telegenov, K., & Shintemirov, A. (2014). An open-source 3D printed underactuated robotic gripper. In *2014 IEEE/ASME 10th International Conference on Mechatronic and Embedded Systems and Applications (MESA)*, pp. 1–6.

Umedachi, T., & Trimmer, B. (2014). Design of a 3D-printed soft robot with posture and steering control. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2874–2879.

Yu, C. -H., & Nagpal, R. (2010). A self-adaptive framework for modular robots in dynamic environment: Theory and applications. *The International Journal of Robotics Research,* doi:10.1177/0278364910384753.

Zhang, C., Anzalone, N. C., Faria, R. P., & Pearce, J. M. (2013). Open-source 3D-printable optics equipment. *PLoS ONE, 8*, e59840.

Zhang, W., Che, D., Liu, H., Ma, X., Chen, Q., Du, D., & Sun, Z. (2009). Super under-actuated multi-fingered mechanical hand with modular self-adaptive gear-rack mechanism. *Industrial Robot Journal, 36*, 255–262.

Ziaeefard, S., Ribeiro, G., & Mahmoudian, N. (2015). GUPPIE, underwater 3D printed robot a game changer in control design education. In *American Control Conference (ACC)*, pp. 2789–2794.

Workplace Emotion Monitoring—An Emotion-Oriented System Hidden Behind a Receptionist Robot

Paulo Gurgel Pinheiro, Josue J.G. Ramos, Vander L. Donizete, Pedro Picanço and Gustavo H. De Oliveira

Abstract Systems capable of recognizing and analyzing humans emotions can play an important role in many domains, especially for commercial or productivity purposes. Emotion-oriented systems are widely used in commercial area, but there is still a lack of them being applied as automatic tool in companies. One challenge is to highlight the impact these systems might have on employee satisfaction due to privacy issues and the sense of being monitored. In this work we propose a workplace emotion monitoring system capable of recognizing workers emotions, quantifying them and to provide data for productivity analysis. The system is able to capture the average office humor without using individual and private information. To validate the solution, the system was set up at a R&D Center during 10 workdays. The system was able to capture the emotions providing data for worker's productivity and job satisfaction analysis.

Keywords Emotion-oriented system · Facial expressions · Emotions · Monitoring workplace · Robotics · Human–robot interaction

1 Introduction

Employee Monitoring, the act of monitoring employee during their activities on the organizations, is an important analysis tool to track and increase productivity and behaviors. This monitoring has been the subject of studies but and the can be used for safety purposes using biometric sensors or CCTV systems (TopReview 2015),

P.G. Pinheiro (✉) · J.J.G. Ramos · V.L. Donizete · P. Picanço · G.H. De Oliveira
Robotics and Vision Computing Division of CTI, Rod D Pedro I Km 143, 6,
Campinas, SP 13069-901, Brazil
e-mail: paulo.pinheiro@cti.gov.br

J.J.G. Ramos
e-mail: josue.ramos@cti.gov.br

or a tool to improve the productivity as by improving the work environment quality by using questionnaires as in Houston Chronicle (2015) due to its important for the productivity and to its thin line between impersonal monitoring and invasion of privacy. Monitoring techniques may use personal information for analysis, such as the person's name or his/her registration, but others can be indifferent highlighting the overview of the corporation without going into the personal matters.

Systems capable of detecting human emotion can be useful in many areas. Monitoring employees in workplaces as in a call center might allows the manager to make some probabilistic statement about the relationship between emotion (e.g. anger) and problem behavior (e.g. rudeness) (Bickmore and Picard 2005). Also they can be applied in companies to point out the relationship between the average emotion of workers (e.g. enthusiasm) to weekdays (e.g. days which the daily meal of the company's restaurant is great) or to important events that might happen in the company (e.g. some collective promotion).

In this work, we propose an emotion-oriented system capable of monitoring the emotions at workplace without considering an individual particularly, but all employees in general. Here, the goal is to study the relationship between the average of the emotions of employees and hours of the day and the day of week. The proposed system claims to be a general tool to be used for others pattern recognition system that has to consider the emotion state at workplaces as an input.

The use of them in workplaces may impact on employee satisfaction due to privacy issues and the sense of being monitored. That's why the first challenge is to define what technology to use for data acquisition and the second one is how this technology will be inserted into the environment to be as natural as possible without the employees feeling their privacy are being invaded.

As the vector of the system, we will use ANA, the receptionist virtual robot that is already known to employees and is already blended in the environment, not featuring an external or new interference. Figure 1 shows the user interacting with ANA while the robot captures his emotions in background.

The proposed system aims to be a tool capable of recognizing workers emotions, quantifying them and to provide data for elaborate analysis. To validate the solution, the study was held in a group of R&D in Brazil during the first 10 weekdays of November 2015. The system captured the emotions of workers, linking the emotion average rate to each hour of the day.

2 Background—Emotion Analyzer Systems

According to Heylen et al. (2011) the emotion-oriented systems can be classified into tree main groups: systems that can recognize emotions, stimulate human emotion and express emotions. Although the proposed system is able to express emotions (e.g. ANA, the robot, can perform dozens of facial expressions) and stimulate human emotions (ANA might pop up to the user some controversial or

Fig. 1 ANA, the receptionist robot, placed at the corridor of department, can capture and analyze the emotion of the worker in real-time

biased news), the aim of this work relies on the first group, a system to recognize and analyze human emotions.

Unlike people, machines can read facial expressions and from there infer emotions much more quickly and for long periods of time with almost no limitations on the amount of data stored. These emotion-oriented systems can isolate the environmental context and focus only on the emotions of the person, without getting involved by external influences.

For this reason, these systems have been used for commercial and productivity purposes. In Lewinski et al. (2001), users' facial expressions of happiness were automatically analyzed by FaceReader algorithm distinguishing video advertisements between two categories, amusing and non-amusing. They demonstrated, using participants' self-reports and the results from algorithm that advertisement effectiveness of amusing video ads relates to happiness, and not with any other basic emotion.

During the 2015 Super Bowl, Emotient Analytics (Mone 2015) identified and tracked the faces of 30 volunteers at a local bar while they watch the game. The system could identify the changing emotional state of each individual over time whether they were amused, ambivalent, or surprised. The system matched the emotional states to what was happening on screen.

Systems to recognize and analyze human emotions are widely used in the commercial (advertising) area, but there is still a lack of these models being applied as automatic methods in companies. In Schiopu (2015), the author describes the effects of emotions (positive and negative) felt by employees on job satisfaction. The emotions felt have direct effect on employee job satisfaction and intention to work in the same company. This system could be automated to operate for long periods reaching more people in the workplace. An automatic tool could also

benefit studies covering long periods of time, such as in Staw et al. (1994) who analyzed 272 employees over 18 months to study the link between positive emotion and favorable outcomes at the workplace.

3   Proposed System

Based on the requirements posed on the previous section, the proposed system should be able to capture the average humor without considering any privacy issue, which is using individual personal information. In other words, the system must be able to extract the sentiment information of a person and use it to calculate the average humor without making use of his/her identity or history of interactions with the system. Figure 2 shows the system architecture and its components.

Interaction Manager (IM) is responsible for the human-robot interactions and receptionist services of the robot. We will not go into the details of IM, since this work focus on the processes of capturing emotions, persistence, analysis and data presentation. After the system captures the user's emotions, it analyzes and categorizes them in real time, as we will describe later. The system also saves the emotional information for possible future analysis. The information about the current emotion and the average of emotions feeds the Interaction Model that might consider them to change the interaction behavior (e.g. if the person who is interacting is sad, the robot can avoid to perform overly happy expressions). Finally, a graphical interface shows the partial results. We will cover details about each one of these modules in this section.
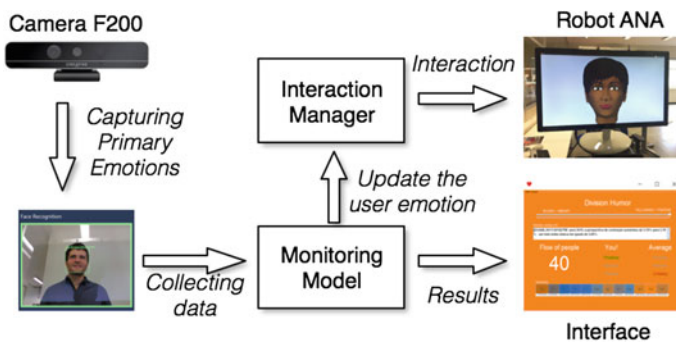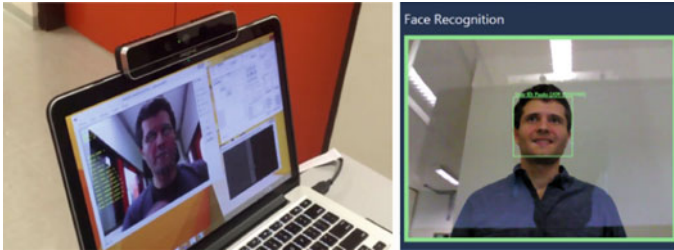

Fig. 2 Architecture overview

Fig. 3 Camera Intel F200, and face and emotion recognition screen

3.1 Capturing Emotions

To capture emotions of employee we use an Intel RealSense 3D Camera (Realsense 2015). This is a stand-alone camera that uses depth-sensing technology and can be attached to a desktop or laptop computer or it can be included in laptops and tablets. It consists of a conventional camera, an infrared laser projector, an infrared camera, and a microphone array. The camera features facial analysis, tracking 78 points of the face, inferring emotions and sentiments. Figure 3 shows the camera and the emotion capture in real time.

The algorithm supports 6 primary emotions: Anger, Disgust, Fear, Joy, Sadness and Surprise. The proposed system applies a filter to extract the basic sentiments of these emotions, categorizing them into: negative, positive or neutral. Figure 4 shows on its the first, neutral expressions, and on the other lines, all kind of positive sentiment expressions. Figure 5 also shows a set of neutral expressions on the first line and expressions recognized as negative sentiment on the following lines.

3.2 Collecting Data of Human–Machine Interaction

Data collection has been a challenge when involving non-verbal parts, such as emotion in conversation (Bickmore and Picard 2005). Many applications require video recordings, transcription and analysis, that are in general time consuming. In this work we will use the emotion data not as an input to be treated off-line, rather, the capture and storage of emotions will be performed in real-time.

To make this possible, the system will use the concept of timeline to organize a list of events in chronological order. Each interaction between the user and the robot creates an event (slot). Each slot is saved on the timeline at the point corresponding to the time it happened. Emotions extracted from the user, during that interaction, are inserted in the correspondent slot. When the user goes away, the slot is close and data saved. To retrieve the information, the system loops back through the timeline structure until the correspondent date pushing the data.

Fig. 4 Neutral faces on the first line followed by example of positive sentiment faces
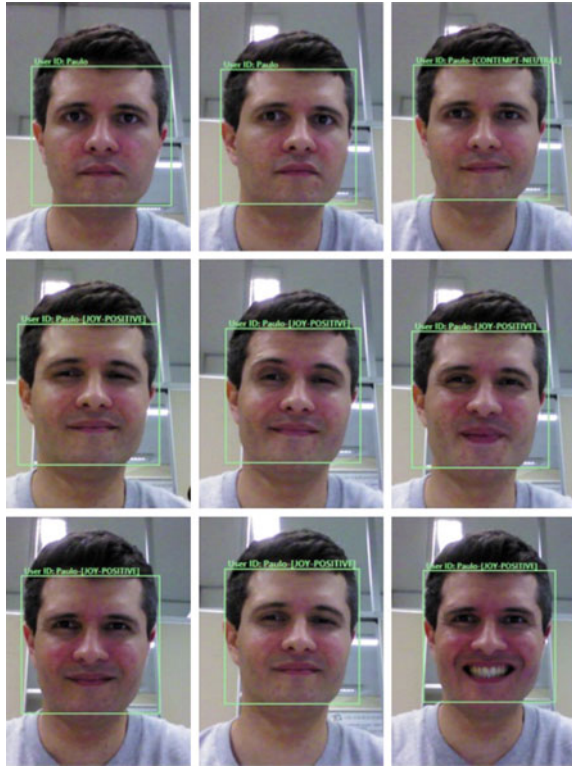

3.3　Monitoring Strategy

As a monitoring emotion strategy, the system will categorize the emotions into three major groups: neutral, positive and negative sentiment. Since the objective is not to get a fine grained emotion average and since the system will be used for a long term (days, weeks), we believe this classification is enough to extract the information we are interested in. As a proof of concept the proposed system will use a score based algorithm to quantify humor/emotion. Other metric algorithms can be considered in the future depending on the application.

In case of using the score based algorithm, at the very first beginning of the day, the algorithm defines a value of 0 to workplace humor for every hour. A negative humor interaction is defined when the emotions expressed by user during the interaction has a negative balance. Let's suppose the user has interacted for 20 s with the robot that captures 10 expressions from user: NEUTRAL, NEUTRAL, POSITIVE, POSITIVE, NEUTRAL, NEUTRAL, NEGATIVE, NEGATIVE, NEGATIVE, NEGATIVE. The balance of this interaction is −2, featuring a negative humor interaction. This interaction will count −2 on the global value of the humor for that hour of day. The opposite applies to a positive balance. Neutral

Fig. 5 Neutral faces on the first line followed by examples of negative sentiment faces
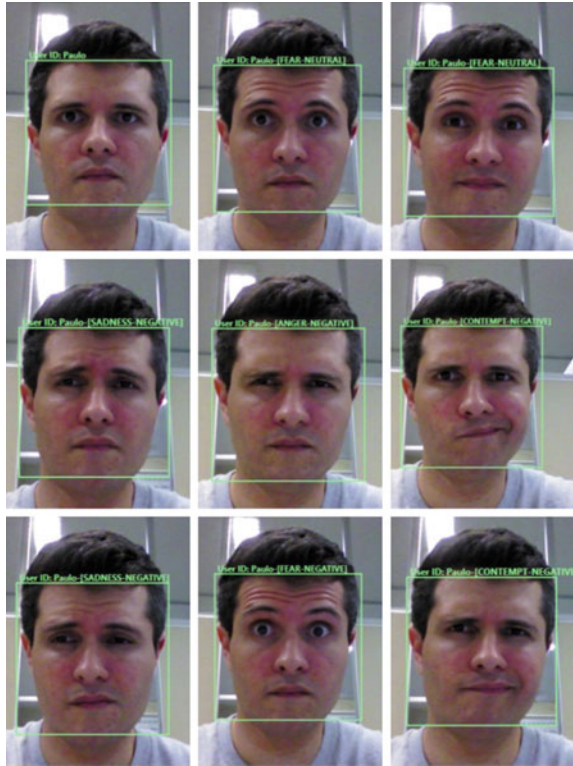
humor interaction means the balance of negative and positive emotions is zero. There is no minimum time to define an interaction. An interaction will last the time the user spends in front of the robot.

The score-based algorithm runs in real-time. However, others algorithms, based or not on score, or even processed offline can be used with the proposed platform. As described in the Sect. 3.2, the timeline saves all the emotions extracted from the interactions. Also, since the timeline structure can be accessed in real-time, other algorithms can replace the proposed one here to meet each domain's needs.

3.4 Graphic Interface

Figure 6 shows the scale of workers' emotion interface. At the top of interface, a range slider shows the current sentiment of the interaction. Positive expressions take the cursor to the right, and negative ones, to the left. Right below, we have a

text area with some daily news, the people flow information (number of people who went next to the robot), the average humor of the user on the current interaction, and the average of department's humor perceived so far. In the lower part, the interface shows the humor average score for each hour of day.

To highlight the dynamic nature of the variation of humor, we define a visual feedback using colors for positive and negative interactions. The variation of colors is used as a non-verbal communication, besides facilitating the representation of emotions along the day. Figure 6 shows the background variation color.

