

A cooperative robotic system based on multiple sensors to construct metallic structures

P. Gil · J. Pomares · S. Puente · F. A. Candelas ·
G. J. García · J. A. Corrales · F. Torres

Received: 29 October 2008 / Accepted: 25 February 2009 / Published online: 18 March 2009
© Springer-Verlag London Limited 2009

Abstract This paper describes a multisensorial robotic system to automatically construct metallic structures. Two robots must work cooperatively in the same workspace to perform the task. The robots are automatically guided using visual and force sensor information. A new time-independent visual force control system which guarantees the adequate robot behaviour during the construction of the structure is described. During the construction of the structure, a human operator works cooperatively with the robots in order to perform some tasks which cannot be automatically developed by the robots. To do so, a new human–robot cooperation approach is described in order to guarantee the human safety. The correct behaviour of the different subsystems proposed in the paper is demonstrated in Section 6 by the construction of a real structure composed of several metallic tubes and different types of pieces to join them.

Keywords Visual servoing · Visual force control · Human–robot interaction · Human tracking

1 Introduction

Automatic assembly processes involve different disciplines such as assembly sequence generation, assembly interpretation, robot positioning techniques based on different sensors and handling of objects of the assembly.

Different approaches in assembly–disassembly sequence generation, planning in real time in the manufacturing environment and techniques for assembly–disassembly interpretation have been widely developed in the last years. Directed graphs with precedence relations [1] are commonly used to determine the movement sequence required to develop the task. Generally, the products are described by databases and/or computer-aided design (CAD) models which represent their main features [2].

Traditionally, the monitoring process in assembly applications has been done by off-line processing techniques. First, the tasks to perform the assembly are planned. Later, a machine vision or another sensorial system is used to automatically check a product at the end of the assembly line, looking for the problems or defects that have occurred. Currently, visual sensory techniques are widely considered by researches for manufacturing process automation. Over the last few years, these techniques have been used for inspection and handling of objects [3], for estimation of pose with range data and three-dimensional image processing [4] or with stereo vision [5]. However, the use of sensorial processes can be wider than the simple inspection of the assembled product. Sensors are an important subject for an intelligent manipulation of objects in situations with a high degree of randomness in the environment. Sensors increase the ability of a robot to be adapted to its workspace. In summary, the sensors system for the assembly process should extract the necessary information about situations or events which take place during the development of the planned task. For example, image sequences from cameras can be used to describe what is happening in the scene or to determine how a robot is moving. The sensors which detect if unexpected events appear during the development of the tasks, such as the presence of a human who crosses the trajectory planned for

P. Gil (✉) · J. Pomares · S. Puente · F. A. Candelas ·
G. J. García · J. A. Corrales · F. Torres
Physics, Systems Engineering and Signal Theory Department,
University of Alicante,
P.O. Box 99, 03080 Alicante, Spain
e-mail: Pablo.Gil@ua.es

a robot, are also important. In previous works, some basic techniques of visual servoing and visual detection of target using image processing [6] have been applied to monitor the movement of a robot and guide it to achieve the suitable position in the disassembly and manipulation processes. In addition, nowadays, human–robot interaction is used more and more to help in the modelling and localisation of objects [7], and control techniques are applied more to manipulate and insert objects [8] in assembly processes. In the near future, any manufacturing system will have to integrate a multisensorial system into the work cell.

In this paper, a multisensorial assembly system to construct metallic structures is proposed. This system presents two important advantages over classic assembly systems, mainly due to the interaction between the human and the robots. Firstly, in this system, the human will perform assistance tasks in the manipulation and positioning of objects. These tasks are monitored with a localisation system based on several sensors. This localisation system permits avoiding collisions between the human operator and the robots when they share the environment and their workspaces are crossed. Secondly, the robots are guided using a new time-independent visual force control system whilst they construct the metallic structure. This time-independent visual force control is robust and works correctly although interruptions in the trajectory take place, for example when the human crosses the trajectory planned for the movement of the robot. In addition, a visual–force interaction matrix has been modelled to guarantee the coherence between visual and force information during the tracking of trajectories.

Finally, another important aspect is the extensive use of sensors in the different phases of the task, such as inertial sensors, force sensors, radio frequency sensors, time of flight sensors and cameras. In this paper, we will show the usefulness of combining tasks of assembly and inspection between robots to perform some tasks such as the manipulation of a bar by a robot whilst another robot with a range camera detects the adapted position in the insertion task. Furthermore, a scenario, where a complex metal structure is built, has been used to prove the correct operation of an assembly process. Thus, it is demonstrated how a human interacts with robots which share a common workspace without collisions and safeguarding his integrity when they do cooperative tasks and how the robots can track a suitable trajectory although the references for the guidance are lost during a brief period of time.

This paper will be structured as follows. The first section describes the hardware that composes the architecture of the system to assembly processes. In Section 3, the main aspects of the time-independent visual control techniques are commented. In Section 4, the strategy to localise and the model of human–robot interaction is presented. After-

wards, the sequence of the needed tasks to execute an assembly process to build a metallic structure is described. Finally, in Section 6, the practical experiments to develop each task and the adopted strategies are explained. In addition, the last section shows the results and conclusions of the proposed intelligent multisensorial system.

2 Architecture

This paper proposes a subdivision of the implemented system into several subsystems. Each subsystem works in different phases of the manipulation tasks in the assembly processes, and frequently, all subsystems work jointly (Fig. 1).

- Visual servoing subsystem: This subsystem is composed of a 7 *df* Mitsubishi PA-10 robot and a PhotoFocus MV-D752-160 camera mounted on the robot's end-effector (see eye-in-hand configuration in Fig. 1). It is employed to guide the robot by using visual information. Thus, servoing techniques are applied for optimally viewing the objects and suitably approaching the objects in each relevant task. The visual servoing approaches proposed in this paper have several advantages with respect to the previous ones. These advantages are indicated in Section 3.
- Visual force control subsystem: The previous subsystem has been improved adding a JR3 force sensor. This subsystem is employed during insertion tasks to control not only the robot position but also the robot interaction force. In this case, visual force techniques are applied to measure with accuracy the contact between objects during the insertion process (this approach is described in Section 3.3). Some experiments show the validity of the visual force control system to guide the movement of the robot and control the manipulation of objects in each planned task. On the one hand, the basic task of

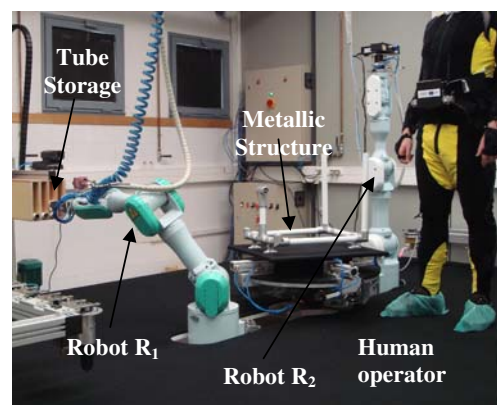


Fig. 1 System architecture

the visual servoing system is to control the pose of the robot's end-effector by using information extracted from images of the scene. On the other hand, the force information is used to control the handling and grasping of objects which are manipulated.

- Cooperation subsystem: In different phases of the task, two PA-10 robots, R_1 and R_2 , both with a force sensor in their end-effectors, are used. Both robots are equipped with a tool interchanger in order to employ the required tools during the task (gripper, robotic hand, screwdriver, 3D camera, etc.). During the assembly, the system must determine which robot performs the main task and which robot helps in its execution. The tasks to be developed by each robot are distributed using a task planner [6, 9], which determines the actions for the assembly process. Robot R_1 has a parallel gripper, a screwdriver and a vacuum available, and robot R_2 has a Barret-hand and a range camera available. The 3D-range camera (PMD-Vision 19K) is used to monitor, inspect and detect the parts of the metallic pieces which compose the structure by using image processing techniques. Furthermore, the robots share with a human the workspace in order to develop the construction of the metallic structure. Section 4 presents a human–robot interaction approach where the task of the robot is modified depending on the human location in order to

maintain a safety distance. An inertial human motion capture system (GypsyGyro-18 from Animazoo) and an indoor localisation system based on ultra-wideband pulses (UWB from Ubisense) are used to localise precisely the human operator who collaborates in the assembly task. Thereby, the movement of the human who interacts with the robot at the same workspace is tracked and the positions of all her/his limbs are determined in real time whilst the human–robot interaction takes place.

To show how each subsystem works in an assembly process, the building of a complex metallic structure has been considered. Thus, a description of the tasks of the manufacturing system to build the complex metal structure will be shown in Section 5 and the applications of each subsystem will be shown in the Section 6.

3 Visual robot guidance

In this section, an approach to guide the robot using visual information is presented. To do this, it is necessary to track previously defined trajectories using a visual servoing system which is based on an eye-in-hand camera configuration.

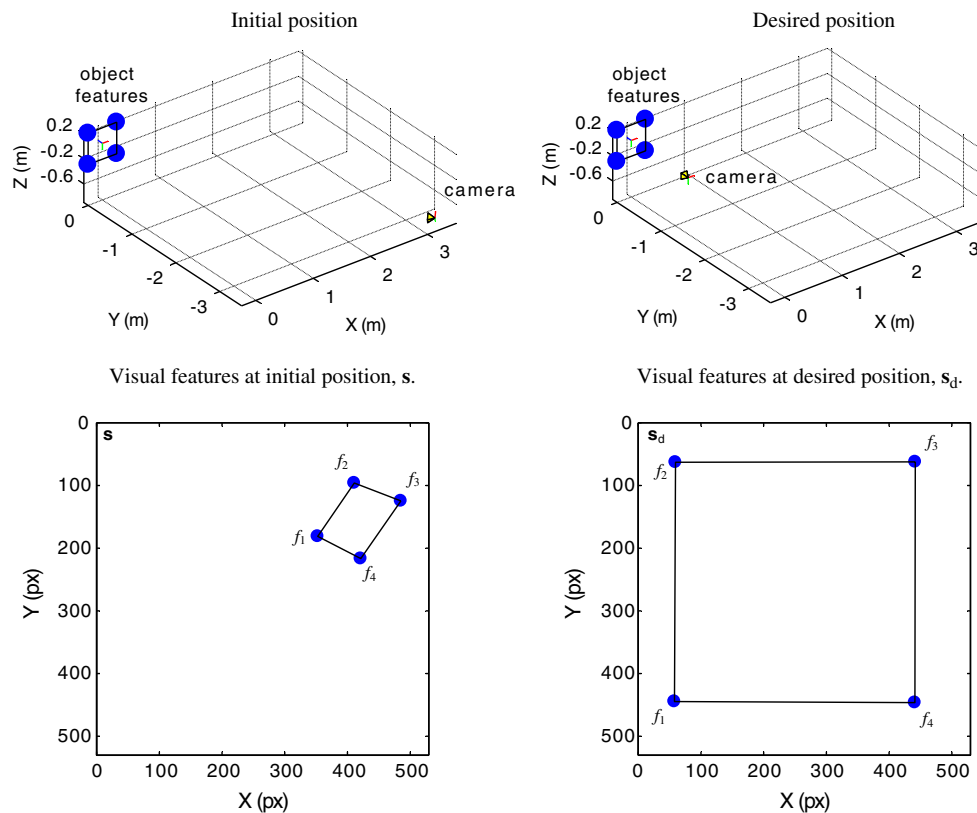


Fig. 2 Example of a visual servoing task

In a robotic task, the robot has to be frequently positioned at a fixed location with respect to the objects in the scene. However, the positions of these objects are not always controlled. Therefore, it is not possible to establish beforehand the precise location of the robot end-effector which will accomplish the task correctly. Visual servoing is a technique that allows positioning a robot with respect to an object using visual information [10].

3.1 Image-based visual servoing

Basically, the image-based visual servoing approach consists of extracting visual data from an image acquired from a camera and comparing it with the visual data obtained at the desired position of the robot. By minimising the error between the two images, it is possible to control the robot to the desired position (see the visual servoing task represented in Fig. 2).

A visual servoing task can be described by an error function \mathbf{e}_t which must be regulated to 0:

$$\mathbf{e}_t = \mathbf{s} - \mathbf{s}^* \quad (1)$$

where \mathbf{s} is a $M \times 1$ vector containing M visual features corresponding to the current state, whilst \mathbf{s}^* denotes the visual features values in the desired state. \mathbf{L}_s represents the interaction matrix which relates the variations in the image with the variation in the velocity of the camera:

$$\dot{\mathbf{s}} = \mathbf{L}_s \cdot \dot{\mathbf{r}} \quad (2)$$

where $\dot{\mathbf{r}}$ indicates the velocity of the camera.

By imposing an exponential decrease of \mathbf{e}_t ($\dot{\mathbf{e}}_t = -\lambda \mathbf{e}_t$), it is possible to obtain the following control action for a classical image-based visual servoing:

$$\mathbf{v}_c = -\lambda \hat{\mathbf{L}}_s^+ (\mathbf{s} - \mathbf{s}^*) \quad (3)$$

where $\hat{\mathbf{L}}_s^+$ is the pseudo-inverse of an estimation of the interaction matrix [10] and $\lambda > 0$ is a proportional control gain.

Image-based visual servoing uses only the visual data obtained from an image to control the robot movement. This system is adequate to position a robot from an initial point to a desired location, but it cannot control intermediate 3D positions of the end-effector. The behaviour of these systems has been proven to be robust in local conditions (i.e. in conditions in which the initial position of the robot is very near to its final location) [11]. However, in large displacements, the errors in the computation of the intrinsic parameters of the camera [12], or in the estimation of the distance to the object [13], can lead the system to a local minimum. A solution to this problem is to reach the correct location by following a desired path.

3.2 Time-independent method for tracking image paths using visual servoing

The method proposed to track trajectories in the image must present a time-independent behaviour. In a time-dependent tracking system, the references depend on the time [14]. Thereby, if the robot interacts with an object placed in its workspace, the system continues sending visual references to it even though the robot cannot move. Once the obstruction ends, the references that have been sent up to the moment are lost and the path is not tracked correctly. On the contrary, time-independent tracking systems do not lose any reference, and thus, they permit to follow the complete path.

Only two techniques have been found in the literature with this time-independent behaviour when tracking an image path using visual servoing [15, 16]. In [16], the visual servoing system proposed to track the trajectory in the image moves the robot to a position in the path using a classical image-based visual servoing control law. Once the error is cancelled (the robot velocity is 0), the system guides the robot to the next position in the path. This process is repeated until the robot has visited all the references in the path. Thus, this system is not able to maintain a constant velocity during the tracking. In [15], a visual servoing approach based on movement flow is presented. The movement flow is a vector field that indicates the direction in which the desired features to be used by an image-based visual servoing system must be located. Thus, the tracking of the trajectory is performed. This approach has several problems which must be solved in order to be applied in a real workspace. The movement flow-based visual servoing system has a correct behaviour in the 3D space. However, when the velocity is increased, this behaviour is not guaranteed. Moreover, it is not possible to specify the desired velocity to do the tracking. Another problem is that the system has an oscillatory behaviour when difficult trajectories are being tracked.