As the non-verbal communication, the proposed system sets two rule color marks, a strong yellow/orange to represent the state of being more positive and a strong blue to represent the more negative humor state. According to Zajonc (1976), yellow is able to represent positive emotions, since it is the color nearest to the light, carrying the nature of brightness and having a serene and softly exciting character over people. In the same study, the blue was described as the color having a peculiar and powerful effect on the eye, but on the negative side. Its appearance would represent a contradiction between excitement and repose. In the pop culture, the yellow color is easily used to represent positive things (e.g. yellow smile face), and the blue due its informal meaning of being associated with melancholy, sad, or depression. Figure 6 also shows the tracker of emotions per hour of day.
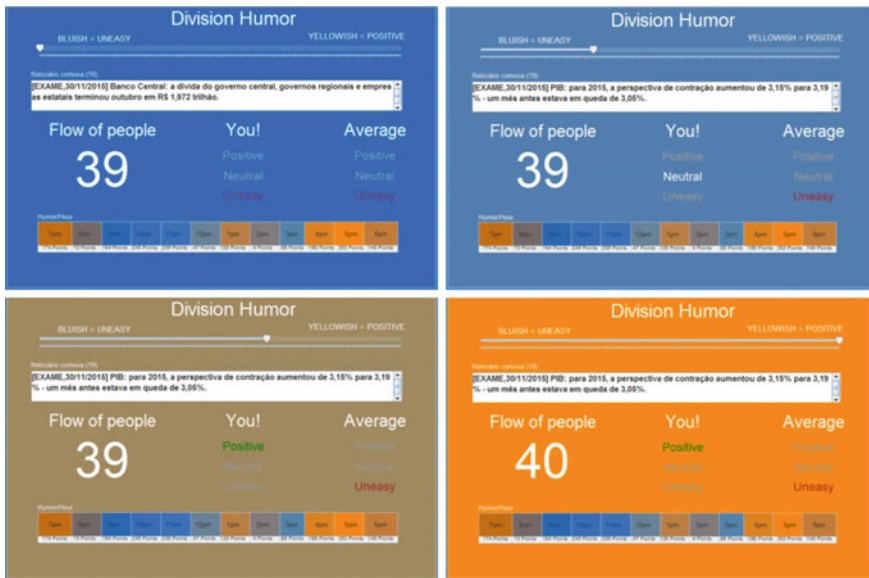

Fig. 6 Interface showing at the *top*, the scale of humor of the interaction (range slider), at the *bottom*, the daily emotion, and as the background color, the corresponding emotion rate, ranging from extreme *blue* (sadness) to extreme *yellow/orange* (happiness)

Fig. 7 ANA, the receptionist robot is used as a vector to capture workplace emotions decreasing the feeling of invasion of privacy of the workers
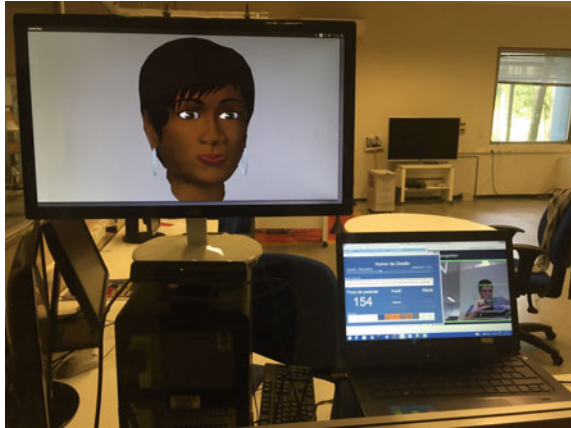

4 ANA, the Receptionist Robot

ANA is a receptionist robot designed to interact with people using verbal and nonverbal communication (Trovato et al. 2015). ANA plays the receptionist role of the department, receiving passersby and giving them information about general questions, people extension number, schedules, and delivering personal messages. This robot version is built upon the receptionist robot originally developed by Prof. Reid Simmons's team at Carnegie Mellon University (Kirby et al. 2005). ANA is already well known by employees and it is already part of the environment, being a natural and camouflaged vector for our work. It decreases the users feeling about invasion of privacy or of being monitored by a system. Figure 7 shows the default setup of ANA.

In addition to the receptionist services, ANA presents daily news and the weather forecast. Such information attracts the attention of the workers increasing the number of interactions. These attractions are carefully chosen to be neutral and not to influence the employee's humor.

5 Experiments

A conceptual study was performed with the purpose of demonstrating the practical use of the developed platform and its ability to recognize user emotions, to quantify them and to provide data for potential analysis. To demonstrate the platform, the robot was setup on the corridor at the department collecting emotions of the employee from 8 AM to 6 PM during the first 10 weekdays of the month of November. Even if the worker did not interact with the robot, but his/her face was detected, the expressions were captured and considered. The system calculates the emotion score for each hour based on the score-based algorithm described in

Fig. 8 Score per hour on weekdays
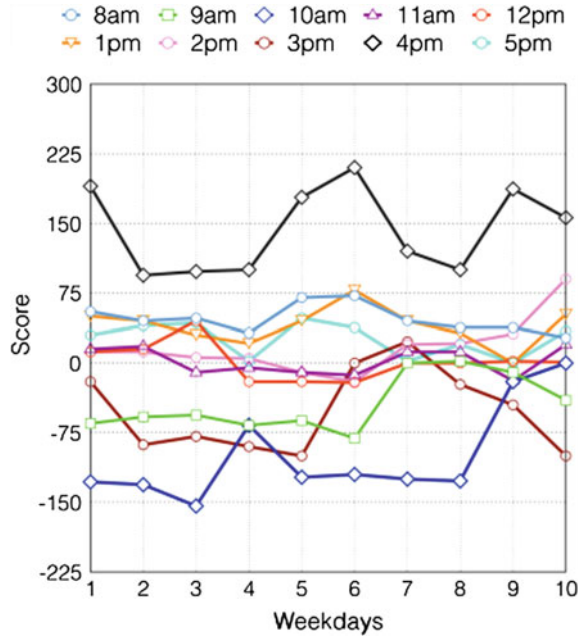
Sect. 3.3. The study was held at the Division of Robotic and Visual Computing in CTI (Centro de Tecnologia da Informação Renato Archer) in Campinas, Brazil.

5.1 Results

Figure 8 shows 10 weekdays and the score average for each hour of the day. The strokes on the lines (circle, square, diamond) has no meaning besides making the graph more readable.

According to the graphic, we can notice that in most of the hours (8 AM, 11 AM, 12 PM, 1 PM, 2 PM, 5 PM) the workplace sentiment tends to the neutrality, ranging from −21 points (12 PM on Day 4) to +90 points (2 PM on Day 10). The line corresponding to 9 AM remained negative during the first six days after tending to neutrality in the last four days. The 4 PM line stands out, having a positive score reaching the peak every day, while the 10 AM line reaches the lowest values for 7 out the 10 days.

For a closer look, Fig. 9 insulates the morning hours and their score, while Fig. 10 insulates the afternoon hours. The 8 AM line appears as the morning time of day which the sentiment is more positive, followed by 11 AM. The 9 AM and 10 AM lines feature more negative score.

Figure 11 shows the average sentiment of each day along the 10 days. The study comprises two weeks, starting from Monday. On the second week, the daily score

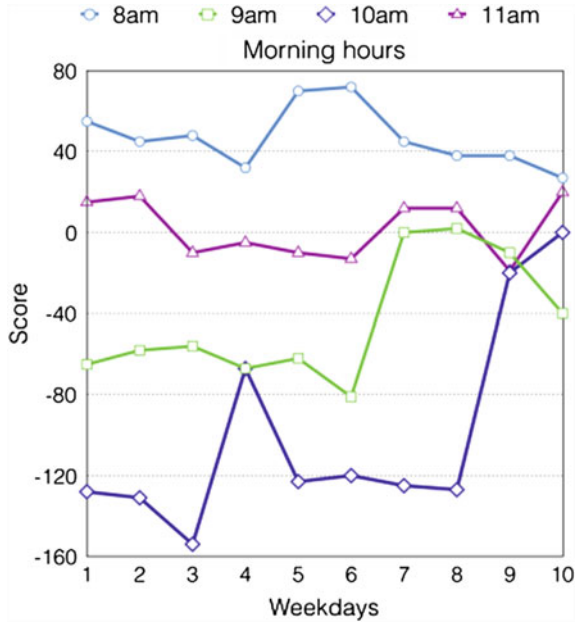Fig. 9 Sentiment on the mornings during the 10 weekdays

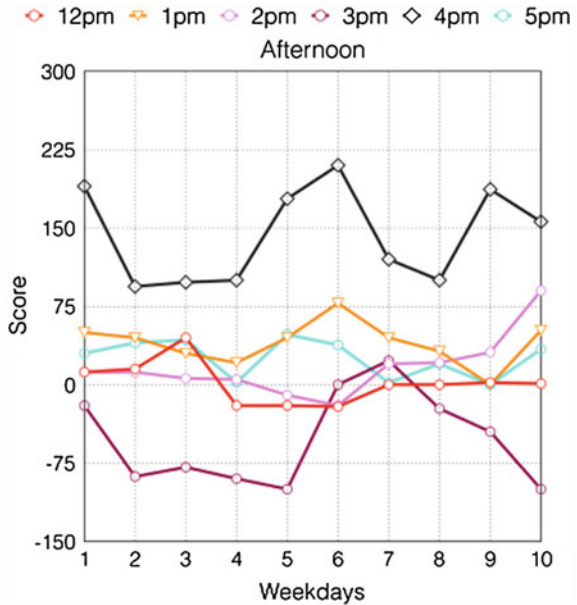Fig. 10 Sentiment on the afternoon during the 10 weekdays

Fig. 11 Sentiment average of each day
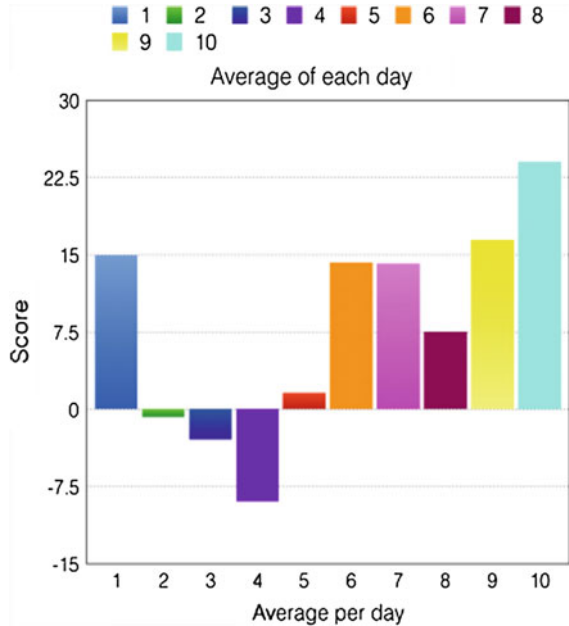
average was positive for every day. Unlike the first week, which on average, only 2 days were positive.

5.2 Discussion

Based on the experiments it was possible to extract spot information about the emotional state of the people in two weeks of work:

- 10 AM: the hour of the day which people had performed more negative emotions.
- 4 PM: the hour of the day people were more positive.
- 3 PM: the hour of the first week, which the system captured more negative emotions.
- 11 AM and 12 PM: the hours featuring more neutral sentiment.
- On the first week, only Monday and Friday were positive days.
- On the second week, 5 out of 5 days had positive score average.

These are some patterns that catch the eye when looking at the charts, but due to the small universe is not possible to state, for example, that all Mondays, regardless the weeks, tend to have a similar positive average, or the sentiment at 11 AM would always be neutral.

Although the small amount of data, some assumptions can be made, for example, since the work shift ends at 5 o'clock, it would not be a surprise to expect that in the last work hour of day (4:00 PM–4:59 PM) the sentiment would be much more positive (average of 143 points) than the sentiment in the first hour of the day (8:00 AM–8:59 AM), with average of 47 points.

Another assumption is about the difference sentiment between first week and the second one. On average, the first week had 3 negative days out of 5 days, while the second week had none. This research did not consider events that may have happened on the workplace and able to influence the emotions of the workers. Project approvals, important meetings, weather or even the restaurant day menu are some events that may be able to change the sentiment of the people.

6   Conclusion

In this work, we proposed an emotion-oriented system capable of recognizing and analyzing workers' emotions at workplace, linking emotion average score to hours of day. The use of emotion monitoring system at workplaces has the challenge to overcome the feeling the workers have of being monitored. To bypass this issue, the proposed emotion-oriented system had as a vector the robot ANA. ANA is a receptionist robot that is already known to employees and is already blended in the environment, not featuring an external interference.

During two weeks at the Division of Robotic and Visual Computing, in CTI, the robot captured, as a background process, the worker's emotions, categorizing emotions into three main groups, positive, negative and neutral. The system could calculate the average of sentiment for each hour of day. This data was used to realize what hours of day the worker's were more positive, negative or neutral. The same was performed for weekdays and weeks. As a proof of concept, the system has used a score-based algorithm running in real-time. Other algorithms can be integrated into the proposed platform to meet specific domain's needs though. The system was able to run day and night capturing facial expressions whenever the worker's face was detected, even if s/he does not interact with the robot. The system has proven to be robust, running day and night even after the end of the day shift.

The presented research is a first step for future ones. Here, it was used a two-week working window to validate the model. The next step is to use the presented model to capture emotions for a period of, at least, two months and then cross the emotion information with the department calendar events. We expect to see if there is any relationship between the emotions of the people in the workplace with the recorded events of the department's agenda. Other front line is to refine the three basic sentiments used (positive, neutral and negative) into more precise sentiments like nervousness, stress, restlessness and relaxation. Also, this work did not address issues about data security. As a future work, solutions should be developed to minimize the risks of infringing emotional privacy.

References

Bickmore, T., & Picard, R. (2005). Establishing and maintaining long-term human–computer relationships. *ACM Transaction Computer Human Interaction, 59*, 21–30.

Heylen, D., Bevacqua, E., Pelachaud, C., Poggi, I., Gratch, J., & Schröder, M. (2011). Generating listening behavior. In *Emotion-oriented systems* (pp. 321–347). Springer.

Houston Chronicle. (2015). *How to monitor employee motivation, satisfaction & performance*. http://bit.ly/emp_satif. Last visit: November 2015 [Online].

Kirby, R., Broz, F., Forlizzi, J., Michalowski, M. P., Mundell, A., Rosenthal, S., et al. (2005). Designing robots for long-term social interaction. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 2199–2204).

Lewinski, P., den Uyl, T. M., & Butler, C. (2001). Automated facial coding: Validation of basic emotions and FACS AUs in FaceReader. *Journal of Neuroscience, Psychology, and Economics*, 7(4), 227–236.

Mone, G. (2015). Sensing emotions. *Communications of the ACM*, 58(9), 15–16.

RealSense. (2015). *Intel RealSense TM 3D Camera* (*F200*). https://software.intel.com/en-us/realsense/f200camera. Last visit: November 2015 [Online].

Schiopu, A. F. (2015). Workplace emotions and job satisfaction. *International Journal of Economic Practices and Theories*, 5(3), 277–282.

Staw, B. M., Sutton, R. I., & Pelled, L. H. (1994). Employee positive emotion and favorable outcomes at the workplace. *Organization Science*, 5(1), 51–71.

TopReview. (2015). *Employee monitoring software reviews*. In http://employee-monitoring-software-review.toptenreviews.com/. Last visit: December 2015 [Online].

Trovato, G., Ramos, J. G., Azevedo, H., & Moroni, A., Magossi, S., Ishii, I., et al. (2015). Olá, my name is Ana: A study on Brazilians interacting with a receptionist robot. In *17th International Conference on Advanced Robotics, ICAR* (pp. 66–71).

Zajonc, A. G. (1976). Goethes' theory of color and scientific intuition. *American Journal of Physics*, 44(4), 327–333.

Optimum Control for the Vehicle Semi-active Suspension System

Ayush Garg, Akshay Arvind and Bhargav Gadhvi

Abstract This paper presents a novel approach in optimization of semi-active suspension system based on ride comfort and road handling characteristics. Semi-active suspension is capable of providing both ride comfort and road handling of the vehicle by optimization of various parameters. The model used for study is quarter car semi-active model. The entire analysis part of the paper has been carried out using Heat Transfer Search and Teaching-Learning based optimization algorithm. It also throws some light on comparison of semi-active and passive suspension based on their performance on similar road conditions by plotting graph of linear acceleration of sprung mass with respect to time.

Keywords Suspension · Passive · Active · Semi-active · Optimization · Heat transfer search (HTS) algorithm · Teaching-learning based optimization (TLBO)

1 Introduction

Today, the automotive industry is growing rapidly not only in mechanical engineering but also in every other field of engineering. Suspension is considered as one of the key system in vehicles and has been highly investigated and researched in a long time due to their contribution in vehicle ride and it's handling (Shiao 2010).

Suspension system is a part of vehicle assembly consisting of tires, springs, linkages and shock absorbers which connect the wheels to the vehicle and provide relative motion for the vehicle to move (Alexandru and Alexandru 2011). The main purpose of suspension system is to absorb the vibrations due to road disturbances so as to provide ride comfort to the passengers as well as allow good road handling or

A. Garg (✉) · A. Arvind · B. Gadhvi
Department of Mechanical Engineering, Pandit Deendayal Petroleum University,
Gandhinagar, Gujarat, India
e-mail: ayushgarg1511@gmail.com

holding and overall vehicle stability so that it does not deviate from its path of motion (Agharkakli et al. 2012). A good suspension system should provide good vibration isolation, i.e. small acceleration of the body mass and a small "rattle space", which is defined as the maximal allowable relative displacement between the vehicle body and various suspension components (Mohan Rao et al. 2010).

The objective of a vehicular suspension are, (a) to carry the weight of the vehicle body and provide comfortable and safe rides over many types of road surfaces, (b) to isolate the body from the external and internal disturbances, (c) to keep a firm contact between the road and the tires ensuring good vehicle handling (Mulla 2013). To achieve these objectives, lot of development has been taken place in order to incorporate proper tradeoff between the performance parameters.

Active suspension systems have been continuously examined and investigated since a long time. These suspensions are found to be quite successful but consumes very high power and failure in this system may result in handling problems (Gowda and Chakrasali 2014). Semi-active suspension system requires considerably less power than the active suspension. That's why a lot of attention is been given to semi-active suspension to achieve a favorable compromise between ride control and vehicle handling (Alexandru and Alexandru 2011).

This paper deals with analysis of quarter car model as it represents the most basic features of one fourth of a vehicle. The investigation of quarter car model can be done by control of spring stiffness or damping coefficient. This paper deals with use of damping coefficient as control parameter for minimization of acceleration of sprung mass.

In this paper efforts have been put into minimize the vertical acceleration of sprung mass by using Heat transfer search (HTS) and Teaching-Learning based optimization (TLBO) algorithm. These newly developed algorithms has been found to maintain a good balance between exploration and exploitation. The entire optimization process is carried out in MATLAB using ODE45 as functional tool to solve the equations. In the concluding section of the paper, the variation of vertical acceleration of semi active suspension and passive suspension has been plotted with respect to time and the results have been compared.

2 Types of Vehicle Suspension

2.1 Passive Suspension

It is one of the most basic and elementary type of suspension in which the parameters are generally fixed being chosen to achieve certain level of compromise between road handling and ride comfort (Mulla 2013). This type of suspension consists of an energy dissipating element, which is the damper, and an energy-storing element, which is the spring. Since the parameters can't be changed,

the performance of passive suspension depends on the type of road profile and the values of these parameters (Gadhvi and Savsani 2014).

The limitation of passive suspension is, it produces heavily damped or too hard suspension it will transfer a lot of road input or throwing the car on unevenness on the road. Again, if it is lightly damped or soft suspension it will reduce the stability of vehicle in turns and will swing the car (Agharkakli et al. 2012).

Compared to passive, semi-active suspension offers a trade-off between ride comfort and road handling. Due to these serious limitations, the passive suspensions are not used nowadays generally.

2.2 Active Suspension

An active suspension system has the ability to store, dissipate and to introduce energy to the system. It can change its parameters depending upon the operating conditions (Gandhi et al. 2015). Hence it provides a good compromise between ride comfort and road handling.

The special part of active suspension is that it consists of a force actuator which is controlled by a controller as shown in Fig. 1. The controller will receive the data from sensors (Patil et al. 2014). The force actuator will calculate either add or dissipate energy from the system, with the help of sensors as input to reduce the acceleration of sprung mass or rattle space (Mohan Rao et al. 2010).

However, in spite of so many advantages, the main limitations of active suspension are that it consumes a lot of power and the degree of complexity is high.

Fig. 1 Quarter car model of active suspension (Segla and Reich 2007)
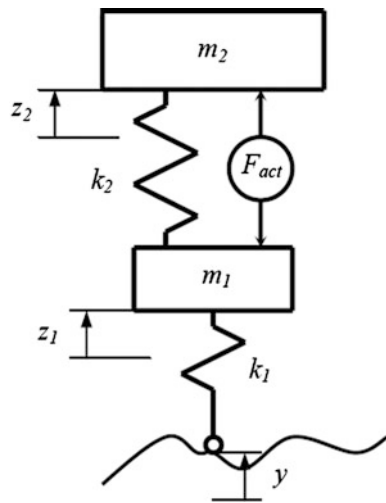
In comparison to the active suspension, a semi-active suspension system requires less power, is less complex and more reliable and can provide considerable improvement in vehicle ride quality.

2.3 Semi-active Suspension

The semi-active suspension system is the latest type of suspension available which utilizes an adjustable damper in the place of force actuator. In this type of system, the conventional spring element is retained, but the damper is replaced with a controllable damper as shown in Fig. 2 (Segla and Reich 2007). In semi-active suspension system, the controller determines the level of damping based on a control strategy, and automatically adjusts the damper to achieve that damping (Gadhvi and Savsani 2014). It can offer a compromise between the simplicity of passive systems and cost of higher performance fully active suspension system.

Due to these advantages, semi-active suspension systems are getting more attention in the development of suspension system. Thus, major focus is given to the optimized control for the semi-active suspension system in this paper.

3 Heat Transfer Search Algorithm

There have been many metaheuristic optimization algorithms developed to solve a wide variety of real life problems. All these algorithms are population-based where a group of solutions carry out the search process.

Fig. 2 Quarter car semi-active model (Segla and Reich 2007)
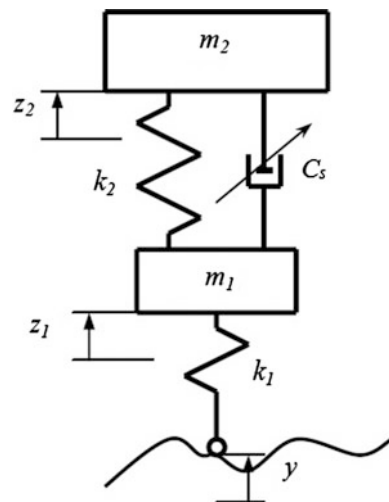
However, for any algorithm, Exploration and Exploitation are the two most important parts that governs the success rate of that algorithm (Patel and Savasni 2015). Exploration is the process of abrupt movement in the search space to cover it entirely while exploitation is the process to refine certain areas of explored search space. Pure exploration enhances the capacity of any algorithm to produce new solutions with less precision while pure exploitation increases the possibility of trapping at the local optimum solution during the search process (Patel and Savasni 2015). Therefore, a perfect balance between these two conflicting parameters is always required for the better performance of any algorithm. Keeping these things in mind, the HTS algorithm is used in this paper to incorporate proper trade-off between exploration and exploitation (Patel and Savasni 2015).

HTS algorithm inspired from natural law of thermodynamics which states that "Any system always try to achieve equilibrium state with its surroundings" (Patel and Savasni 2015). If any system at higher/lower temperature interacts with the surrounding maintained at lower/higher temperature, the system always try to achieve temperature level of the surrounding by heat transfer to establish thermal equilibrium. In this way, the system always tries to reduce the thermal imbalance that exists between the system and surrounding or within the system itself to attain a thermal equilibrium state (Patel and Savasni 2015).

Similarly, in HTS algorithm during optimization, if the difference in the solution (i.e. objective function difference) exists within the population, solution tries to improve its value (Patel and Savasni 2015). This improvement can be done by considering the difference between the current solution and either of the best solution, other random solution from the population or the mean value of solution from the population. Like other optimization algorithms, HTS is also a population-based technique, which uses a set of solutions to attain the global optimum solution. The population in HTS is analogous to the clusters of molecules that take part in heat transfer process (Patel and Savasni 2015).

Hence, this newly developed HTS algorithm is successfully used in the analysis part of this paper.

4 Teaching-Learning Based Optimization Algorithm

Teaching-Learning based Optimization (TLBO) algorithm is a global optimization method originally developed by Rao and Patel (2013). It is a population-based iterative learning algorithm that exhibits some common characteristics with other evolutionary computation (EC) algorithms. TLBO algorithms simulate the teaching–learning phenomenon of a classroom to solve multi-dimensional, linear and nonlinear problems with appreciable efficiency (Rao and Patel 2013). The proper tuning of the algorithm-specific parameters is a very crucial factor affecting the performance of optimization algorithms. The improper tuning of algorithm-specific parameters either increases computational effort or yields the local optimal solution. Considering this fact, TLBO was developed which requires only common

controlling parameters like population size and number of generations for its working (Rao and Patel 2013).

This optimization method is based on the effect of the influence of a teacher on the output of learners in a class. A group of learners constitute the population and the different design variables in TLBO are analogous to different subjects offered to learners and the learners result is analogous to the 'fitness', as in other population-based optimization techniques. As the teacher is considered the most learned person in the society, the best solution so far is analogous to Teacher in TLBO (Rao and Patel 2013).

The process of TLBO is divided into two parts. The first part consists of the "Teacher phase" and the second part consists of the "Learner phase". The "Teacher phase" means learning from the teacher and the "Learner phase" means learning through the interaction between learners. The TLBO algorithm has gained wide acceptance among the optimization researchers (Rao and Patel 2013).

In this way, this TLBO technique is employed along with HTS to achieve great results.

5 Mathematical Modelling of Semi-active Suspension for Quarter Car Model

As explained earlier, quarter car model is used for the analysis of this paper as it is simple and can capture some of the very important characteristics of the full model (Gadhvi and Savsani 2014). From the results of this model, the analysis can be expend into full car model.

In this section, the details about the quarter car model, its equations of motion and the various selected parameters are discussed. The purpose of this mathematical modelling is to obtain a state space representation of the quarter car model.

5.1 Quarter Car Semi-active Suspension

5.1.1 Figure of Quarter Car

Figure 2 shows a basic two-degree-of freedom system representing the model of the quarter car semi-active suspension. In the above figure, the symbol m_1 is the unsprung mass, m_2 is the sprung mass, C_s is the damping coefficient, k_1 is tire stiffness, k_2 is suspension stiffness, z_1 & z_2 are vertical displacement of unsprung mass and sprung mass respectively and y is road profile (Segla and Reich 2007).

Here, the damper in semi-active suspension is controlled by a controller and is designed to improve the performance of the suspension as compared to its passive counterpart.

5.1.2 Equations of Motion

These dynamic equations are simply derived from the Newton's law (Segla and Reich 2007),

For m_1,

$$m_1\ddot{z}_1 + C_s(\dot{z}_2 - \dot{z}_1) + k_2(z_1 - z_2) + k_1(z_1 - y) = 0 \qquad (1)$$

For m_2,

$$m_2\ddot{z}_2 + C_s(\dot{z}_2 - \dot{z}_1) + k_2(z_2 - z_1) = 0 \qquad (2)$$

5.2 Formulation of Optimization Problem

$$f = \int_{t=0}^{t=t_{max}} |\ddot{z}_2(t)|dt \qquad (3)$$

5.3 Road Profile and Its Equation

The suspension system is subjected to excitation by a circular bump as shown in Fig. 3 of the height (h) = 0.05 m and length (L) = 3 m (Segla and Reich 2007) (Table 1).

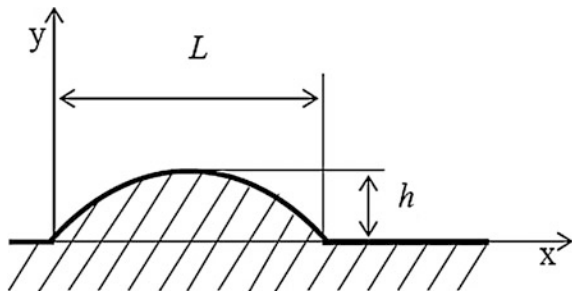Fig. 3 Road profile and its dimensions (Segla and Reich 2007)

Table 1 Numerical parameters taken for the quarter car suspension model (Segla and Reich 2007)

Description	Variable	Value	Unit
Unsprung mass	m_1	55	kg
Sprung mass	m_2	400	kg
Tire stiffness	k_1	218,408.2	N/m
Suspension stiffness	k_2	40,624.2	N/m
Ride velocity	v	20	m/s

Table 2 Values obtained of C_s and f using HTS and TLBO

Run	HTS		TLBO	
	Cs (N/m)	f (m/s^2)	Cs (N/m)	f (m/s^2)
Run 1	2000	1.8662	2000	1.8662
Run 2	2000	1.8662	2000	1.8662
Run 3	2000	1.8662	2000	1.8662
Run 4	2000	1.8662	2000	1.8662
Run 5	2000	1.8662	2000	1.8662

5.4 Range of Values of Decision Variable

The lower and upper limit of C_s are taken as 500 and 2000 (N s/m) respectively. These values are taken after considering the average general values of medium weight vehicles.

6 Results and Discussion

As stated earlier, the optimization process is carried out using two algorithms—HTS and TLBO. The results of these two algorithms are shown in Table 2.

From the above result table, it is found that the minimum value of acceleration of sprung mass is 1.8662 m/s^2 with the suspension damping coefficient as 2000 N s/m.

6.1 Plots of Passive and Semi-active Suspension

Figures 4 and 5 are plotted in MATLAB which shows the comparison of the passive and semi-active suspensions under the same conditions.

These plots shows the variation of linear acceleration of sprung mass with respect to time for both type of suspensions.

Passive suspension system has fixed parameters so the amplitude of acceleration of sprung mass is very high initially and then it settles down after some time.
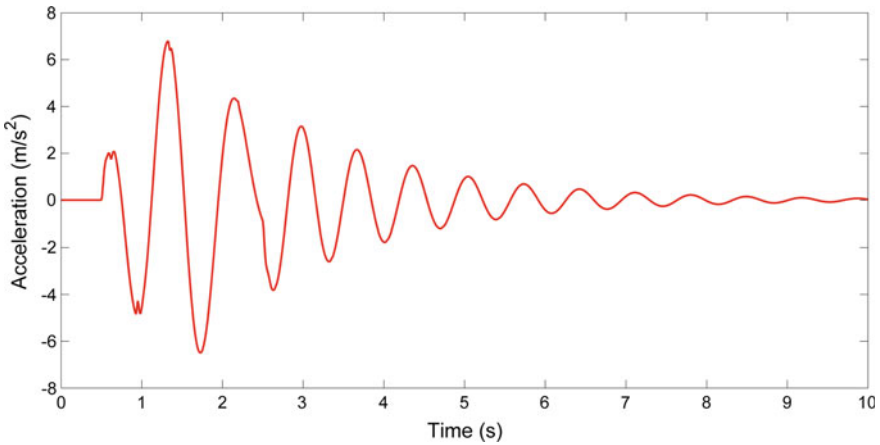
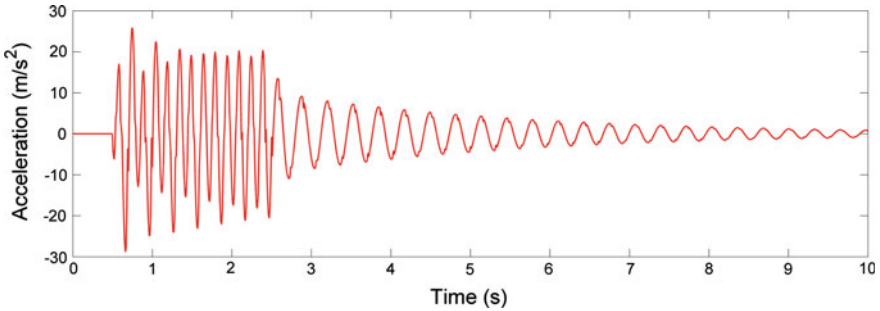Fig. 4 Car acceleration with semi-active suspension

Fig. 5 Car acceleration with passive suspension

Whereas in case of semi active suspension, the amplitude of acceleration is limited up to a very small range. It clearly points out that semi-active suspension has the ability to provide good road handling in different bad road conditions.

Also, the graph of semi active suspension system settles down to nearly zero value much before the time taken by passive suspension system. So the settling-time of semi-active suspension is lower than the passive which indicates that the effect of vibrations generated due to road disturbances doesn't last longer in this system.

The reason for the better performance of semi active suspension can be deduced from the fact that a varying damping coefficient is used in this system whose value is continuously adjusted depending on the nature of road profile encountered.

Thus, we can conclude that semi active suspension system provides better ride comfort and road handling than passive suspension system.

7   Conclusion

Semi-active and passive suspension systems were studied and a detailed comparative analysis was made between the two using a two degree-of-freedom quarter car model. Linear acceleration was taken as the parameter to judge the ride quality. Semi-active suspension has been found to be more efficient as acceleration value is reduced to more than thrice as compared to its passive counterpart. By including an active element in the suspension, it is possible to improve its performance rather than using purely passive elements.

The objectives of the project has been achieved. Thus, by using semi-active suspension the rider can achieve good road handling as well as good ride comfort.

Another important noticeable point is that the results obtained by both the algorithms are exactly same. The reason of this is that the case taken in this study is a simple one and only one parameter is optimized in this paper. Both the algorithms—HTS and TLBO are advanced algorithms and they have promising characteristics but since the problem here is simple, the results obtained are same. This concept can be extended to other models (half car, full car etc.) where the results can be compared by both the algorithms.

References

Agharkakli, A., Sabet, G. S., & Barouz, A. (2012). *International Journal of Engineering Trends and Technology*, *3*(5).

Alexandru, C., & Alexandru, P. (2011). *International Journal of Mechanics*, *5*(4).

Gadhvi, B., & Savsani, V. (2014). ASME 2014 International Mechanical Engineering Congress and Exposition.

Gandhi, Y., Mehta, V., Patel, M., Gadhvi, B., & Markana, A. (2015). *Journal of Aeronautical and Automotive Engineering*, *33*(7).

Gowda, D. V., & Chakrasali, S. (2014). International Conference on Recent Trends in Signal Processing, Image Processing and VLSI.

Mohan Rao, T. R., Rao, G. V., Rao, K. S., & Purushottam, A. (2010). *IJRRAS*.

Mulla, A. A. (2013). International Conference on Emerging Trends in technology & Its Applications.

Patel, V. K., & Savasni, V. J. (2015). *Information Sciences, 324*, 217–246.

Patil, K. S., Jagtap, V., Jadhav, S., Bhosale, A., & Kedar, B. (2014). *IOSR Journal of Mechanical and Civil Engineering*, 6–14.

Rao, R. V., & Patel, V. (2013). *Scientia Iranica, 20*(3), 710–720.

Segla, S., & Reich, S. (2007). 12th IFToMM World Congress.

Shiao, Y. (2010). SICE Annual Conference, 2070–2082.

Depth Control of AUV Using a Buoyancy Control Device

**Mahdi Choyekh, Naomi Kato, Yasuaki Yamaguchi,
Ryan Dewantara, Hidetaka Senga, Hajime Chiba,
Muneo Yoshie, Toshinari Tanaka and Eiichi Kobayashi**

Abstract A new method for depth control was developed for a spilled oil and blow out gas tracking autonomous buoy robot called SOTAB-I by adjusting its buoyancy control device. It is aimed to work for any target depth. The new method relies on buoyancy variation model with depth that was established based on experimental data. The depth controller was verified at sea experiments in Toyama bay in Japan and showed good performance. The method could further be adapted to altitude control by combining the altitude data measured from bottom tracking through a progressive depth control. The method was verified by a simulating program and showed that the algorithm succeeded to bring the robot to the target altitude.

Keywords AUV · Depth control · Buoyancy device

1 Introduction

Oil spills produced by accidents from oil tankers and blowouts of oil and gas from offshore platforms cause tremendous damage to the environment as well as to marine and human life (Seymour and Geyer 1992). To prevent oil and gas that are accidentally released from deep water from spreading and causing further damage to the environment over time, early detection and monitoring systems can be deployed to the area where underwater releases of the oil and gas first occurred. Monitoring systems can provide a rapid inspection of the area by detecting chemical substances and collecting oceanographic data necessary for enhancing the accuracy of simulation of behavior of oil and gas. An autonomous underwater vehicle (AUV) called the spilled oil and gas tracking autonomous buoy system (SOTAB-I) is being developed to perform on-site measurements of oceanographic

M. Choyekh (✉) · N. Kato · Y. Yamaguchi · R. Dewantara · H. Senga ·
H. Chiba · M. Yoshie · T. Tanaka · E. Kobayashi
Department of Naval Architecture and Ocean Engineering,
Osaka University, Yamadaoka 2-1 Suita, Osaka 565-0871, Japan
e-mail: mahdi.choyekh@gmail.com

data as well as dissolved chemical substances using underwater mass spectrometry (Choyekh et al. 2014).