The method described here must be able to control the desired tracking velocity in order to solve the problems detected in these previous time-independent methods. To do so, the desired path $\mathbf{T} = \{\mathbf{s}^k / k \in 1 \dots N\}$ (with \mathbf{s}^k being the set of M points or visual features observed by the camera at instant k , $\mathbf{s}^k = \{\mathbf{f}_i^k / i \in 1 \dots M\}$) is sampled, and then these references are sent to the system as the desired references for each moment. In this way, the current and the final positions are very close together, and the system takes advantage of the good local behaviour of image-based visual servoing. From the initial set \mathbf{s}^1 of image features observed at the initial camera position, it is necessary to find an image configuration which provides the robot with the desired velocity $|\mathbf{v}_d|$ by iterating over the set \mathbf{T} . For each image configuration \mathbf{s}^k , the corresponding camera velocity

is determined considering an image-based visual servoing system (at this first stage $\mathbf{s} = {}^1\mathbf{s}$):

$${}^k\mathbf{v} = -\lambda \hat{\mathbf{L}}_s^+ (\mathbf{s} - {}^k\mathbf{s}). \quad (4)$$

This process continues until $|{}^k\mathbf{v}|$ is greater than the desired velocity, $|{}^k\mathbf{v}_d|$. At this moment, the set of features ${}^k\mathbf{s}$ will be the desired features to be used by an image-based visual servoing system (see Eq. 3). However, the visual features, ${}^j\mathbf{s}$, which provide the desired velocity, are between ${}^k\mathbf{s}$ and ${}^{k-1}\mathbf{s}$. To obtain the correct image features, the method described in [17] is employed.

Therefore, once the control law represented in Eq. 4 is executed, the system searches again for a new image configuration which provides the desired velocity. This process continues until the complete trajectory is tracked.

3.3 Visual force control

Now, we consider the task of tracking a path using visual and force information. The visual loop carries out the tracking of the desired trajectory in the image space. To do this, as it has been described in Section 3.2, the method to track trajectories in the image is employed:

$$\mathbf{v}_c = -\lambda \hat{\mathbf{L}}_s^+ (\mathbf{s} - {}^j\mathbf{s}) \quad (5)$$

where ${}^j\mathbf{s}$ is the set of features in the path obtained by the system to maintain the desired velocity.

Prior to defining the visual force controller employed, the meaning of the force–image interaction matrix, \mathbf{L}_{FI} , is described. To that end, considering \mathbf{F} as the interaction forces obtained with respect to the robot end-effector and \mathbf{r} as the end-effector location, the interaction matrix for the interaction forces, \mathbf{L}_F , is defined in the following way:

$$\mathbf{L}_F = \frac{\partial \mathbf{F}}{\partial \mathbf{r}} \rightarrow \mathbf{L}_F^+ = (\mathbf{L}_F^T \mathbf{L}_F)^{-1} \mathbf{L}_F^T = \frac{\partial \mathbf{r}}{\partial \mathbf{F}}. \quad (6)$$

From this last relationship and by applying Eq. 2 is obtained:

$$\begin{aligned} \dot{\mathbf{s}} &= \mathbf{L}_s \cdot \frac{\partial \mathbf{r}}{\partial \mathbf{t}} = \mathbf{L}_s \cdot \frac{\partial \mathbf{r}}{\partial \mathbf{F}} \cdot \frac{\partial \mathbf{F}}{\partial \mathbf{t}} = \mathbf{L}_s \cdot \mathbf{L}_F^+ \cdot \dot{\mathbf{F}} \rightarrow \dot{\mathbf{s}} \\ &= \mathbf{L}_{FI} \cdot \dot{\mathbf{F}} \end{aligned} \quad (7)$$

where $\mathbf{L}_{FI} = \mathbf{L}_s \cdot \mathbf{L}_F^+$ is the force–image interaction matrix. This matrix is estimated using exponentially weighted least-squares [18].

As it has been described in previous works [15], in order to guarantee the coherence between visual and force information, it is necessary to modify the image trajectory through the interaction forces. Therefore, in an application in which it is necessary to maintain a constant force with a given object, the image trajectory must be modified depending on the interaction forces. To do so, using the

matrix \mathbf{L}_{FI} , the new desired features used by the controller during the contact will be:

$$\mathbf{s}_d = {}^j\mathbf{s} + \mathbf{L}_{FI} \cdot (\mathbf{F} - \mathbf{F}_d). \quad (8)$$

Applying Eq. 8 to Eq. 3, the system is able to track a previously defined path in the image being compliant with the surface of the interaction object:

$$\mathbf{v}_c = -\lambda \hat{\mathbf{L}}_s^+ (\mathbf{s} - \mathbf{s}_d). \quad (9)$$

4 Heterogeneous cooperation

A heterogeneous cooperative architecture allows a team of robots and humans to improve the efficiency of their performance when they are working together with the same aim in the same workspace. The cooperative architecture chooses the team member which will perform each task better than any other team member.

For the assembly process, a hierarchical model of the product is generated using the relations among components, assemblies and sub-assemblies, as it is explained in [1]. Based on this model, the task planner generates a group of rules that guarantees that the task can be carried out successfully. A tree of actions, which shows the order and parallelism of the actions, is created from these rules. Afterwards, a generic decision tree, which includes the different decision trees for each action [9], is used by the task planner to compute the best robot to perform each action. In addition, the computation of the grasping considers the quantity of robots required to execute the task and the possibility of human collaboration based on weight distribution [19].

The task planner computes a free-collision path among robots and thus obtains the trajectory for each robot. The task planner considers the following aspects: the location of each robot in order to reduce the tool changes and the time for completing each action in order to minimise the global time required to perform the target task. But the task planner does not contemplate the capabilities and location of humans that could interact in the same industrial workspace. Consequently, the human–robot interaction may be dangerous for humans due to the possibility of collisions with robots or with heavy objects.

In this section is presented an approach based on different sensors to determine the global position of the human skeleton and location of all its limbs to avoid collisions between human and robots in the assembly process.

4.1 Sensors for human tracking

Motion capture systems [20] are able to track precisely the movements of the main limbs of a human being. Users of

these systems usually wear markers (or sensors) near each joint of their body. The motion capture system calculates their movements by comparing positions and angles between markers. Therefore, motion capture systems guarantee the safety of human operators in industrial environments.

An inertial motion capture system has been used in the work described in this paper. This system is composed of 18 small inertial sensors attached to a Lycra suit which is worn by the human operator. Each inertial sensor integrates miniaturised accelerometers and gyroscopes whose measurements are combined [21] to obtain the orientation (roll, pitch and yaw) of the operator's limb to which the sensor is attached. These inertial sensors neither suffer from occlusions like cameras nor from important magnetic interferences like magnetic trackers. The measurements obtained from the inertial sensors are applied to a skeleton (Fig. 3a) which represents the structure of the body of the operator to be tracked. The joint rotations of the skeleton are very precise because they are directly obtained from the inertial sensors measurements which have a resolution of 0.1° and an accuracy of 1° in yaw and 0.25° in roll and pitch. The global position of the skeleton in the environment (3D coordinates of the hips node of the skeleton) is estimated from the legs' rotation angles by a software algorithm which determines the length of the steps. However, this algorithm accumulates some error through time because foot strikes are not always correctly detected [22].

Due to this global translational error accumulated by the inertial motion capture system, an additional localisation system is needed. In particular, an UWB localisation system is used to correct this error. It is composed of two hardware

elements: four sensors which are installed at fixed positions in the industrial workspace (Fig. 3b) and a small tag (of similar size to a credit card) which is carried by the human operator. This tag sends UWB pulses to the sensors which estimate the global position of the operator by time difference of arrival and angle of arrival techniques [23].

Although the position measurements obtained from this UWB system are quite accurate (with mean errors smaller than 0.15 m), the sampling rate of the system is not high enough (only 5–9 Hz) to track quick human movements. Therefore, it is not possible to use only the measurements obtained from the UWB system as global position of the skeleton because there will appear jumps in the human trajectory when the operator is walking quickly. The combination of the UWB measurements with the translational measurements from the inertial motion capture system solves this problem (Fig. 4). On one hand, the UWB measurements are used to correct the error accumulated by the inertial measurements. On the other hand, between each pair of UWB measurements, the inertial measurements are used in order to maintain a high sampling rate (30 Hz) which can capture quick human movements. The fusion algorithm used to combine the translational measurements of both tracking systems was detailed in a previous work [22].

Finally, the rotational measurements registered by the motion capture system, the global position of the operator estimated by the fusion algorithm and the joint angles of the robots obtained from their controller are represented over a CAD model in order to show the configuration of the human–robot interaction at each instant of the assembly task (Fig. 3c).

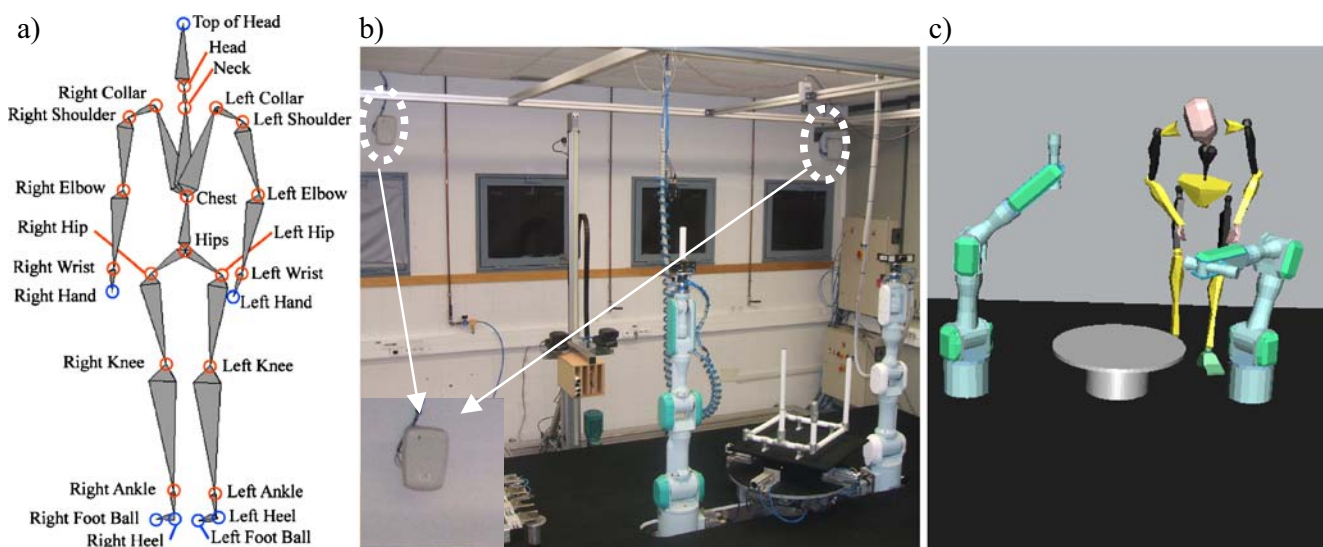


Fig. 3 a Skeleton structure where the measurements of the inertial motion capture system are applied. b UWB sensors of the indoor localisation system. c CAD model for 3D representation of the motion data from the human and the robots

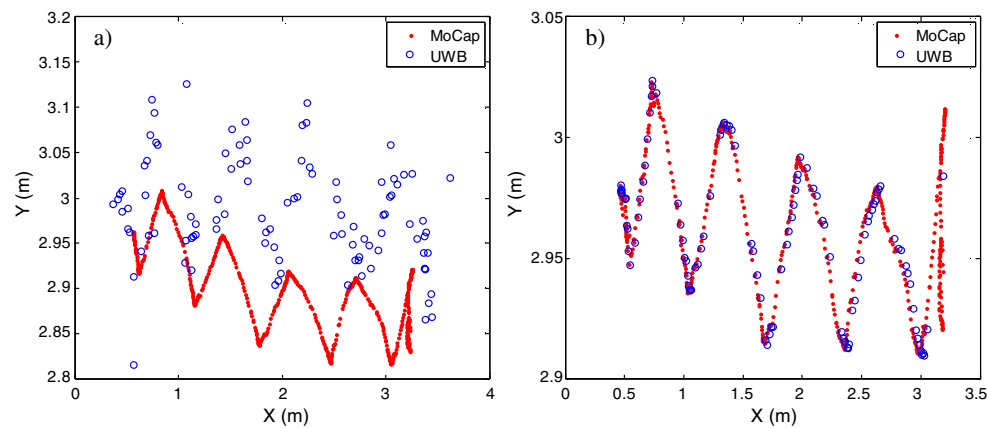


Fig. 4 **a** Original measurements of the human. **b** Position estimates obtained with the fusion algorithm

4.2 Robot behaviour for safe human–robot interaction

The global position of the human operator in the industrial workspace is obtained from the fusion algorithm described in [22]. The rest of the rotational measurements from the inertial sensors of the motion capture system are applied on the skeleton structure (Fig. 3a) in order to represent the movements of the main limbs of the operator. Thereby, the absolute position of all the nodes of the skeleton can be determined, and thus, the full body of the human operator is completely located in real time.

This full-body tracking makes possible the close interaction between human operators and robots. A safety behaviour has been implemented in the robot controller in order to guarantee the safety of the human operator. The visual force controller not only registers the movements of the robot but also the measurements from the human tracking systems (inertial motion capture system and UWB system). Every time the human tracking measurements are received, the controller calculates the Euclidean distance from every node of the skeleton of the human operator to the end-effector of the robot. The controller stores the smallest distance and compares it to a safety threshold (1 m) which represents the maximum distance between the hands of the operator and her/his body. Whilst this distance is greater than the safety threshold, the robot will continue tracking the planned path by visual servoing using the method described in Section 3.2. When the distance between one node and the robot is smaller than the safety threshold, the robot will stop tracking the initial path and will move away from the human operator in a linear path. The robot will try to maintain the distance between the end-effector and the closest node of the operator's skeleton. When the human operator goes away and the human–robot distance returns to be greater than the threshold, the visual servoing tracking is reactivated to follow the original path. Here, a time-independent behaviour is indispensable to

complete satisfactorily the tracking task, as it was described in Section 3.

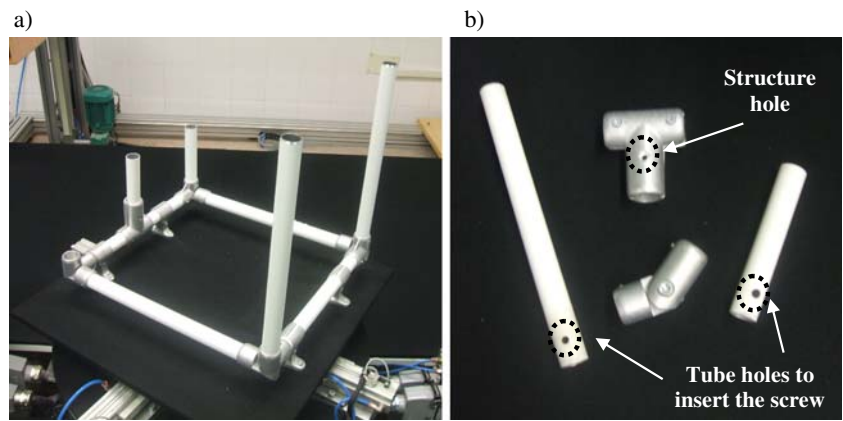
5 Description of the manufacturing process

In Fig. 5a, the metallic structure which will be assembled using the manufacturing system is shown. The task involves the assembly of the metallic structure composed by tubes and T-connectors (Fig. 5b). The insertion of the tubes will be performed automatically by the robotic manipulators. However, the addition of the T-connectors at the end of the tubes is a complex operation which cannot be performed by only two robots. In this case, the collaboration of a human operator is required.