SOTAB-I has three main surveying modes. At the first stage, SOTAB-I performs the water column survey by adjusting its buoyancy. The rough mode is used to collect rough data on physical and chemical characteristics of plumes by repeating descending and ascending on an imaginary circular cylinder centered at the blowout position of oil and gas through the variation of buoyancy and movable wings' angles. Finally, in case the UMS detects a high concentration of any particular substance, a precise guidance mode will be conducted to track and survey its detailed characteristics by repeating descending and ascending within the plume. The photograph mode enables us to have a large visual overview of the area around the blowout position of oil and gas by taking pictures of the seabed and making image mosaicking. SOTAB-I moves laterally using horizontal thrusters along diagonal lines of a polygon with a radius of 5 m centered on the blowout position of oil and gas. Therefore, the depth control is a very important task in the surveying effort which requires particular attention.

There are many challenges and constraints associated with depth control of underwater vehicle. For instance, at-sea experiments require enormous financial and logistic resources limiting the experiments time. Hence, it is important that the program should be easy to implement and repeatedly verified by simulating programs before its real deployment. On the other hand, environmental constraints like a considerable variation of the density of water between the sea surface and the seabed bring complications in the control because they lead to the variation of the neutral buoyancy value of the robot. Even if the neutral buoyancy of the robot is determined accurately at a certain spatial condition, there is no guarantee that the robot will keep its vertical position due to the up-welling and down-welling water currents. Other constraints are represented by the hardware limitations. In fact, the buoyancy device employed has three controlling states: it can be controlled to increase the buoyancy, decrease it or stay idle. However, it is not possible to change the rate of variation directly. In addition, the rate of change of buoyancy is relatively slow, not symmetric in both directions and vary with depth. Moreover, the change of the buoyancy variation orientation is not instantaneous, there is a lag time of 2 s between each change of state. The oil level sensor has also an inaccuracy within ±0.1 %.

Previously, a PID controller was developed for depth control (Kato et al. 2015). It gave good performances and small overshoot, but only for a depth range up to 100 m. Beyond that limit, significant overshoot was reported. The previous controller relied on a very accurate determination of the neutral buoyancy. In addition, the PID control parameters were not adaptive. Besides, it doesn't enable to freeze the robot at the target depth. For the lacks mentioned before, it is necessary to develop a new controller that overcomes the shortening and take in consideration the environmental and hardware constraints.

A new method for depth control was developed. It is aimed to work for any target depth and to freeze the robot at the target depth. The method relies mainly on

the buoyancy variation model with depth established based on tank and at-sea experimental data.

2 Description of SOTAB-I and Its Buoyancy Control Device

2.1 SOTAB-I Overview

The SOTAB-I is 2.5 m long and weighs 325 kg. It can be submerged in water as deep as 2000 m. It is able to descend and ascend by adjusting its buoyancy using a buoyancy control device while changing its orientation through two pairs of movable wings. The SOTAB-I can also move in horizontal and vertical directions using two pairs of horizontal and vertical thrusters. The arrangement of devices and sensors installed on SOTAB-I is shown in Fig. 1.

When the robot floats on the sea surface, a wireless local area network (WLAN) and an iridium satellite communication transceiver module are used for data transmission. When the robot is underwater, the user on the mothership and the SOTAB-I can communicate through the acoustic modem.

The robot tracking on the sea surface is ensured by a global positioning system (GPS) receiver that serves to determine the robot's absolute position. In the case where the robot is submerged, tracking is ensured by the ultra-short baseline (USBL) system. The vertical position of the robot in the water column is given by depth data from the CTD sensor. When the robot is within the bottom tracking altitude from the seabed, the Doppler velocity logger (DVL) is able to measure robot's velocities. The robot motion and orientation are given by the compass and the Inertial Measurement Unit (IMU). An ADCP is employed to measure the magnitude and orientation of water current layers. SOTAB-I is also fitted with a UMS to determine the characteristics and physical properties of the dissolved gas and oil (Short et al. 2006). To obtain a visual representation of blowouts of plumes of gas on the seabed, the robot is equipped with a camera.

2.2 Buoyancy Control Device

In the buoyancy control device (Fig. 2), an oil hydraulic pump injects and extracts oil between the external oil bladder and the internal oil reservoir. A motor valve serves to automate opening and closing cycles and a brake is used to lock the pump. Table 1 defines all the parameters related to the buoyancy control that will be used in the next sections.
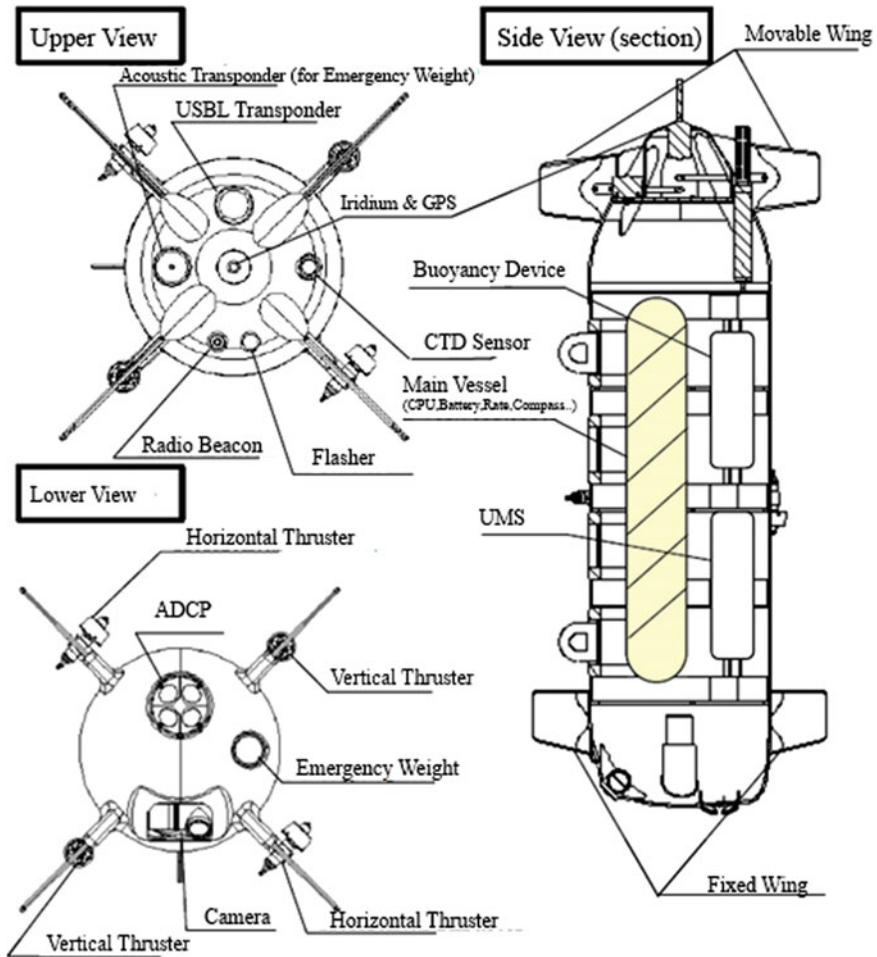
Fig. 1 Arrangement of devices and sensors installed on SOTAB-I

3   Establishment of the Buoyancy Model

The objective in this section is to establish a time model and a buoyancy model. The time model is needed for the depth control of the robot. It enables to estimate the time needed for SOTAB-I to change its buoyancy from its current value to a target value. The buoyancy model is needed for the simulating program. It enables to estimate the variation of the buoyancy value from its initial value every sampling period. We consider establishing a model for the buoyancy variation from 20 to 85 % up to 1000 m water depth based on experimental results in a pressure tank and at-sea.
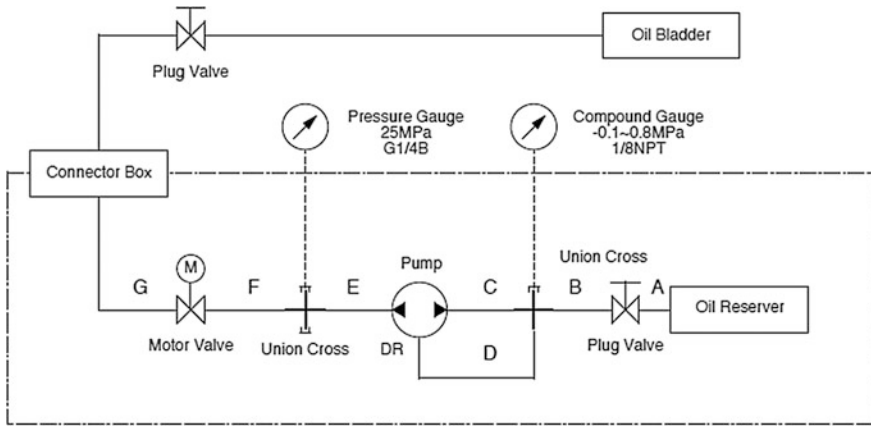
Fig. 2 Buoyancy device

Table 1 Definition of the buoyancy control parameters

Symbol	Definition
D	Current depth (m). D > 0
D_t	Target depth (m)
D_m	Margin of tolerance around D_t
S	Current speed (m/s). S > 0 $\rightarrow$ Robot descending
B	Current buoyancy (%). Range: 20 > 95 %
B_t	Output target buoyancy
B_n	Neutral buoyancy
B_m	Margin of tolerance around the neutral buoyancy B_n
T_c	Time needed to change the buoyancy from B to B_t in (s)
T_r	Time needed to reach the target depth based on the current speed of the robot
T_m	Time margin used for security purpose. It compensates eventual inaccuracy in the buoyancy model

3.1 Experiments Results of Buoyancy Variation

3.1.1 Pressure Tank Experiments

Pressure tank experiments were performed to calculate the time necessary for changing 7500 cc of oil in the reservoir in both directions. OUT → IN direction is when the oil hydraulic pump injects and extracts oil from the external oil bladder and injects it to the internal oil reservoir. IN → OUT is the opposite direction.

Based on Fig. 3, the time needed to change the buoyancy corresponding to 7500 cc in the OUT > IN direction is constant till 10 MPa (∼1000 m water depth). From 10 to 20 MPa, it can be modeled as a linear function. The time difference

Fig. 3 Relationship between the pressure and the buoyancy variation time
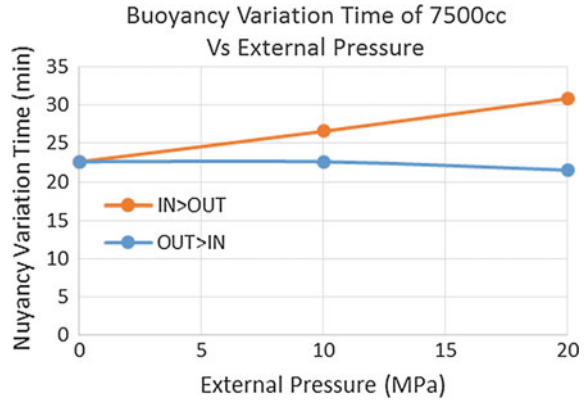
between the full scale variation of buoyancy at 0 and 20 MPa is less than 9 min. For the IN > OUT direction, The Buoyancy variation time is almost same from 0 up to 10 MPa. Beyond that limit, it becomes slightly faster.

3.1.2 At-sea Experiments

Figure 4a confirms the results obtained in the pressure tank. The buoyancy variation rate is almost same from 0 to 100 m. Hence, the time and buoyancy models in the OUT > IN direction can be modeled as a linear function. Figure 4b dates from an experiment on the 20th of March 2015 in Toyama Bay. It shows the buoyancy variation in the IN > OUT direction from 20 to 85 % at 1 and 700 m water depths. It shows that buoyancy variation can be represented under the form of a 3rd degree polynomial function as shown in Table 2.

3.2 Model of the Buoyancy Variation with Depth

Based on the experimental data, the buoyancy variation in the OUT > IN direction can be modeled as a linear function.

$$Tc = -15.122 * (B_t - B) \tag{1}$$

In the IN > OUT direction, the main parameter that contributes considerably in the change of the buoyancy variation speed is the depth. The model can be generalized and written under the form of 3rd degree polynomial function that depends on the depth D.
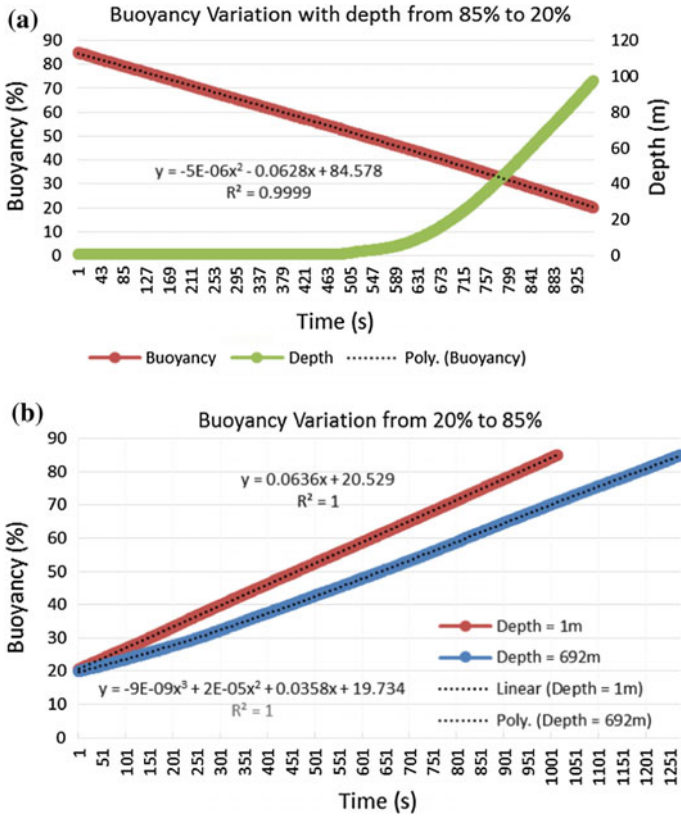
Fig. 4 Buoyancy variation with depth: **a** OUT > IN, **b** IN > OUT

Table 2 Buoyancy variation

OUT > IN	1–700 m	$Tc = -15.122 * (B_t - B)$
IN > OUT	1 m	$Tc = 15.122 * (B_t - B)$
	700 m	$Tc = 0.0015 * (B_t^3 - B^3) - 0.2688 * (B_t^2 - B^2) + 34.065 * (B_t - B)$

$$Tc = C_3(D) * (B_t^3 - B^3) - C_2(D) * (B_t^2 - B^2) + C_1(D) * (B_t - B) \qquad (2)$$

Linear interpolation and extrapolation of the coefficients a, b and c are used to determine the buoyancy model at a certain depth based on the models established for depths equal to 1 and 700 m. Equation 3 shows the formula used to calculate $C_i(D)$ for.

Table 3 Comparison between buoyancy variation time obtained by the model and experiments results

Experiment	Depth	Orientation	Range of variation (%)	Time (experiment) (s)	Time (model)	Time/Range (s/1 %)
5/25/2015 2:28:43	Air	OUT > IN	95 → 21	1121	1119 s	−0.03
		IN > OUT	20 → 94	1166	1164 s	−0.03
11/27/2014 9:20:56	0–95 m	OUT > IN	79 → 20	859	892 s (50 m)	+0.56
	700 m	IN > OUT	20 → 85	1274	1340 s (700 m)	+1.01
11/28/2014 9:48:23	0–42 m	OUT > IN	78 → 31	703	710 s	+0.15
	0–100 m	IN > OUT	31 → 79	854	823 s (50 m)	−0.65
3/20/2015 14:12:33	95–155 m	IN > OUT	20 → 30	178	199 s (155 m)	+2.1
					184 s (95 m)	+0.6
	155–198 m	IN > OUT	30 → 40	176	192 s (155 m)	+1.6
					193 s (198 m)	+1.5
	198–210 m	IN > OUT	40 → 49	170	169 s (198 m)	−0.11

$$C_i(D) = (C_i(1) - C_i(700)) * D/(700 - 1), \quad i = \{1, 2, 3\} \tag{3}$$

Knowing that

- At 700 m $C_3 = 0.0015$; $C_2 = 0.2688$; $C_1 = 34.065$
- At 1 m $C_3 = 0$; $C_2 = 0$; $C_1 = 15.122$.

3.3 Comparison Between Buoyancy Variation Model and Experimental Data

Data of buoyancy variation time in several ranges were collected from previous at-sea experiments of SOTAB-I and compared to the values obtained by the model. To estimate the accuracy of the model, we defined the ratio Time/Range to estimate the time deviation of the model in seconds for every 1 % of buoyancy variation. Table 3 summarizes the obtained results.

Table 3 shows the estimated time of buoyancy variation is close to the values obtained from experiments. The maximum deviation was obtained when the buoyancy variation is between 20 and 30 %.

4 Depth Control

4.1 Control Algorithm

The depth control algorithm is shown in Fig. 5. The TcTr control serves to smoothly bring the robot to the target depth. Once the target depth is reached, the

Fig. 5 Depth control process
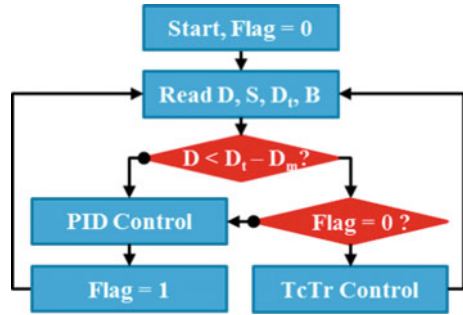
PID control which is used to freeze the robot there. In this paper, the TcTr control is mainly focused.