The assembly sequence produced by the task planner is shown in Fig. 6. Furthermore, the real execution of the different tasks to carry out the assembly sequence is illustrated in Fig. 7. These tasks are the following:

- Task 1. The visual force control system is employed to insert an aluminium tube in its holder. Firstly, the robot grasps a tube, and subsequently, the visual servoing system is employed to guide the robot until the insertion begins (Fig. 7a). At this moment, a visual force control strategy is employed to assure that the insertion is performed correctly (Fig. 7b). The image trajectory is generated using four laser points projected on the table. These points will be the extracted features for the visual servoing system (see Fig. 7a).
- Task 2. In this task, the cooperation between both robots is required. The tube has a hole which must coincide with the hole in the structure (see the structure hole in Fig. 5b). To do so, robot R_1 rotates the tube whilst robot R_2 supports a range camera. The range camera determines the moment in which both holes coincide (Fig. 7c), and finally, once the tube has been

Fig. 5 a Metallic structure.
b Detail of the tubes and T-connectors which compose the structure



correctly orientated using task 2, a screw is inserted in order to fix the tube in the structure (Fig. 7d).

- Task 3. Robot–robot cooperation. In this task, a robot gives a tube to the other robot (Fig. 7e). Robot R_1 has grasped the tube from the store (Fig. 1) to give it to robot R_2 , and robot R_2 transports it to the suitable position.
- Task 4. As it has been previously indicated, the insertion of T-connectors at the end of the tubes requires the collaboration of a human operator. This task is performed by the human simultaneously to another insertion task developed by the robot as in task 1 (Fig. 7f).

described in Sections 2 and 3 are employed. The visual servoing system is used to guide the robot towards the point in which the tube insertion is done. When the insertion begins, the visual force control approach is employed in order to guarantee the correct insertion. This visual force control system not only makes the robot track a given desired trajectory but it also compensates the interaction forces obtained during the insertion. In this section, in order to illustrate the correct behaviour of the system, an example of insertion is described.

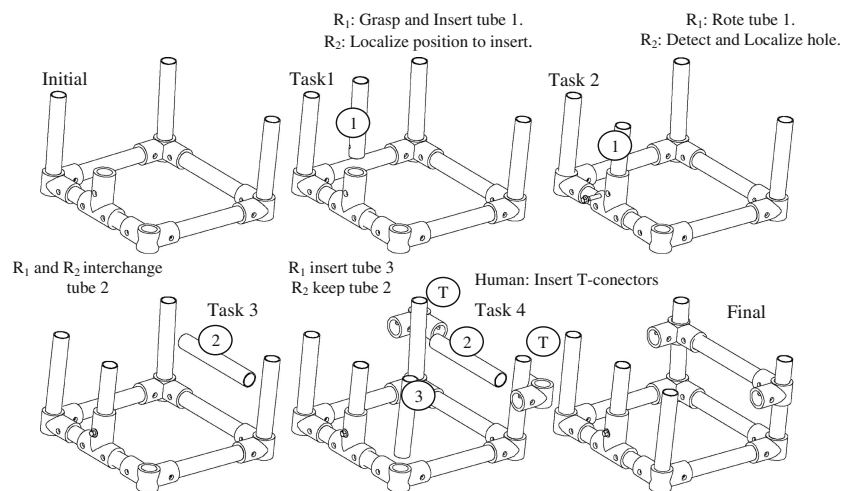
Figure 8 shows the 3D path required to perform one of the assemblies to construct the structure previously described. The desired path has been modified by considering the forces measured at the end-effector of the robot (see Fig. 9). In this way, the robot is able to introduce the tube into the aluminium holder correctly. Figure 10 shows the desired image path and the path modified by the visual force controller described in Section 3. The task can be accomplished thanks to the force–image interaction matrix [18] which permits modifying the desired image trajectory. The trajectory in the image space is recomputed online based on this matrix. The forces in Z-axis are regulated to

6 Results

6.1 Task 1: Tube insertion using visual force controller

In this phase, the task involves the insertion of an aluminium tube in its holder. This is a basic task in the construction of the structure. To do this, the approaches

Fig. 6 CAD of the assembly sequence to construct the metallic structure from task planner



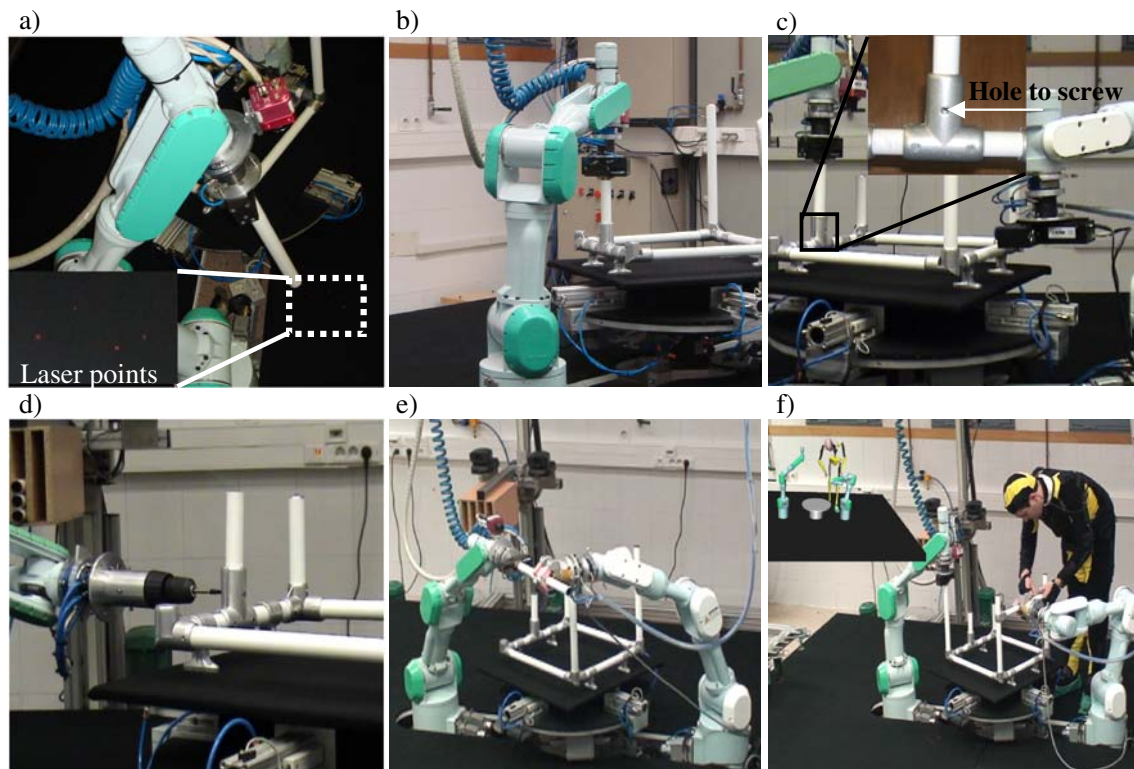


Fig. 7 Execution of tasks. **a** The visual servoing system guides the robot to begin the insertion. **b** A visual force control strategy is employed to carry out the insertion task correctly. **c** A range camera is used to determine the hole in which the screw must be inserted. **d**

Screwing in order to fix the tube into the holder. **e** Robot–robot cooperation. A robot gives a tube to the other robot. **f** Human–robot interaction

10 N during the insertion. This value has been empirically determined to assure the correct insertion. Once the force sensor has no measures different from the tool and the tube weights, the insertion is completed when the forces in Z-axis are lower than -10 N.