CTD sensor enables to measure the depth (D) and the vertical speed (S) of the robot. At every moment, it is possible to have an estimation of the time needed to reach the target depth (Tr) using Eq. 4.

$$Tr = ((Dt - D)/S) \qquad (4)$$

The buoyancy variation model established in Sect. 3 of this paper enables to estimate the time (Tc) needed for changing the robot buoyancy from it is current value to the neutral buoyancy. The first step is based on the continuous estimation of the time to reach (Tr) and the time to change (Tc) while adjusting decreasing the buoyancy value, till the stop condition is reached.

- If the estimation of the Tr > Tc it means that it is possible to increase the vertical speed of the robot since we have enough time to change the buoyancy to its neutral level. Hence, we decrease the target buoyancy value.
- In the case where the Tr = Tc or less, then it means we have just enough time to change the buoyancy to the neutral level before the robot reaches its target depth. Hence, we start to increase the buoyancy of the robot progressively.

We introduce Tm, which corresponds to the time error margin used to compensate eventual inaccuracies in the buoyancy device model. The condition for stopping to increase the robot vertical speed becomes Tr ≤ Tc + Tm. Figure 6 shows the TcTr control inputs and outputs diagram and Fig. 7 illustrates it flow chart.
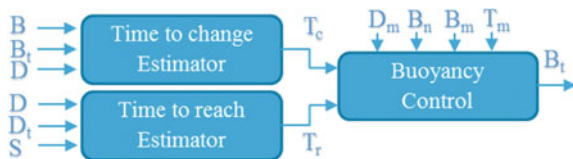
Fig. 6 TcTr input/output diagram
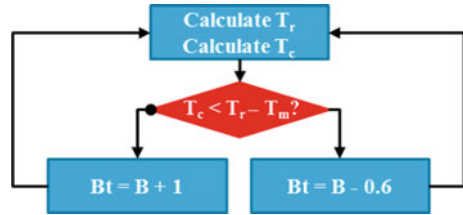
Fig. 7 TcTr control flowchart

Table 4 TcTr parameters configuration

Parameter	Value
Target depth (Dt)	400 m
Depth margin (Dm)	1 m
Neutral buoyancy (Bn)	51.0 %
Buoyancy margin (Bm)	2.0 %
Time margin (Tm)	30 s

4.2 Experiments Results

Sea experiments were conducted in Toyama Bay, Japan on June 11th, 2015 on the board of the Wakashio-maru of the National Institute of Technology, Toyama College. The experimental site was located at 36° 52′N, 137° 11′E with a water depth of around 560 m. In this sea experiment, the depth control with time estimation scheme was deployed, with a target depth of 400 m. The configuration of the depth control parameters is listed in Table 4.

As shown in Fig. 8, the TcTr control succeeded to bring the robot to the target depth smoothly with a vertical speed near 0 m/s. The buoyancy device was able to adjust the buoyancy so that it could achieve neutral buoyancy when reaching the target depth.

5 Progressive Depth Control

One of SOTAB-I operating modes is the photograph mode. To take pictures of the blow out position, SOTAB-I needs to approach the seabed around 4 m altitude. Hence, an altitude control is needed. One way is to use thrusters to control the altitude. However, there is a risk that they mix up the sediments on the seabed which influences the transparency of the water. In addition, it will disturb the water flow, causing some inaccuracies in the water current measurement. In this section we suggest a second method that consists in the adaptation of the depth control algorithm detailed in the previous section to altitude control. This is possible by combining CTD sensor depth information with the DVL altitude measurement

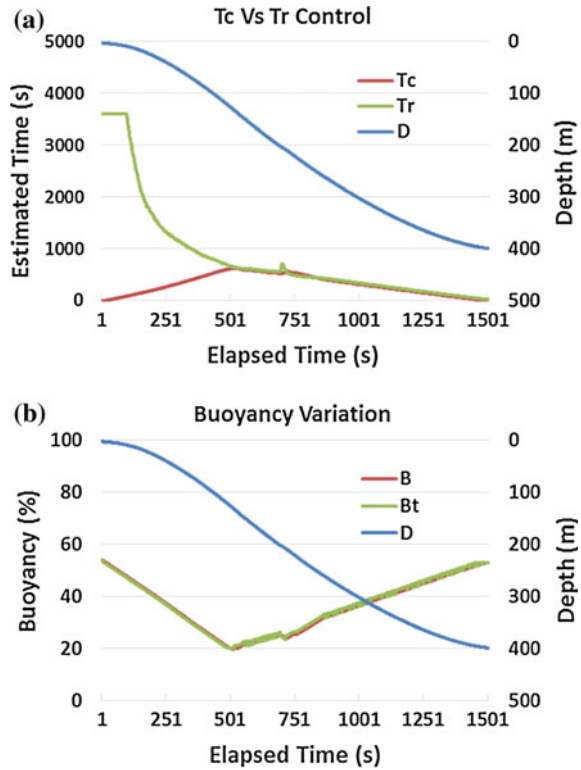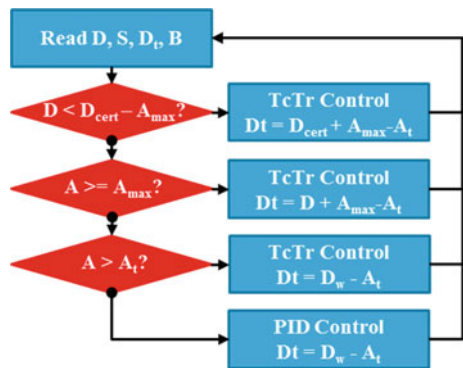Fig. 8 At-sea experiments results of the depth control

Fig. 9 Altitude control flowchart

when bottom tracking is active enabling the measurement of water depth. Figure 9 illustrates the flow chart of the progressive depth control.

Step 1: The robot dives with a fast speed until the depth of the robot reaches the depth limit (D_{cert}) of the "certain zone". The "Certain Zone" is the zone

in which we are sure that the robot will not be within the bottom tracking altitude (A_{max}) equal to 24 m. D_{cert} is input by the user on board before starting the descent. In this first step, the depth control with time estimation scheme is used. At first, the buoyancy control device will decrease the buoyancy. After the buoyancy level becomes lower than the neutral buoyancy of the robot, SOTAB-I will start diving. The buoyancy control device will continue to reduce its buoyancy level down to 20 %, which is set as the minimum buoyancy level of the buoyancy control device, with maximum speed. Then it will increase again its buoyancy level close to the neutral buoyancy level. The purpose of this strategy is that the robot should have enough time to change its buoyancy level to its neutral buoyancy when reaching the target depth. In this step, the target depth Dt is set as a fixed value, which is equal to the certain zone limit plus the DVL range A_{max}, minus the target altitude (A_t).

Step 2: When the robot reaches the certain zone limit, the variable target depth control is started. After passing the certain zone limit, there is a chance that the DVL will detect the seabed and output its current altitude (A). Therefore, from this viewpoint, the buoyancy change is limited up to the time needed to reach the target depth Tr. In this step, the depth control with time estimation is still being used. However, the target depth Dt is set equal to the current depth D plus the DVL range A_{max} minus the target altitude At. The target depth will continuously change as the depth D of the robot decreases. Hence, it is a depth control with variable target depth. At this point, the buoyancy level of SOTAB-I is already close to the neutral buoyancy. Therefore, there will not be much change in the buoyancy level to ensure that the robot is able to stop when reaching the target depth, as shown in Step 2. As a result, the robot will dive at a steady speed.

Step 3: If the seabed is detected by DVL, the target depth Dt is set equal to the water depth D_w minus the target altitude A_t. The water depth D_w is defined as the sum of the depth D measured by CTD and the altitude A measured by DVL. This step is also carried out by using the depth control with time estimation scheme.

Step 4: When the robot is within the range of the target depth plus or minus the depth margin Dm, the depth control method is switched from the depth control with time estimation to the PID depth control method. The depth margin Dm is usually set around 1 m as a compensation in the control mechanism of the buoyancy device. SOTAB-I will stay within the target depth for a certain period of time, which has been set on the timer. When the timer reaches zero, the robot will start ascending.

Simulation result of the algorithm (Fig. 10) shows that that the robot could reach the target altitude from the seabed with a slow speed close to 0 m/s and a buoyancy near the neutral value.
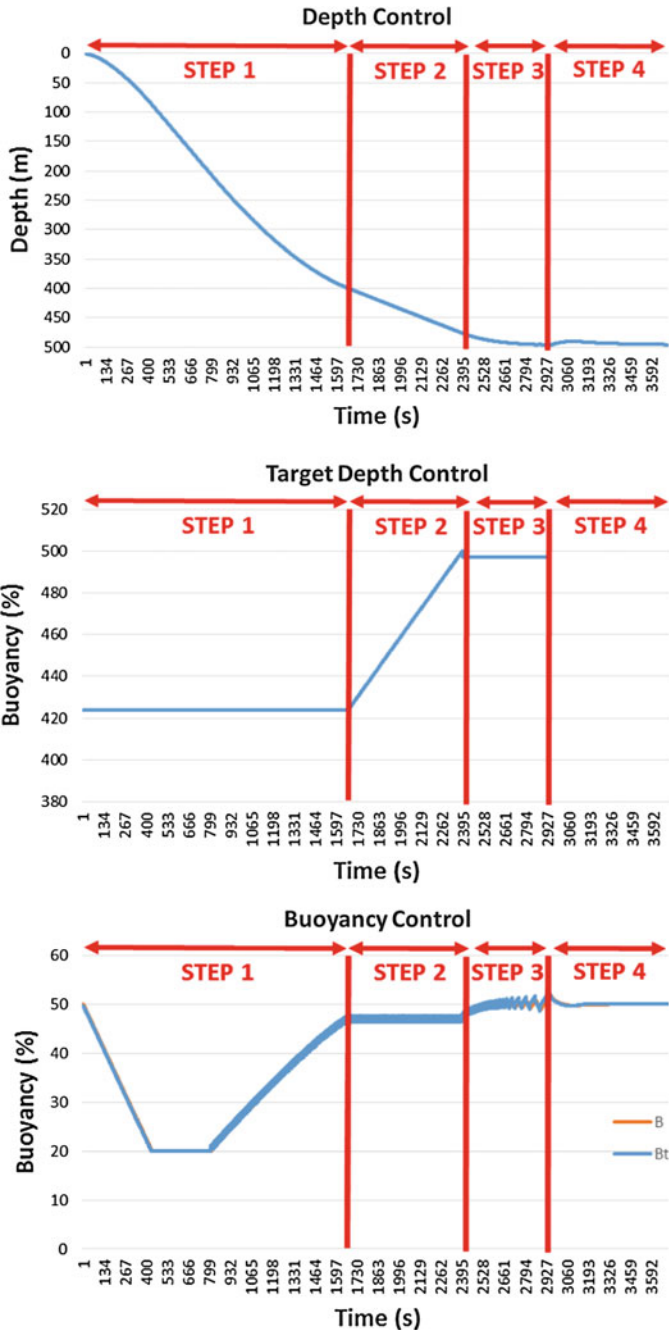
Fig. 10 Simulation results of the altitude control

6   Conclusions

A new method for depth control using the buoyancy control device was developed. A model of the buoyancy variation with time was established. It was built starting from the results obtained on tank and at-sea experiments. The depth control algorithm is based on the comparison between the time estimated for the robot to change its buoyancy from its current value to the neutral value, and the time expected for the robot to reach the target depth. The method was demonstrated at-sea experiments in Toyama Bay in Japan in June 2015. It showed the ability of the control algorithm to smoothly bring the robot to the target depth without a significant overshoot. The algorithm is characterized by its flexibility and doesn't require a strict determination neutral buoyancy value. A margin of inaccuracy can be customized before performing the dive.

The method could be further adapted to perform an altitude control needed for the photograph mode. A progressive depth control algorithm based on 4 steps was established. The results of the simulation showed that it worked properly. Further tests at-sea are needed to practically confirm the obtained results.

References

Choyekh, M., et al. (2014). Vertical water column survey in the Gulf of Mexico using autonomous underwater vehicle SOTAB-I. *Marine Technology Society Journal*, 88–101.

Kato, N., et al. (2015). Autonomous spilled oil and gas tracking buoy system and application to marine disaster prevention system. In Interspill 2015, Amsterdam, Netherland.

Seymour, R. J., & Geyer, R. A. (1992). Fates and effects of oil spills. *Annual Review of Energy and the Environment*, 261–283.

Short, R. T., Toler, S. K., Kibelka, G. P. G., Rueda Roa, D. T., Bell, R. J., & Byrne, R. H. (2006). Detection and quantification of chemical plumes using a portable underwater membrane introduction mass spectrometer. *Trends in Analytical Chemistry, 25*(7), 637–646.

DOB Tracking Control for Systems with Input Saturation and Exogenous Disturbances via T-S Disturbance Modelling

Xiangxiang Fan, Yang Yi and Yangfei Ye

Abstract In this paper, the anti-disturbance dynamical tracking problem is investigated for a class of systems subject to input saturation and unknown disturbances under the framework of disturbance-observer-based-control (DOBC). In order to expand the application scope of exogenous disturbances, T-S fuzzy models are employed to describe those complex nonlinear disturbances, and the corresponding disturbance observer is also well designed. The PI-type composite controller with the estimates of disturbance is designed to ensure the system stability and the convergence of tracking error to zero. Meanwhile, an estimation of domain of attraction can also be described by the level set of the Lyapunov function. Finally, a simulation example for flight control systems with nonlinear disturbances is given to verify the effectiveness of the proposed schemes.

Keywords Disturbance-observer-based control · Input saturation · T-S fuzzy model · Tracking control

1 Introduction

It is well known that exogenous disturbances widely exist in controlled systems, see for instance Yang and Tsubakihara (2008), Chen et al. (2000), Yang et al. (2013), Guo and Chen (2005). Thus, many advanced control approaches have been employed to handle disturbance attenuation and rejection problem, such as adaptive dynamical compensation (Marino and Tomei 1995), H_∞ control (Schaft 1992), output regulator theory (Isidori 1990), and so on. On the other hand, disturbance-observer-based control (DOBC) theory was studied in late 1980s and the basic idea

X. Fan · Y. Yi (✉) · Y. Ye
College of Information Engineering, Yangzhou University,
Huayang West Road, Yangzhou, Hanjiang District, Jiangsu, China
e-mail: yiyangcontrol@163.com

of the DOBC scheme is to construct an observer to estimate the unknown disturbance, then a feed-forward compensator plus conventional control laws are applied to reject the disturbance. However, in most of DOBC results, such as Yao and Guo (2013), Sun and Guo (2014), and Wei et al. (2015), the exogenous disturbances are assumed to be generated by a linear exogenous system, which seriously limits the types of disturbances.

According to Takagi and Sugeno (1985), the Takagi-Sugeno (T-S) fuzzy model, as a powerful tool for approximating complex nonlinear systems, has been intensively investigated during the past decades. By introducing a family of fuzzy IF-THEN rules, the traditional linear control systems are applied to approximate and analyze the nonlinear control systems. Moreover, many typical nonlinear systems, such as descriptor systems (Zhang et al. 2007a), networked control systems (Zhang et al. (2007b), stochastic systems (Yi et al. 2009) and time-delay systems (Qiu et al. 2011), can be modelled and controlled by the designed T-S fuzzy models. Meanwhile, some mature control methods including adaptive control (Tong and Li 2012; Tong et al. 2012), and fault estimation (Zhang et al. 2012), can also be considered based on T-S fuzzy models.

On the other hand, control input saturation is a common problem in a wide range of practical control systems, since the signal that an actuator can implement is impossible to be unlimited. Due to the existence of input saturation, system control performance as well as stability will be greatly effected (Luo and Zhang 2008). Moreover, the problem of controller design for systems with input saturation is widely investigated in recent years, see Hu et al. (2002), Zuo et al. (2010). The control problem for systems with saturated input, which including control input and disturbance input, was developed in Wei et al. (2015).

Motivated by the above observations, this paper considers the disturbance observer based tracking control problem for the nonlinear systems with input saturation and the T-S fuzzy modelling problem for nonlinear irregular disturbances. Following the disturbance observer design based on T-S disturbance models, the composite anti-disturbance controller for systems subject to input saturation is proposed by combining the estimation of disturbance with PI control algorithm, such that there exists an initial condition domain ensuring that for every initial condition from this domain, the stability and the favorable tracking performance of augmented systems can achieved by the designed optimization algorithm in present of disturbances. Finally, simulations for a flight control system are given to show the efficiency of the proposed approach.

2 Problem Formulation

In this paper, we study anti-disturbance control problem for the following uncertain system with input saturation:

$$\begin{cases} \dot{x}(t) = A_0 x(t) + B_0 sat[u(t) + d(t)] + F_{01} f_{01}(x(t), t) \\ y(t) = C_0 x(t) + F_{02} f_{02}(x(t), t) \end{cases} \tag{1}$$

where $x(t) \in \mathbb{R}^n, d(t) \in \mathbb{R}^m, u(t) \in \mathbb{R}^m$ and $y(t) \in \mathbb{R}^{p_1}$ are the state, the unknown disturbance, control input and the measurement output, respectively. A_0, B_0, C_0 and F_{01} are coefficient matrices. $sat(\cdot)$ is the standard saturation function with $sat(\cdot)_i = sgn(\cdot)\min\{\cdot, 1\}$. $f_{0i}(x(t), t)$ is known nonlinear functions.

Assumption 1 For any $x_j(t) \in R^n (j = 1, 2), f_{0i}(x(t), t)$ satisfies

$$\begin{cases} f_{0i}(0, t) = 0 \\ \|f_{0i}(x_1, t) - f_{0i}(x_2, t)\| \leq \|U_i(x_1 - x_2)\| \end{cases} \tag{2}$$

where U_i are known constant weighting matrices.