6.2 Task 2: Robot–robot coordination to fix the tube to the structure

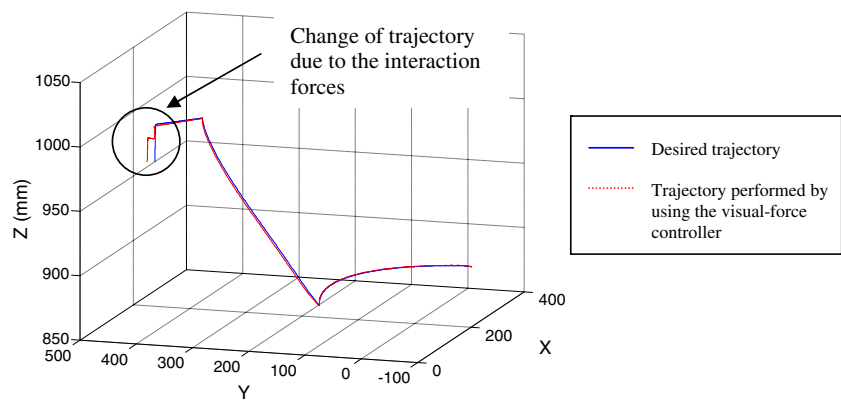
Once the tube has been inserted, a screw must be inserted to join the new tube with the structure. Before the insertion of the screw, it is necessary to make coincident the hole in the structure and the hole in the tube. As previously described in Section 5, robot R_1 rotates the tube until its hole coincides with the structure's hole.

In order to distribute tasks between the robots, the global planner is employed (see Section 4). To perform these tasks, the global planner generates two tasks: “detecting the hole” (T_1) and “inserting the screw” (T_2). Task T_1 is divided into two actions: “location of the holder's hole” (A_{11}) and “rotating the tube to find the hole” (A_{12}). Task T_2 has only one action: “inserting the screw” (A_2). Once the actions to be performed have been generated, the task planner has to distribute them among the robots. Consid-

ering the tools available, action A_{11} must be performed by robot R_2 . To perform action A_{12} , both robots are required, robot R_1 to rotate the tube using the gripper and robot R_2 to locate the hole with the range camera. Action A_2 must be performed by robot R_1 because it is the one that has the screwdriver.

Action A_{11} has to be performed previously to the tube insertion because in other cases, the hole will not be visible to locate it. The position of the hole of the tube holder is approximately known, using a CAD model of the workspace. With that information, the robot has to position the range camera in front of the hole. According to the geometric restrictions, the trajectory planner determines the movements of the robot to maximise the visibility of the hole [24]. Figure 11 shows the sequence of images captured by the range camera along the movement of the robot. In that sequence, the hole in the structure is located, maximising its visibility. Initially, the tube is not visible at all in the image. With the movement of the camera, the visibility of the structure is increased, improving the visibility of the hole that is the target of that action.

Once this action is done, robot R_1 has to insert the tube in the structure. After this, the tube must be orientated to achieve the correct visibility of the hole. Whilst robot R_2

Fig. 8 3D evolution of the end-effector in a tube insertion task

holds the camera, robot R_1 must rotate the tube. These are the actions assigned by the task planner to each robot. If the hole is not visible, the tube must be orientated looking for the correct orientation to have accessibility of the hole for inserting the screw. This last action is performed in a cooperative way; one robot is required to rotate the tube and the other one is used to control the range camera. Figure 12 shows the sequence of range images during the performance of that action. Initially, with the tube inserted in the structure, the hole previously located is not visible. Due to this, the tube has to be rotated until the hole is perfectly visible by the range camera. The sequence shows the images during the rotation of the tube until the hole is visible.

Once the tube is properly oriented, robot R_1 changes the gripper for a screwdriver to insert the screw in the hole. To carry out the insertion, the algorithms described in previous research are applied [25]. Screwing has been performed using the force sensor of robot R_1 to detect the moment when the screw is inserted in the structure to hold both parts of the structure. A control loop is used to control the speed of the robot end-effector according to the force measurement from the force sensor. Figure 13 shows the force applied when the screwing is performed. Initially, there is no force because the screw is not inserted in the hole. When the screw is inserted, the force increases, and that means that it has arrived to the end of the hole. The screwdriver is turned on and produces a wave form force whilst the screw is being screwed. When the force reaches a greater absolute value, the screw is fixed at the end of the hole and the screwdriver must be removed.

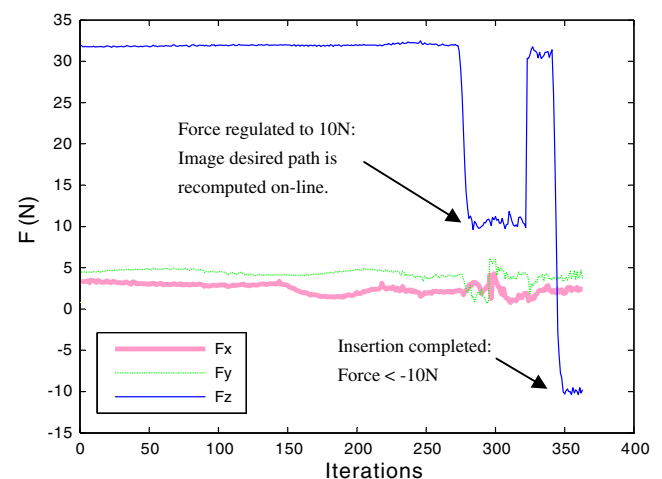
Once this task is performed, robot R_2 has to change the range camera for a robotic hand to perform the next task, and the other robot changes the screwdriver for a parallel gripper.

6.3 Task 3: Robot–robot cooperation

With the new tools, the next task to complete the structure requires that one robot holds a tube on the structure. To perform this task, first of all, the robot has to take the tube.

Due to the fact that the tube is out of the robot's work area, the other robot has to take the tube and give it to the robot. To perform this, the grasping point has to be computed. To compute the grasping point, a 3D model of the object is used. With that model, a hyperplane is selected to compute the skeleton of that plane. The grasping point will be in that skeleton, and it is selected considering the restrictions due to non-accessible parts of the object as well as the restrictions of the tasks and the tool used [26]. The grasping point for a group of robots is computed, taking into consideration the load balance of forces among them. Grasping points are selected as far as possible among the robot that holds the object and the nearest point to the geometric centre of the object. Due to this, the grasping point of an object can be recomputed if its restrictions change.

According to the grasping point algorithm, robot R_1 holds the tube, as is shown in Fig. 14a. In this figure, it can be seen that one part of the tube is not accessible because it is inside a storage box of tubes. In Fig. 14b, the new grasping point for robot R_2 is computed when the tube is

**Fig. 9** Forces measured at the robot end-effector during the tube insertion task

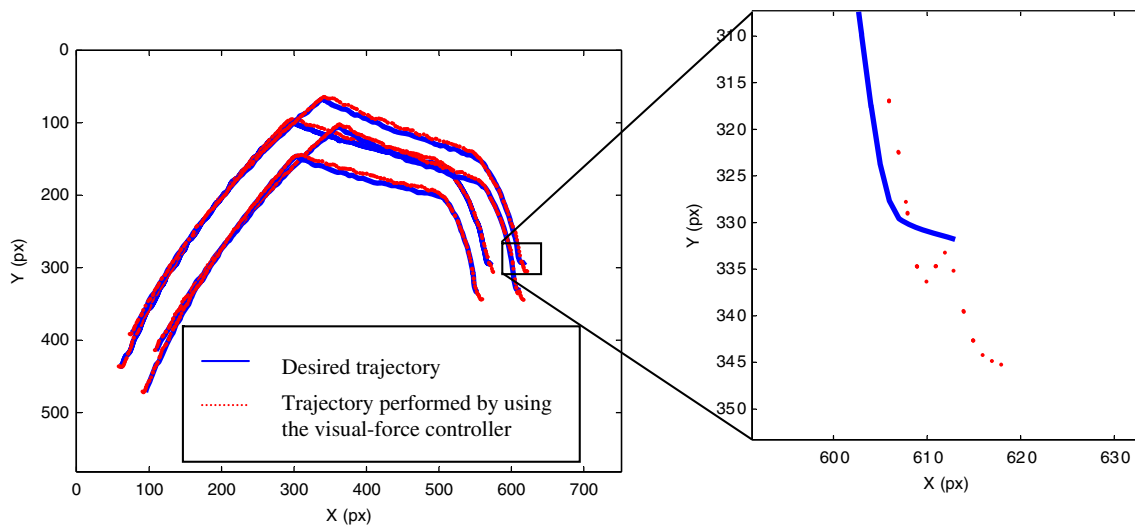


Fig. 10 Online modification of the features in the image in an insertion task by using the visual force controller

completely free outside of the storage box and thus there are no accessibility restrictions.

Once the specified robot has the tube, it is situated in its location in the structure and the human can perform his/her part of the task consisting in attaching the tube to the structure. At the same time, the other robot has to insert another tube in the structure. During the execution of these tasks, the system has a feedback to control that both tasks are performed whilst avoiding collisions and guaranteeing the safety of the human.

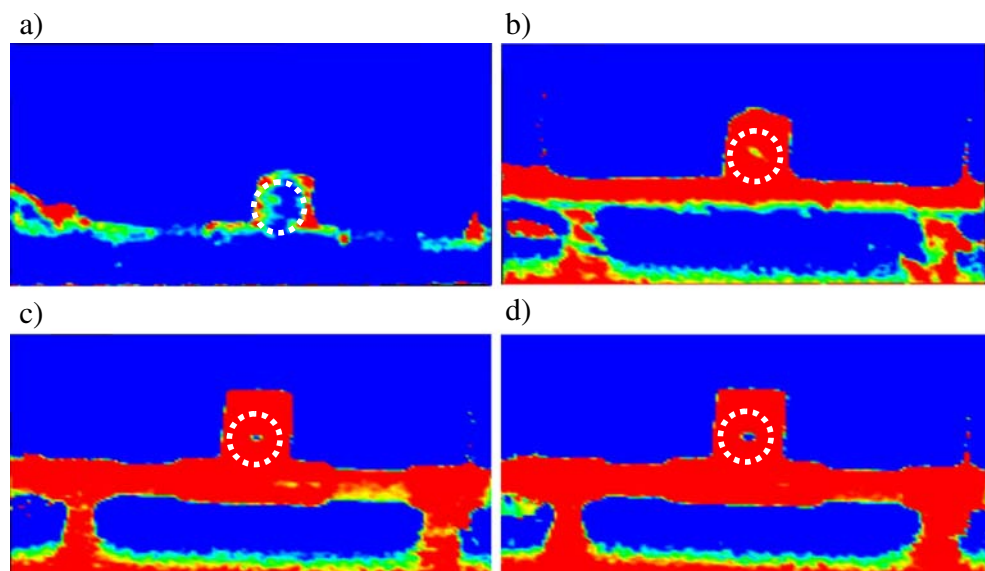
6.4 Task 4: Human–robot cooperation

The whole sequence of this interaction task is depicted in Fig. 15. Some frames of the sequence not only include a photograph of the task but also the 3D CAD representation

of motion data captured by the robot controller (the human skeleton obtained from the tracking systems and the joint angles of the robots). The interaction task can be summarised in the following steps:

1. Frame 1: One of the robots, R_2 , is holding the tube in the place where the human operator is going to add the T-connectors, whilst the other robot, R_1 , carries a tube in order to insert it in the structure.
2. Frames 2 and 3: When the distance between the human operator and the end-effector of the robot is smaller than the safety threshold, the robot controller stops the task of robot R_1 (which is inserting a new tube) and executes the safety behaviour described in Section 4.2. This safety behaviour makes the robot move away from the human operator in a linear path in order to maintain

Fig. 11 Range images during the location of the holder's hole. **a** Robot R_2 is near the structure. **b** The robot is in front of the structure. The hole is visible but the robot is not well orientated. **c** The hole is visible. **d** Final position. The hole is visible and the robot is correctly orientated



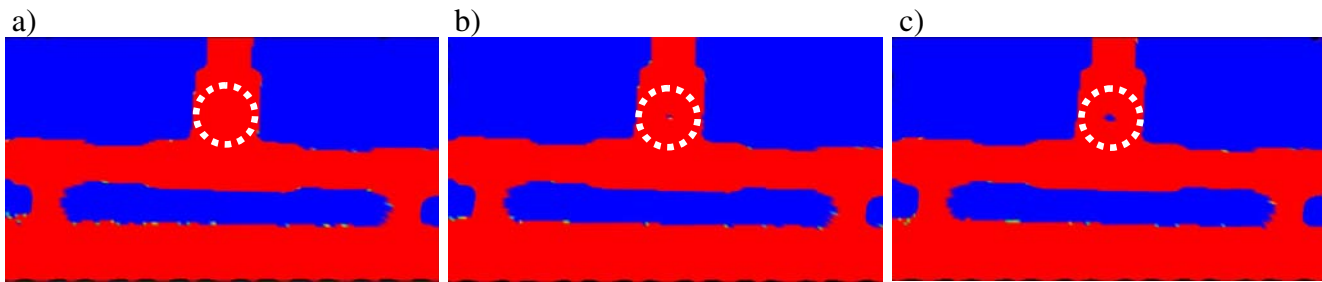


Fig. 12 Range images whilst rotating the tube to find the hole. **a** Both holes do not coincide. The robot R_1 begins the rotation of the tube. **b** Both holes begin to be coincident. The robot continues the rotation. **c** Both holes are perfectly coincident

the safety distance (Fig. 16). Whilst robot R_1 is performing the safety behaviour, the human operator puts the two parts of the T-connector at the end of the tube and screws them. This safety behaviour is not executed in robot R_2 because it does not move whilst holding the tube and there is no risk of collision.

3. Frame 4: When the human operator finishes assembling the first T-connector and goes away from the metallic structure, robot R_1 abandons the safety behaviour and continues with the stopped task for inserting the tube in the structure.
4. Frame 5: Whilst the human operator assembles the second T-connector at the other end of the tube, the safety behaviour is not executed because the safety distance threshold is not exceeded. Therefore, robot R_1 continues inserting the tube whilst the human operator screws the T-connector.
5. Frame 6: When the human operator ends the screwing of the T-connector and the robots finish their tasks, the robots return to their home position.