Different from common disturbances such as harmonics and constant disturbances, T-S fuzzy models are employed to model more complex disturbances. The unknown nonlinear disturbance $d(t)$ is supposed to be generated by the following T-S fuzzy system with r rules:

Plant rule j: If θ_1 is A_1^j, and ... and θ_n is A_n^j, then

$$\begin{cases} \dot{w}(t) = W_j w(t) \\ d(t) = V_j w(t) \end{cases} \tag{3}$$

where $w(t) \in \mathbb{R}^q$, W_j, V_j are known coefficient matrices. θ_l and A_i^j are the premise variables and the fuzzy sets, respectively. n is the number of If-Then rules and the premise variables. By fuzzy blending, the global fuzzy model are as follows:

$$\begin{cases} \dot{w}(t) = \sum\limits_{j=1}^{r} h_j(\theta) W_j w(t) \\ d(t) = \sum\limits_{j=1}^{r} h_j(\theta) V_j w(t) \end{cases} \tag{4}$$

where $\quad w_j(\theta) = \prod_{i=1}^{n} A_i^j(\theta_i), \quad h_j(\theta) = w_j(\theta) \Big/ \sum_{j=1}^{r} w_j(\theta), \quad h_j(\theta) \geq 0 \quad$ and $\sum_{j=1}^{r} h_j(\theta) = 1$.

Assumption 2 (A_0, B_0) is controllable and $(W_j, B_0 V_j)$ is observable.

In order to realize the tracking performance of system, we introduce an extended state variable:

$$z(t) := \left[x^T(t), \int\limits_0^t e^T(\tau) d\tau \right]^T \tag{5}$$

where $e(t) := y(t) - y_d$ and y_d is the desired system output. Furthermore, we can construct that

$$\dot{z}(t) = Az(t) + B_1 sat[u(t) + d(t)] + Cy_d + F_1 f_{01}(z(t), t) \tag{6}$$

where $A = \begin{bmatrix} A_0 & 0 \\ C_0 & 0 \end{bmatrix}, B_1 = \begin{bmatrix} B_0 \\ 0 \end{bmatrix}, C = \begin{bmatrix} 0 \\ -I \end{bmatrix}, F_1 = \begin{bmatrix} F_{01} \\ 0 \end{bmatrix}.$

According to (6), PI-type control input is designed as follows:

$$u_{PI}(t) = K_p x(t) + K_I \int_0^t e(\tau) d\tau, \quad K = [K_P \quad K_I] \tag{7}$$

where K_P, K_I are controller gains to be determined later.

Lemma 1 (Hu and Lin 2001). *Given K and H in $\mathbb{R}^{m \times (n+p_1)}$. For a $z \in \mathbb{R}^{n+p_1}$, if $z \in L(H)$, then*

$$sat(Kz) = co\{D_i Kz + D_i^- Hz, i \in Q\} \tag{8}$$

where D_i is a diagonal matrix with each element of the diagonal being either 1 or 0, $D_i + D_i^- = I$, $Q = \{1, \ldots, 2^m\}$, and $co(\cdot)$ denotes the convex hull of a set.

Let $P_1 \in \mathbb{R}^{(n+p_1) \times (n+p_1)}$ be a positive definite matrix. We denote

$$\Omega(P_1) = \{z \in \mathbb{R}^{n+p_1} : z^T P_1 z \leq 1\} \tag{9}$$

and a symmetric polyhedron

$$L(H) = \{z \in \mathbb{R}^{n+p_1} : |H^l z| \leq 1, l \in Q_m\} \tag{10}$$

where $Q_m = \{1, \ldots, m\}$ and H^l is the lth row of the matrix H.

3   Controller Design and Stability Analysis

We construct the following disturbance observer to estimate disturbance $d(t)$:

$$\hat{d}(t) = \sum_{j=1}^{r} h_j(\theta) V_j \hat{w}(t), \hat{w}(t) = v(t) - Lz(t) \tag{11}$$

$$\dot{v}(t) = \sum_{j=1}^{r} h_j(\theta) \{ (W_j + LB_1 V_j)[v(t) - Lz(t)] - L[-Az(t) - B_1 u(t) \\ - Cy_d - F_1 f_{01}(z(t), t)] \} \tag{12}$$

where $\hat{d}(t)$ and $\hat{w}(t)$ are the estimations of $d(t)$ and $w(t)$, respectively. $v(t)$ is the auxiliary variable, L is the observer gain to be determined.

Define $e_w(t) := w(t) - \hat{w}(t)$ and introduce variable

$$\sigma(t) = \begin{bmatrix} z^T(t) & e_w^T(t) \end{bmatrix}^T \tag{13}$$

According to Lemma 1, and combining (4), (8) with (10), for $\forall \sigma(t) \in L(H)$ with $H = \left[H_1, \sum_{j=1}^{r} h_j(\theta) V_j \right]$, the saturation can be expressed as

$$sat(u(t) + d(t)) = \sum_{i=1}^{2^m} \delta_i \left(D_i K + D_i^- H_1 \right) z(t) + \sum_{j=1}^{r} h_j(\theta) V_j e_w(t) \tag{14}$$

Combing (5), (11), (12) with (14), yields

$$\dot{e}_w(t) = \sum_{j=1}^{r} h_j(\theta)(W_j + LB_1 V_j)e_w(t) + \sum_{i=1}^{2^m} \delta_i LB_1 D_i^-(H_1 - K)z(t) \tag{15}$$

By integrating the estimation of disturbance with PI-type control input, the composite controller is inferred as

$$u(t) = -\hat{d}(t) + Kz(t), \quad K = \begin{bmatrix} K_P & K_I \end{bmatrix} \tag{16}$$

Integrating disturbance error (15) with system (6), we can obtain that

$$\dot{\sigma}(t) = \bar{A}\sigma(t) + \bar{C}y_d + \bar{F}_1 f_{01}(\sigma(t), t) + \bar{F}_2 f_{02}(\sigma(t), t) \tag{17}$$

where

$$\bar{A} = \begin{bmatrix} A + \sum_{i=1}^{2^m} \delta_i B_1 (D_i K + D_i^- H_1) & \sum_{j=1}^{r} h_j(\theta) B_1 V_j \\ \sum_{i=1}^{2^m} \delta_i LB_1 D_i^- (H_1 - K) & \sum_{j=1}^{r} h_j(\theta)(W_j + LB_1 V_j) \end{bmatrix}$$

$$\bar{C} = [C \quad 0]^T, \ \bar{F}_1 = [F_1 \quad 0]^T, \ \bar{F}_2 = [F_2 \quad 0]^T, \ f_{0i}(\sigma(t), t) = f_{0i}(x(t), t)$$

Theorem 1 *For system* (17), *if there exist matrices* $Q_1 = P_1^{-1} > 0$, $P_2 > 0$, R_1, R_2, R_3 *and scalars* $\mu_1 > 0$, $\mu_2 > 0$, $\mu_3 > 0$ *satisfying*

$$\begin{bmatrix} \Pi & [R_3B_1\delta_i(R_2-R_1)]^T + B_1V_j & Q_1U_1^T & Q_2U_2^T \\ * & sym(P_2W_j + R_3B_1V_j) & 0 & 0 \\ * & * & -\mu_2I & 0 \\ * & * & * & -\mu_3I \end{bmatrix} < 0 \qquad (18)$$

and

$$\begin{bmatrix} 1 & R_2^l & V_j^l \\ * & Q_1 & 0 \\ * & * & P_2 \end{bmatrix} \geq 0 \qquad (19)$$

where R_2^l and V_j^l are the lth row of the matrix R_2 and V_j, respectively, and $\Pi = sym(AQ_1 + B_1D_iR_1 + B_1D_i^-R_2) + \mu_1CC^T + \mu_2F_1F_1^T + \mu_3F_2F_2^T$. Then the system (17) under the composite controller (16) is stable and the tracking error satisfies $\lim_{t\to\infty} y(t) = y_d$ with $K = R_1Q_1^{-1}, L = P_2^{-1}R_3, H_1 = R_2Q_1^{-1}$. Meanwhile, an estimation of the domain of attraction of the system (17) is given by $\Omega(P)$, where $P = diag\{Q_1^{-1}, P_2\}$.

Proof Consider a Lyapunov function candidate as follows

$$V(\sigma(t)) = \sigma^T(t)P\sigma(t) \qquad (20)$$

Using Assumption 1, it is easy to get that

$$\dot{V}(\sigma(t)) \leq \max_{i\in Q}\left\{\sum_{j=1}^{r} \delta_i\sigma^T(t)\Theta_{ij}\sigma(t)\right\} + \mu_1y_d^2 \leq -\alpha_1\|\sigma(t)\|^2 + \mu_1y_d \qquad (21)$$

where $\alpha_1 > 0$ and

$$\Theta_{ij} = \begin{bmatrix} \Xi_i & P_1B_1V_j + [P_2LB_1D_i^-(H_1-K)]^T \\ * & sym(P_2W_j + P_2LB_1V_j) \end{bmatrix},$$

$$\Xi_i = sym[P_1A + P_1B_1(D_iK + D_i^-H_1)]$$
$$+ \frac{1}{\mu_1}P_1CC^TP_1 + \frac{1}{\mu_2}P_1F_1F_1^TP_1 + \frac{1}{\mu_3}P_1F_2F_2^TP_1 + \mu_2U_1^TU_1 + \mu_3U_2^TU_2.$$

Obviously, if $\|\sigma(t)\|^2 > \alpha_1^{-1}\mu_1y_d^2$, thus $\dot{V} < 0$. It is noted that for any $z(t)$ and $e_w(t)$, we have

$$\sigma^T(t)\sigma(t) \leq \max\{\sigma^T(0)\sigma(0), \alpha_1^{-1}\mu_1y_d^2\} \qquad (22)$$

where $\sigma(0)$ is the initial value of $\sigma(t)$, which implies that system (17) is stable.

Supposed that $\phi_1(t)$ and $\phi_2(t)$ are two trajectories of the closed-loop system (17) corresponding to a fixed initial condition. Defining

$$\bar{\phi}(t) = \phi_1(t) - \phi_2(t) \tag{23}$$

Then we have

$$\dot{\bar{\phi}}(t) = \bar{A}\bar{\phi}(t) + \bar{F}_1(f_{01}(\phi_1, t) - f_{01}(\phi_2, t)) + \bar{F}_2(f_{02}(\phi_1, t) - f_{02}(\phi_2, t)) \tag{24}$$

Choose Lyapunov function as $\Gamma(\bar{\phi}, t) = \bar{\phi}^T P \bar{\phi}$, then

$$\dot{\Gamma}(\bar{\phi}, t) \leq \bar{\phi}^T \Psi \bar{\phi} < 0 \tag{25}$$

Obviously, (25) can be guaranteed by (19) in Theorem 1. It can be verified that the system (24) is asymptotically stable and $\bar{\phi}(t) \to 0$. Thus, it follows that $\lim_{t \to \infty} y(t) = y_d$.

On the other hand, by pre-multiplying and post-multiplying $diag\{I, Q_1^{-1}, I\}$ to both side of (19), we can obtain that

$$\begin{bmatrix} 1 & H_1^l & \sum\limits_{j=1}^{r} h_j V_j^l \\ * & P_1 & 0 \\ * & * & P_2 \end{bmatrix} \geq 0 \tag{26}$$

According to Hu and Lin (2001), (26) implies that $\Omega(P) \subset L(H)$. It means that for any $\sigma(t) \in \Omega(P), \sigma(t) \in L(H)$.

4 Simulation Examples

Consider the A4D aircraft model in Guo and Chen (2005), which is subject to input saturation and known nonlinearity. The longitudinal dynamics of A4D aircraft at a flight condition of 0.9 Mach and 15,000 ft. altitude can be described as (1), where $x(t) = [x_1(t), x_2(t), x_3(t), x_4(t)]^T, u(t)$ represents elevator deflection (deg). The related matrices are given as

$$A_0 = \begin{bmatrix} -0.0605 & 32.37 & 0 & 32.2 \\ -0.00014 & -1.475 & 1 & 0 \\ -0.0111 & -34.72 & -2.793 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, B_0 = \begin{bmatrix} 0 \\ -0.1064 \\ -33.8 \\ 0 \end{bmatrix},$$

$$F_{01} = \begin{bmatrix} 0 & 0 & 50 & 0 \end{bmatrix}^T, F_{02} = 0, C_0 = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix},$$
$$U_1 = diag\begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}, f_{01}(x(t), t) = \sin(2\pi 5t)x_2(t).$$

The nonlinear exogenous disturbance is described by the following T-S fuzzy model:

Rule j: If w_1 is A_1^j $(j = 1, 2)$, then $\dot{w}(t) = W_j w(t), d_1(t) = V_j w(t)$, with $W_1 = \begin{bmatrix} -1 & 2 \\ -5 & 0 \end{bmatrix}, V_1 = \begin{bmatrix} 0 & 4 \end{bmatrix}, W_2 = \begin{bmatrix} 0 & -6 \\ 4 & 0 \end{bmatrix}, V_2 = \begin{bmatrix} 0 & 4 \end{bmatrix}$, and member functions

$$A_1^j = \exp\left(\frac{-(w_1 - a_j)^2}{2\sigma_1^2}\right) \bigg/ \left[\exp\left(\frac{-(w_1 - 1.2)^2}{2\sigma_1^2}\right) + \exp\left(\frac{-(w_1 - 1)^2}{2\sigma_2^2}\right)\right]$$

where $a_1 = 1.2, a_2 = 1, \sigma_1^2 = 0.5, \sigma_2^2 = 1$.

Based on the results of Guo and Chen (2005), we can get a candidate value of R_3:

$$R_3 = \begin{bmatrix} 0 & 50.6065 & -12.7545 & 0 & 0 \\ 0 & -368.9759 & 5.5669 & 0 & 0 \end{bmatrix}$$

By solving inequalities (18), (19) with R_3, we get

$$K = \begin{bmatrix} 0.0003 & 0.4188 & 0.1418 & 0.3739 & 0.0000 \end{bmatrix}$$
$$L = \begin{bmatrix} 0 & 0.0368 & -0.0145 & 0 & 0 \\ 0 & -0.4384 & 0.0059 & 0 & 0 \end{bmatrix}$$

The initial values of system are chosen as $w(0) = [2, 1]^T, x(0) = [0.3, -1.1, 1.4, 0]^T$ and $y_d = 1$.

Figure 1 displays the irregular nonlinear disturbance and its estimation as well as the estimation error, which illustrates the satisfactory tracking ability of the designed T-S fuzzy disturbance observer. Figure 2 shows the trajectories and stability of system states. When the composite controller is applied, the good tracking performance of system output is shown in Fig. 3. It can be seen that the scheme proposed in this paper guarantees system with nonlinear disturbance and input saturation has the satisfactory disturbance tracking ability and the good tracking performance of output, simultaneously.
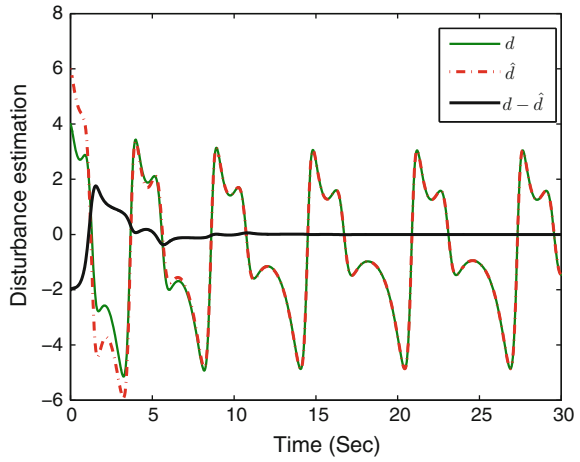
Fig. 1 Nonlinear disturbance and its estimation value

Fig. 2 The state trajectory of A4D systems
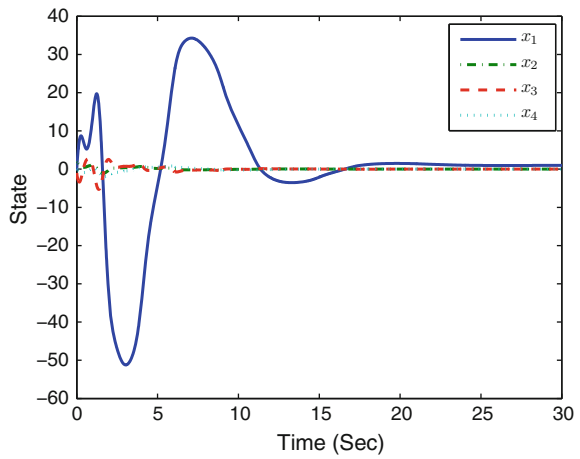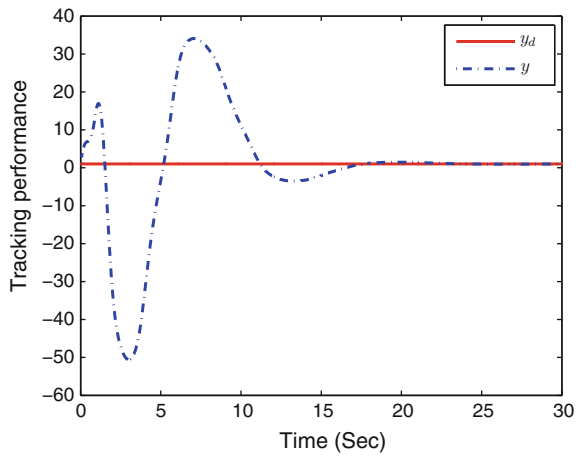

Fig. 3 The trajectory of system output

5   Conclusion

In this paper, the disturbance observer based PI tracking control framework is proposed for a class of nonlinear systems subject to input saturation and unknown irregular disturbances. T-S fuzzy disturbance observer is designed to estimate various complex disturbances described by T-S fuzzy models. Based on LMI convex optimization algorithm, the system stability and the tracking performance are guaranteed with an estimation of domain of attraction. Finally, a simulation example on A4D aircraft model demonstrates the effectiveness of the obtained results.

References

Chen, W. H., Balance, D. J., & Gawthrop, P. J. (2000). A nonlinear disturbance observer for robotic manipulators. *IEEE Transactions on Industrial Electronics, 47*(4), 932–938.

Guo, L., & Chen, W. H. (2005). Disturbance attenuation and rejection for systems with nonlinearity via DOBC approach. *International Journal of Robust and Nonlinear Control, 15*, 109–125.

Hu, T., & Lin, Z. (2001). *Control systems with actuator saturation: Analysis and design*. Boston, MA: Birkhäuser.

Hu, T., Lin, Z., & Chen, B. M. (2002). Analysis and design for discrete-time linear systems subject to actuator saturation. *Systems & Control Letters, 45*, 97–112.

Isidori, A., & Byrnes, C. I. (1990). Output regulation of nonlinear systems. *IEEE Transactions on Automatic Control, 35*(2), 131–140.

Luo, Y. H., & Zhang, H. G. (2008). Approximate optimal control for a class of nonlinear discrete-time systems with saturating actuators. *Progress in Natural Science, 18*(8), 1023–1029.

Marino, R., & Tomei, P. (1995). *Nonlinear control design: Geometric adaptive and robust*. Englewood Cliffs: Prentice Hall.

Qiu, J., Feng, G., & Gao, H. (2011). Nonsynchronized-state estimation of multichannel networked nonlinear systems with multiple packet dropouts via T-S fuzzy-affine dynamic models. *IEEE Transactions on Fuzzy Systems, 19*(1), 75–90.

Schaft, A. J. (1992). L2-gain analysis of nonlinear systems and nonlinear state feedback H_∞ control. *IEEE Transactions on Automatic Control, 37*(6), 770–784.

Sun, H., & Guo, L. (2014). Composite adaptive disturbance observer based control and back-stepping method for nonlinear system with multiple mismatched disturbances. *Journal of the Franklin Institute, 351*, 1027–1041.

Takagi, T., & Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man and Cybernetics, 15*(1), 116–132.

Tong, S. C., & Li, Y. M. (2012). Adaptive fuzzy output feedback tracking backstepping control of strict-feedback nonlinear systems with unknown dead zones. *IEEE Transactions on Fuzzy Systems, 20*(1), 168–180.

Tong, S. C., Li, Y. M., & Shi, P. (2012). Observer-based adaptive fuzzy backstepping output feedback control of uncertain MIMO pure-feedback nonlinear systems. *IEEE Transactions on Fuzzy Systems, 20*(4), 771–785.

Wei, Y., Zheng, W. X., & Xu, S. (2015). Anti-disturbance control for nonlinear systems subject to input saturation via disturbance observer. *Systems Control letters, 85*, 61–69.

Yang, J., Li, S. H., Su, J. Y., & Yu, X. H. (2013). Continuous nonsingular terminal sliding mode control for systems with mismatched disturbances. *Automatica, 49*(7), 2287–2291.

Yang, Z., & Tsubakihara, H. (2008). A novel robust nonlinear motion controller with disturbance observer. *IEEE Transactions on Control Systems Technology, 16*(1), 137–147.

Yao, X. M., & Guo, L. (2013). Composite anti-disturbance control for Markovian jump nonlinear systems via disturbance observer. *Automatica, 49*(8), 2538–2545.

Yi, Y., Zhang, T. P., & Guo, L. (2009). Multi-objective PID control for non-Gaussian stochastic distribution system based on two-step intelligent models. *Science in China-Series F: Information Sciences, 52*(10), 1754–1765.

Zhang, H. B., Shen, Y. Y., & Feng, G. (2007a). Delay-dependent stability and H_∞ control for a class of fuzzy descriptor systems with time delay. *Fuzzy Sets and Systems, 160*(12), 1689–1707.

Zhang, H. G., Yang, D. D., & Chai, T. Y. (2007b). Guaranteed cost networked control for T-S fuzzy systems with time delays. *IEEE Transactions on Systems, Man and Cybernetics C, 37*(2), 160–172.

Zhang, K., Jiang, B., & Shi, P. (2012). Fault estimation observer design for discrete-time Takagi-Sugeno fuzzy systems based on piecewise Lyapunov functions. *IEEE Transactions on Fuzzy Systems, 20*(1), 192–200.

Zuo, Z., Ho, D. W. C., & Wang, Y. (2010). Fault tolerant control for singular systems with actuator saturation and nonlinear perturbation. *Automatica, 46*, 569–576.

Application of H-Infinity Output-Feedback Control with Analysis of Weight Functions and LMI to Nonlinear Nuclear Reactor Cores

Gang Li, Bin Liang, Xueqian Wang, Xiu Li and Bo Xia

Abstract This research is to deal with the nonlinear control issue for power regulations of the pressurized water reactor core in nuclear power plants. Based on modeling a nonlinear pressurized water reactor core using the lumped parameter method, its linearized model is achieved through the small perturbation linearization way. The H_∞ output-feedback control with analysis of weight functions and linear matrix inequalities solving method is used to contrive a robust controller of the linearized core model. Besides, general laws of selecting weight functions in H_∞ control are summarized. The solved H_∞ output-feedback controller is applied to the nonlinear core model. The nonlinear core model and the H_∞ controller construct the nonlinear core power H_∞ control system. Eventually, the nonlinear core power H_∞ control system is simulated, and simulation results show that the nonlinear control system is effective.

Keywords Reactor core · H_∞ control · Weight function · LMI

G. Li · B. Liang (✉) · X. Wang · X. Li · B. Xia
National Laboratory for Information Science and Technology, Tsinghua University,
Haidian District, Beijing, China
e-mail: wychug@163.com

G. Li · B. Liang · X. Wang · X. Li · B. Xia
Graduate School at Shenzhen, Tsinghua University, University Town, Shenzhen, China

G. Li · B. Liang · X. Wang · X. Li · B. Xia
Department of Automation, Tsinghua University, Haidian District, Beijing, China

G. Li · B. Liang · X. Wang · X. Li · B. Xia
Shenzhen Key Lab of Space Robotic Technology and Telescience, University Town,
Shenzhen, China

G. Li · B. Liang · X. Wang · X. Li · B. Xia
Shenzhen Engineering Lab of Precision Geometry Measurement Technology,
University Town, Shenzhen, China

1   Introduction

Energy generated by nuclear power plants (NPPs) occupies an important part of the whole electricity production in the world, and NPPs will undergo sustainable development for future electricity production. In developing technologies of NPPs, researching and improving the control technology of core power in NPPs is necessary for avoiding security risks and enlarging economic potential.

With the advancement of control theories over the decades, the sophisticated and desired control methods have been established, which give a bright future to the reactor core power control. Edwards et al. (1990) designed the controllers under the state feedback assisted control structure to provide the tight control of nuclear reactors. But in Edwards et al. (1990), the core power linear control system is only researched and a relative nonlinear control system is ignored; besides, the controllers designed by using the pole placement technology do not have a good robustness and are not always optimal or even ineffective in response to drastic load maneuvers. However, the H_∞ control is a norm-based robust optimal control strategy proposed by Zames (1981) and Francis (1987). Robustness of this control includes performance robustness as well as stability robustness. This robustness can guarantee applications of H_∞ controllers to a nonlinear plant.

In the paper, the H_∞ output-feedback control strategy (Zames 1981; Francis 1987) with the linear matrix inequalities (LMI) solving method (Yu 2002) is adopted for nonlinear core power regulations of a pressurized water reactor (PWR) NPP. Besides, general laws of selecting weight functions in H_∞ control are summarized as well. Finally, the designed nonlinear PWR core power H_∞ control system is simulated, and conclusions are drawn.

2   Modeling of PWR Core

The point reactor core modeling (Schultz 1961; Kerlin et al. 1976) as a traditional modeling method is adopted to model a PWR core, which is a lumped parameter method. For this modeling, a core is regarded as one point without any space profile, and parameters of the core only vary with time and have nothing to do with space positions. In the research field for reactor control, such modeling is popular (Edwards et al. 1990; Ben-Abdennour et al. 1992; Arab-Alibeik and Setayeshi 2003; Ansarifar and Akhavan 2015).

According to the point reactor core modeling, the nonlinear PWR core is modeled adopting the point kinetics equations with six groups of delayed neutrons and reactivity feedbacks due to control rod movement and variations in fuel temperature and coolant temperature. Main model parameters are given in Table 1. The nonlinear core model is showed as Eq. (1).

Table 1 Main model parameters

Parameter	Name
P_r	Core power level
c_{ri}	ith group normalized precursor concentration
c_r	Normalized precursor concentration
T_f	Fuel average temperature
T_i	Coolant inlet temperature
T_e	Coolant outlet temperature
δrod	Position variation of the control rod (fraction of core length)
δ	Deviation of a parameter from initial steady-state value

$$
\begin{cases}
\frac{dP_r}{dt} = \frac{\rho - \beta}{\Lambda} P_r + \sum_{i=1}^{6} \frac{\beta_i c_{ri}}{\Lambda} \\
\frac{dc_{ri}}{dt} = \lambda_i P_r - \lambda_i c_{ri}, \; i = 1, 2, \ldots, 6 \\
\frac{dT_f}{dt} = \frac{f_f P_0}{\mu_f} P_r - \frac{\Omega}{\mu_f} T_f + \frac{\Omega}{2\mu_f} T_i + \frac{\Omega}{2\mu_f} T_e \\
\frac{dT_e}{dt} = \frac{(1-f_f)P_0}{\mu_c} P_r + \frac{\Omega}{\mu_c} T_f + \frac{2M-\Omega}{2\mu_c} T_i - \frac{2M+\Omega}{2\mu_c} T_e \\
\rho = \rho_{rod} + \alpha_f(T_f - T_{f0}) + \frac{\alpha_c}{2}(T_i - T_{i0}) + \frac{\alpha_c}{2}(T_e - T_{e0}) \\
\delta\rho_{rod} = \alpha_{rod}\delta rod
\end{cases}
\tag{1}
$$

One group delayed neutron model is utilized and the coolant inlet temperature is treated as a constant (Ben-Abdennour et al. 1992; Arab-Alibeik and Setayeshi 2003). The small perturbation linearization methodology is utilized to linearize the nonlinear core model (1). The transfer function and the state equation of the core are separately calculated and expressed by

$$
G = \frac{\delta P_r}{\delta rod} = \frac{\sum_{i=0}^{3} a_i s^i}{\sum_{i=0}^{4} b_i s^i}.
\tag{2}
$$

$$
\begin{cases}
\dot{\mathbf{x}} = A\mathbf{x} + Bu \\
y = C\mathbf{x} + Du
\end{cases}.
\tag{3}
$$

where $u = \delta rod$-the input; $y = \delta P_r$-the output; a_i ($i = 0, 1, 2, 3$)-numerator coefficients; b_i ($i = 0, 1, 2, 3, 4$)-denominator coefficients; s-laplace operator; $\mathbf{x} = [x_1, x_2, x_3, x_4]^T = [\delta P_r, \delta c_r, \delta T_f, \delta T_e]^T$-the state matrix; A-the $R^{4\times4}$ system matrix; B-the $R^{4\times1}$ input matrix; C-the $R^{1\times4}$ output matrix; D-the zero matrix.

A set of parameters of the PWR core is shown in Table 2.

Table 2 A set of core parameters

Item	Value	Item	Value
P_0 (W)	2.2×10^9	α_f (°C^{-1})	-2.9×10^{-5}
f_f	0.95	α_c (°C^{-1})	-6.3×10^{-4}
λ (s^{-1})	0.15	μ_f (J/°C)	2.25×10^7
β	0.0064	Ω (W/°C)	3.94×10^6
Λ (s)	1.6×10^{-5}	μ_c (J/°C)	6.88×10^7
α_{rod}	0.0814	M (W/°C)	7.06×10^7

3   H∞ Output-Feedback Control with Weight Functions and LMI for Core

On the basis of the small gain theorem and analyzing H∞ norms of the sensitivity and complementary sensitivity, the H∞ output-feedback control strategy for the core is achieved adopting suitable weight functions and the LMI solving method (Gahinet et al. 1995).

3.1   H∞ Control Theory

Tracking control schematic of G with disturbances is shown in Fig. 1, where $K(s)$ is a H∞ controller; d is the measurement disturbance; $L(s) = G(s)K(s)$, the open-loop transfer function of this control system.

The closed-loop transfer functions from u_1 to e, u, y_1 are expressed by

$$\begin{cases} S(s) = \frac{e(s)}{u_1(s)} = [I + L(s)]^{-1} \\ R(s) = \frac{u(s)}{u_1(s)} = K_i(s)[I + L(s)]^{-1} = K_i(s)S(s) \, . \\ T(s) = \frac{y_1(s)}{u_1(s)} = L(s)[I + L(s)]^{-1} = I - S(s) \end{cases} \quad (4)$$

where $S(s)$ is the sensitivity function and $T(s)$ is the complementary sensitivity function.

In terms of Figs. 1 and 2 is obtained by introducing three weight functions. The feedback control schematic of G with weight functions is shown in Fig. 2, where $W_i(s)$ ($i = 1, 2, 3$) are three weight functions; u_1 is a reference input; $u = u_2$, a control input of G, namely δrod; $\mathbf{y_1} = [y_{1a} \; y_{1b} \; y_{1c}]^T$; $y_2 = e$, the tracking error; $G_a(s)$ is a generalized plant.

The state equation of $G_a(s)$ is expressed by Eq. (5).

$$\begin{cases} \dot{\mathbf{x}} = A\mathbf{x} + B_1 u_1 + B_2 u_2 \\ \mathbf{y_1} = C_1 \mathbf{x} + D_{11} u_1 + D_{12} u_2 \, . \\ y_2 = C_2 x + D_{21} u_1 + D_{22} u_2 \end{cases} \quad (5)$$
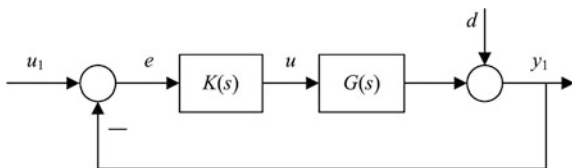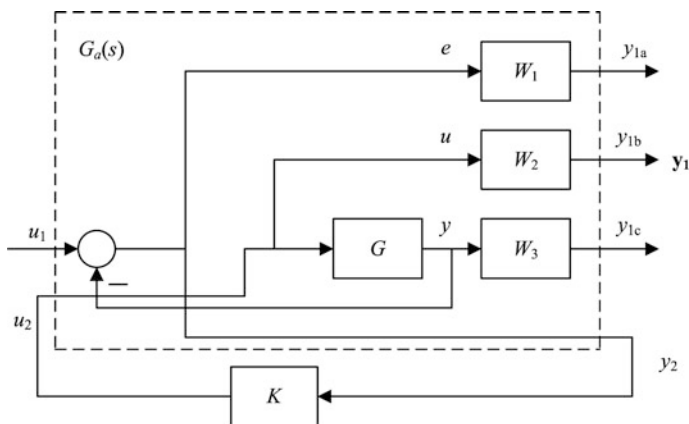
Fig. 1 Tracking control schematic of G with disturbances

Fig. 2 Feedback control schematic of G with weight functions

where $\mathbf{x} \in R^{5\times 1}$, $\mathbf{y_1} \in R^{3\times 1}$, $y_2 \in R^{1\times 1}$; A, B_1, B_2, C_1, C_2, D_{11}, D_{12}, D_{21}, D_{22} are 5×5, 5×1, 5×1, 3×5, 1×5, 3×1, 3×1, 1×1, 1×1 real matrices.

The closed-loop transfer function from u_1 to $\mathbf{y_1}$ is Eqs. (6) or (7).

$$T_{\mathbf{y_1}u_1}(s) = G_{a11} + G_{a12}K(I - G_{a22}K)^{-1}G_{a21} . \tag{6}$$

$$T_{\mathbf{y_1}u_1}(s) = \begin{bmatrix} W_1 S \\ W_2 R \\ W_3 T \end{bmatrix} . \tag{7}$$

where $G_a(s) = [G_{a11}\ G_{a12}; G_{a21}\ G_{a22}]$.

Definition 3.1 (*H_∞ suboptimal design*) For $G_a(s)$ and $\gamma > 0$, $K(s)$ is designed such that the closed-loop system (G_a with K) is stable and

$$\left\| T_{\mathbf{y_1}u_1}(s) \right\|_\infty < \gamma. \tag{8}$$

3.2 Selection of Weight Functions

$\|S(s)\|_\infty$ is a measurement of inhibiting disturbances, $\|R(s)\|_\infty$ one of additive uncertainty, $\|T(s)\|_\infty$ one of multiplicative uncertainty. $W_1(s)$ is the performance weight function, $W_2(s)$ is the controller-output one, $W_3(s)$ is the robustness one. Namely, $W_1(s)$ is a weight function of $S(s)$, $W_2(s)$ is one of $R(s)$, $W_3(s)$ is one of $T(s)$. The use of the functions is to regulate singular values of $S(s)$, $R(s)$ and $T(s)$ through the forms $\|W_1(s)S(s)\|_\infty$, $\|W_2(s)R(s)\|_\infty$ and $\|W_3(s)T(s)\|_\infty$ within corresponding frequency ranges.

In order to make sure that the control system $G_a(s)$ with $K(s)$ possesses the desired ability of inhibiting disturbances and tracking a target and is stable under the structure perturbation and parameter perturbation of $G(s)$, the singular value of $S(s)$ and one of $T(s)$ should be minimum simultaneously. As can be seen in Eq. (4), it is difficult and contradictory to make the singular value of $S(s)$ and one of $T(s)$ minimum simultaneously. Hence, $W_1(s)$ and $W_3(s)$ are selected by a compromise way. Generally, the issue is dealt with by making the singular value of $S(s)$ be as small as possible in a low-frequency range and making the singular value of $T(s)$ be as small as possible in a high-frequency range such that

$$\begin{cases} \bar{\sigma}(S(jw)) \le \bar{\sigma}\left[W_1^{-1}(jw)\right] \\ \frac{1}{\bar{\sigma}(S(jw))} \le |W_1(jw)| \end{cases}. \tag{9}$$

$$\bar{\sigma}(R(jw)) < \bar{\sigma}\left[W_2^{-1}(jw)\right]. \tag{10}$$

$$\begin{cases} \bar{\sigma}(T(jw)) < \bar{\sigma}\left[W_3^{-1}(jw)\right] \\ \frac{1}{\bar{\sigma}(T(jw))} \ge |W_3(jw)| \end{cases}. \tag{11}$$

where $\bar{\sigma}(\cdot)$ is the maximum singular value of $(\cdot)$; $W_1(s)$ and $W_3(s)$ should meet Eq. (12) such that the frequency band of $W_1(s)$ and one of $W_3(s)$ do not have an overlap.

$$\bar{\sigma}\left[W_1^{-1}(jw)\right] + \bar{\sigma}\left[W_3^{-1}(jw)\right] \ge 1. \tag{12}$$

Consequently, $|W_1(s)|$ is as large as possible in a low-frequency band and as small as possible in a high-frequency band; $|W_2(s)|$ is decent such that the use of it can limit control energy and guarantee that the control system has enough bandwidth; $|W_3(s)|$ is as small as possible in a low-frequency band and as large as possible in a high-frequency band.

D_{12} and D_{21} need to be full column rank and full row rank in turn such that K(s) to be solved is a real rational function for SISO plants. $G(s)$, $W_1(s)$, $W_2(s)$ and $W_3(s)$ *$G(s)$ are bounded when $s \to \infty$.

According to G and three weight functions, Eq. (5) is obtained to solve $K(s)$.

3.3 LMI Based Solution for H_∞ Controller

The H_∞ suboptimal design is utilized to solve the H_∞ output-feedback controller K (s) by the LMI toolbox (Gahinet et al. 1995).

The state equation of $K(s)$ is denoted by $(A_K \, B_K \, C_K \, D_K)$. The state equation of the H_∞ closed-loop control system (G_a with K) is Eq. (13).

$$
\begin{cases}
\begin{bmatrix} \dot{\mathbf{x}} \\ \dot{\mathbf{x_k}} \end{bmatrix} = A_c \begin{bmatrix} \mathbf{x} \\ \mathbf{x_k} \end{bmatrix} + B_c u_1 \\[3mm]
\mathbf{y_1} = C_c \begin{bmatrix} \mathbf{x} \\ \mathbf{x_k} \end{bmatrix} + D_c u_1
\end{cases}
. \tag{13}
$$

where $\mathbf{x_K}$ is the state array of $K(s)$; $A_c = [A + B_2 D_K C_2 \quad B_2 C_K ; B_K C_2 \quad A_K]$, $B_c = [B_1 + B_2 D_K D_{21} \quad B_K D_{21}]^{\mathrm{T}}$, $C_c = [C_1 + D_{12} D_K C_2 \quad D_{12} C_K]$, $D_c = D_{11} + D_{12} D_K D_{21}$.

According to Bounded Real Lemma (Yu 2002), $K(s)$ makes the system (13) asymptotically stable and meet Eq. (8) under the necessary and sufficient condition that there exists a symmetric positive definite matrix X_c such that

$$
\begin{bmatrix}
A_c^T X_c + X_c A_c & X_c B_c & C_c^T \\
B_c^T X_c & -\gamma I & D_c^T \\
C_c & D_c & -\gamma I
\end{bmatrix} < 0. \tag{14}
$$

In Eq. (14), X_c and A_K, B_K, C_K, D_K are shown in a nonlinear relationship, and it is difficult to calculate them by the LMI toolbox. Hence, on the basis of Eq. (14), the elimination method (Yu 2002) is adopted to deduce Eq. (15). A_K, B_K, C_K, D_K and γ can be obtained using the LMI toolbox to solve Eq. (15).

$\min \gamma$

$$
\gamma, \Theta, \Xi \; s.t.
\begin{cases}
\begin{bmatrix} N_O & 0 \\ 0 & I \end{bmatrix}^T
\begin{bmatrix}
A^T \Theta + \Theta A & \Theta B_1^T & C_1^T \\
B_1^T \Theta & -\gamma I & D_{11}^T \\
C_1 & D_{11} & -\gamma I
\end{bmatrix}
\begin{bmatrix} N_O & 0 \\ 0 & I \end{bmatrix} < 0 \\[6mm]
\begin{bmatrix} N_C & 0 \\ 0 & I \end{bmatrix}^T
\begin{bmatrix}
A\Xi + \Xi A^T & \Xi C_1^T & B_1 \\
C_1 \Xi & -\gamma I & D_{11} \\
B_1^T & D_{11}^T & -\gamma I
\end{bmatrix}
\begin{bmatrix} N_C & 0 \\ 0 & I \end{bmatrix} < 0 \\[6mm]
\begin{bmatrix} \Theta & I \\ I & \Xi \end{bmatrix} \geq 0
\end{cases}
. \tag{15}
$$

where Θ, Ξ are two symmetric positive definite matrices; N_O (N_C) is a matrix constructed by using a random set of basis vectors in $\ker([C_2 \; D_{21}])$ ($\ker([B_2^{\mathrm{T}} \; D_{12}^{\mathrm{T}}])$) as column vectors, and $\ker()$ is a kernel space.

4 Design of Nonlinear Core Power H∞ Control System

According to the aforementioned H∞ output-feedback control principle, the controller K of G is solved by programming after selecting decent weight functions.

W_1, W_2, W_3, γ of G are shown in Table 3. Singular value plots of S, R, T and weight functions for the frequency $w \in [10^{-2}\ 10^8]$ are shown in Fig. 3, where C1 is the singular value curve of S, C2 is one of W_1^{-1}; C3 is one of R, C4 is one of W_2^{-1}; C5 is one of T, C6 is one of W_3^{-1}; C7 is obtained by the addition of singular values of W_1^{-1} and singular values of W_3^{-1}. As can been seen in Fig. 3, for a frequency, a singular value of W_1^{-1} is greater than one of S, a singular value of W_2^{-1} is greater than one of R, a singular value of W_3^{-1} is greater than one of T, a value of C7 is

Table 3 A set of core parameters

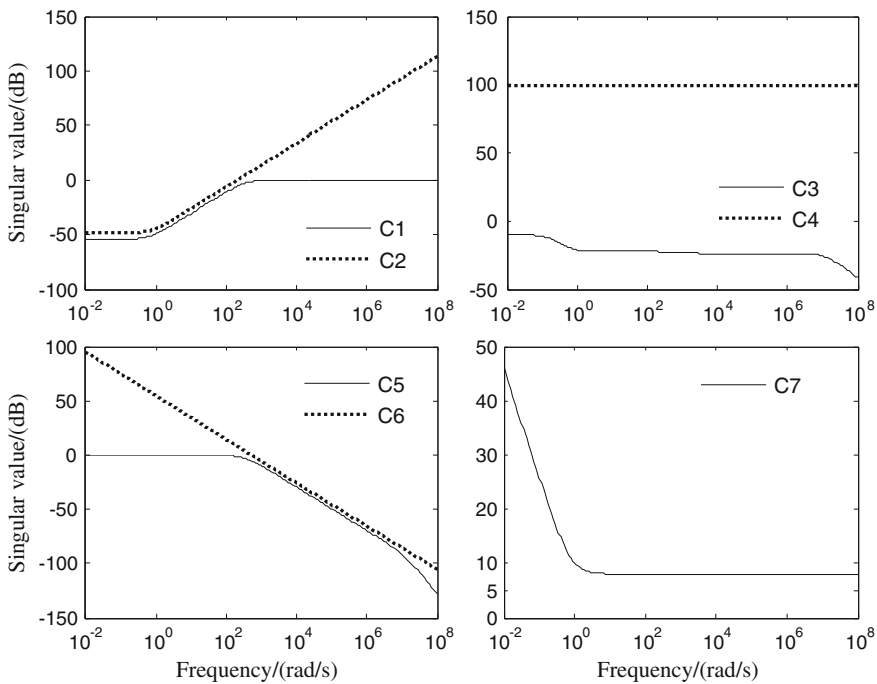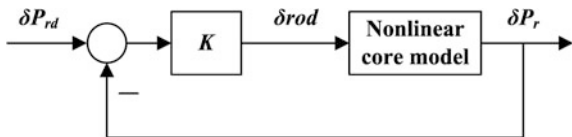| Item	Value
W_1	$263.89/(1.26 \times s + 0.98)$
W_2	1.32×10^{-5}
W_3	$1.25 \times s/663.18$
γ	0.6280

Fig. 3 Singular-value plots of S, R, T and weight functions for G with K

Fig. 4 Schematic of the nonlinear core power H$_\infty$ control system

greater than 1. These indicate that Eqs. (9)–(12) hold. Therefore, K is desirable. K is represented by Eq. (16).

$$K = \sum_{i=0}^{4} c_i s^i \bigg/ \sum_{i=0}^{5} d_i s^i. \tag{16}$$

where $[c_4, c_3, \ldots, c_0] = [8.427\mathrm{e}5 \quad 3.383\mathrm{e}8 \quad 5.833\mathrm{e}8 \quad 3.286\mathrm{e}8 \quad 3.641\mathrm{e}7]$; $[d_5, d_4, \ldots, d_0] = [1 \quad 1.288\mathrm{e}7 \quad 2.586\mathrm{e}7 \quad 1.585\mathrm{e}7 \quad 3.293\mathrm{e}6 \quad 2.18\mathrm{e}5]$, $1\mathrm{e}j = 10^j$ and j is an integer.

The solved controller K is applied to the nonlinear core model. The combination of the nonlinear core model and K is the nonlinear core power H$_\infty$ control system. The schematic of the nonlinear core power H$_\infty$ control system is shown in Fig. 4, where δP_{rd} is the reference input, namely the desired change of power level.

5 Simulation

The nonlinear core power H$_\infty$ control system is simulated considering two typical load variations that are the step process with ± 10 % variations and the ramp process with ± 6 %·min^{-1} variations. The direction from the bottom to the top of core is treated as the positive direction of y-axis for simulation work, the core bottom is the zero point of y-axis and the height of core is 366 cm. When the core power level is at 100 %, the initial rod position is at the top of the core, namely at 366 cm.

When the reference input δP_{rd} is taken as a step load with ± 10 % variations, change trajectories of core parameters during the core following the step load (100 % → 90 % → 100 %) are shown in Fig. 5, where SL1 is the desired load change trajectory for the start of reactor full power level operation, the middle of reactor shut-down operation with a -10 % step variation, 90 % power level operation and start-up operation with a 10 % step variation, and the end of reactor full power level operation again; SL2 is a change trajectory of the core power level; SL3 is a change trajectory of the control rod position in the core.

When the reference input δP_{rd} is taken as a ramp load with ± 6 %·min^{-1} variations, change trajectories of core parameters during the core following the ramp load (100 % → 90 % → 100 %) are shown in Fig. 6, where RL1 is the desired load change trajectory for the start of reactor shut-down operation with a -6 %·min^{-1} ramp variation, the middle of 90 % power level operation, and the end of start-up operation with a 6 %·min^{-1} ramp variation; RL2 is a change trajectory of
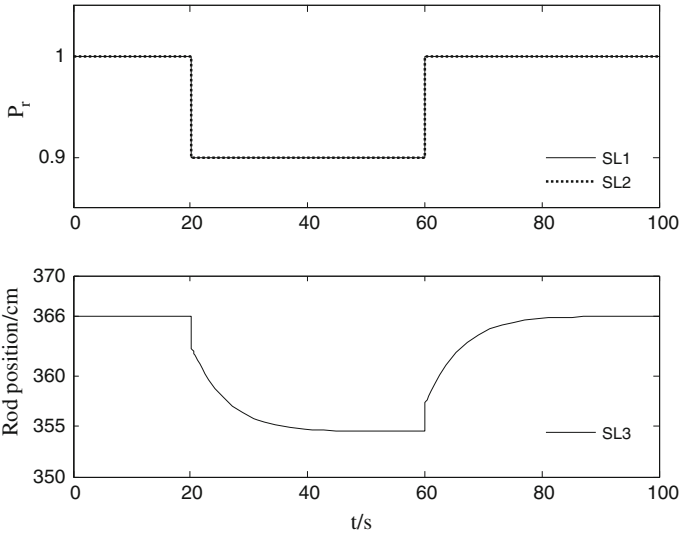
Fig. 5 Change trajectories of core parameters during the core following a step load
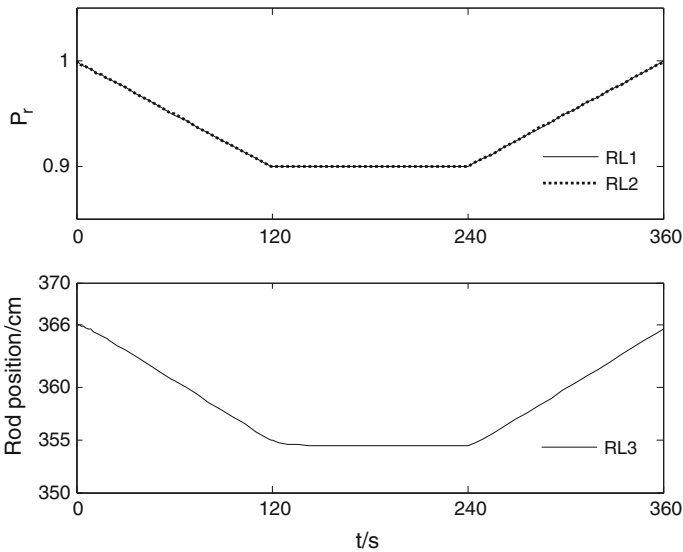

Fig. 6 Change trajectories of core parameters during the core following a ramp load

the core power level; RL3 is a change trajectory of the control rod position in the core.

In terms of Figs. 5 and 6, in reducing the desired load, the control input produced by this nonlinear control system drives the control rod to move with the

inserting direction in the core, and the movement of the rod makes the core power level be decreased by the core satisfactorily following the desired load; in keeping the desired load unchanged, the control input generated by this nonlinear control system keeps the control rod still so that the core power level are retained constant by the core satisfactorily following the desired load; in increasing the desired load, the control input provided by this nonlinear control system drives the control rod to move with the extracting direction, and the movement of the rod makes the core power level be augmented by the core satisfactorily following the desired load.

Consequently, SL2 (RL2) can follow SL1 (RL1) in real time; meanwhile, the processes of SL3 and RL3 are suitable and acceptable.

6 Conclusions

The nonlinear control problem for PWR core regulations is handled modeling the nonlinear model and linear one of this core, and applying the H_∞ output-feedback control with analysis of weight functions and LMI method.

The lumped parameter method and the small perturbation linerization method are used to model the core. Such modeling can be a reference approach for modeling nonlinear systems. In order to implement applications of the controller designed on a linear model to a nonlinear model, the H_∞ controller designed by introducing the LMI method has strong robustness of accommodating model errors, and this design process with LMI is rapid and effective.

The nonlinear PWR core power H_∞ control system possesses the abilities of tracking a target, limiting control energy and guaranteeing robust stability, and is capable of controlling the core power satisfactorily in response to different load changes such as a step or ramp which involve the operations of full power, start-up and shut-down. Hence, the nonlinear PWR core power H_∞ control system is effective.

References

Arab-Alibeik, H., & Setayeshi, S. (2003). Improved temperature control of a PWR nuclear reactor using an LQG/LTR based controller. *IEEE Transactions on Nuclear Science, 50*(1), 211–218.
Ansarifar, G. R., & Akhavan, H. R. (2015). Sliding mode control design for a PWR nuclear reactor using sliding mode observer during load following operation. *Annals of Nuclear Energy, 75*, 611–619.

Ben-Abdennour, A., Edwards, R. M., & Lee, K. Y. (1992). LQG/LTR robust control of nuclear reactors with improved temperature performance. *IEEE Transactions on Nuclear Science, 39* (6), 2286–2294.

Edwards, R. M., Lee, K. Y., & Schultz, M. A. (1990). State feedback assisted classical control: An incremental approach to control modernization of existing and future nuclear reactors and power plants. *Nuclear Technology, 92*(2), 167–185.

Francis, B. A. (1987). *A course in H_∞ control theory*. New York: Springer.

Gahinet, P., Nemirovski, A., Laub, A. J., & Chilali, M. (1995). *LMI control toolbox user's guide*. The MathWorks: Natick.

Kerlin, T. W., Katz, E. M., Thakkar, J. G., & Strange, J. E. (1976). Theoretical and experimental dynamic analysis of the H. B. Robinson nuclear plant. *Nuclear Technology, 30*(3), 299–316.

Schultz, M. A. (1961). *Control of nuclear reactors and power plants*. New York: McGraw-Hill.

Yu, L. (2002). *Robust control—Linear matrix inequalities method*. Beijing: Tsinghua University Press.

Zames, G. (1981). Feedback and optimal sensitivity: Model reference transformations, multiplicative seminorms, and approximate inverses. *IEEE Transactions on Automatic control*. *AC-26*(2), 301–320.