In order to show the time-independent behaviour of the approach described in Section 3.2, a new experiment is described next. In this experiment, if there is no human

detection and the time-independent system to track trajectories in the image acts normally, the tracking is 12 s long. However, in order to expose the difference between the system behaviour and a time-dependent system, Fig. 17 shows an example of a tracking in which the trajectory of the robot is near the human and the tracking is stopped for approximately 3 s. During the presence of the human, the robot goes off the 3D desired path to maintain the safety distance (1 m). Figure 17 shows that once the human detection ends, the system continues with the path tracking from the next reference point in the desired path.

Figure 18 shows the evolution of one of the features in the image obtained with the time-independent method described and with a time-dependent method. The tracking is correctly performed until the moment when the safety behaviour begins. Nevertheless, from the moment when the safety behaviour is disabled, the time-dependent system is not able to return to the exact point in the trajectory in which the obstruction began. This is due to the loss of the temporal references. Therefore, the time-independent method described in Section 3.2 is adequate for the tracking of image paths in tasks in which the robot interacts with a human or with any object of its workspace.

Fig. 13 Screwing force

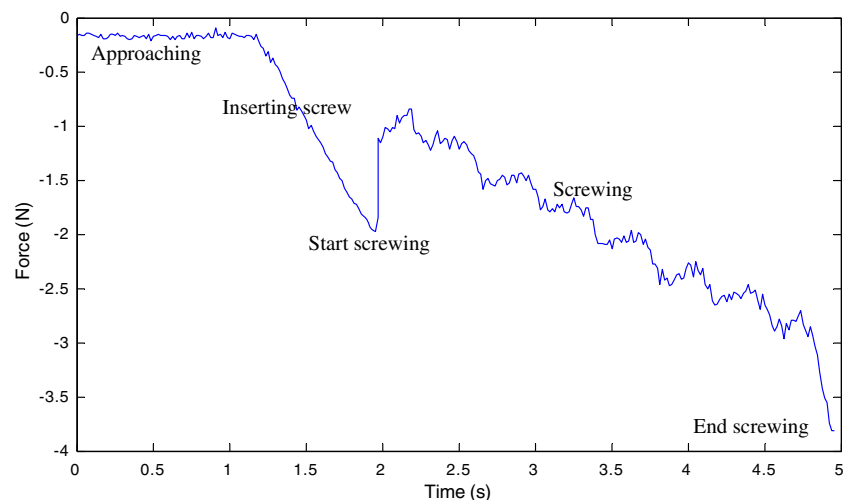
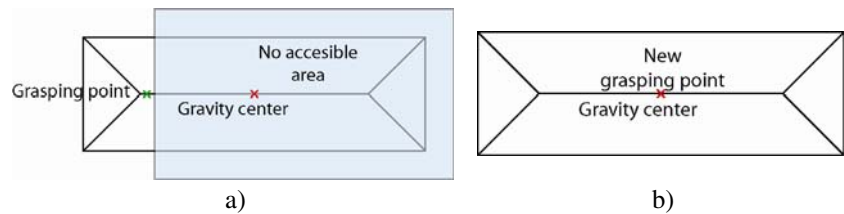


Fig. 14 **a** Grasping point of the tube inside the storage box. **b** Grasping point of the tube for robot R_2



7 Conclusions

The proposed system has the novelty of a high degree of flexibility with an intelligent multisensorial system. This sensorial system is composed of a visual force control subsystem, a time of flight 3D camera, an inertial motion capture subsystem and an indoor localisation subsystem. These sensorial systems are working in this application

cooperatively in order to provide a high degree of flexibility.

In this paper, a robotic system to assemble a metallic structure has been presented. An important aspect of the proposed application is the flexibility provided by the multisensorial system employed. To perform the construction of the metallic structure, two robots are necessary to work cooperatively. Furthermore, the human operator must

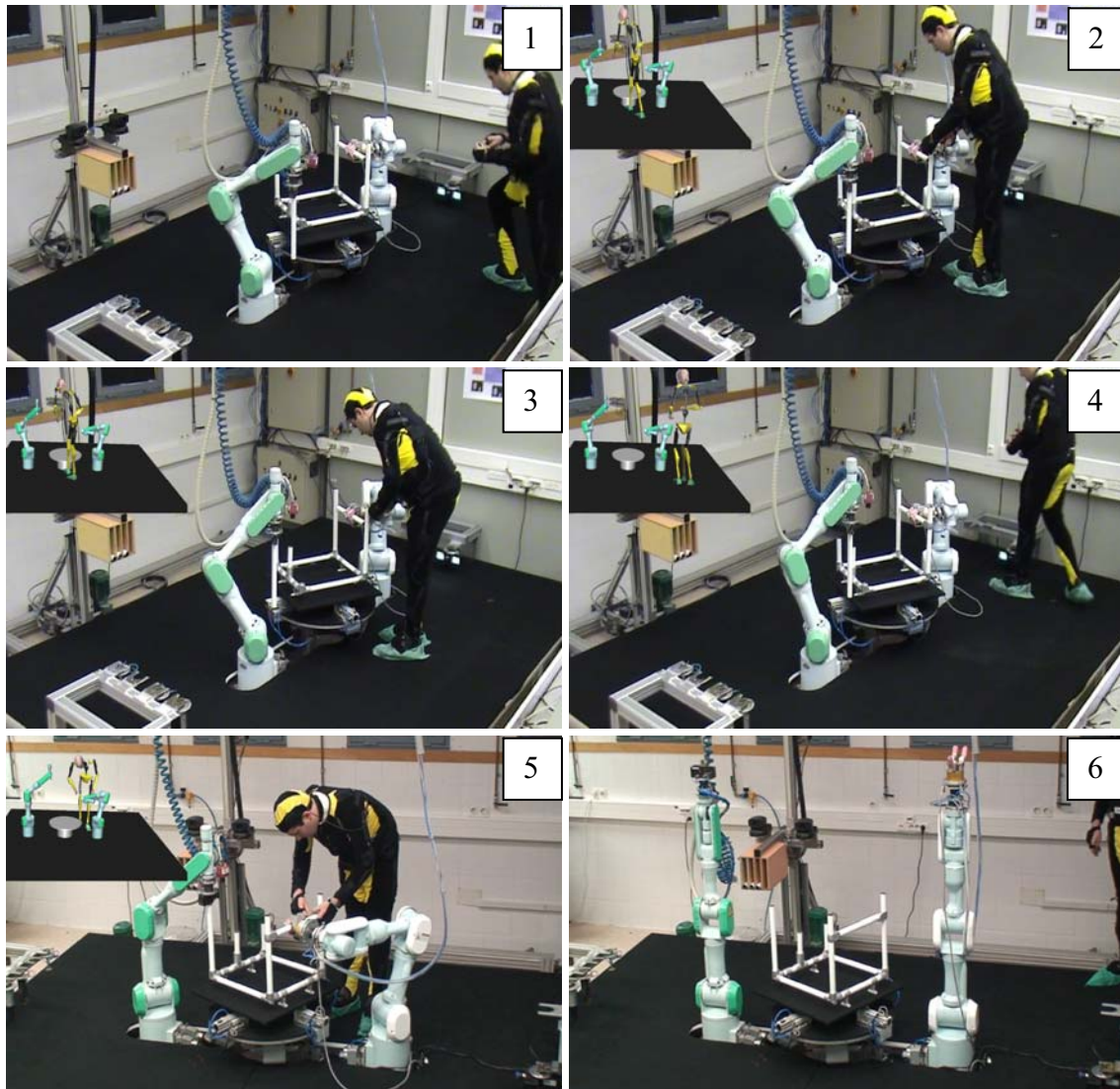


Fig. 15 Sequence of frames of the collaboration task between two robots and a human operator to assemble a metallic structure. A 3D representation of the motion data from the human operator and the robots is depicted in frames 2–5 at the *top left corner*

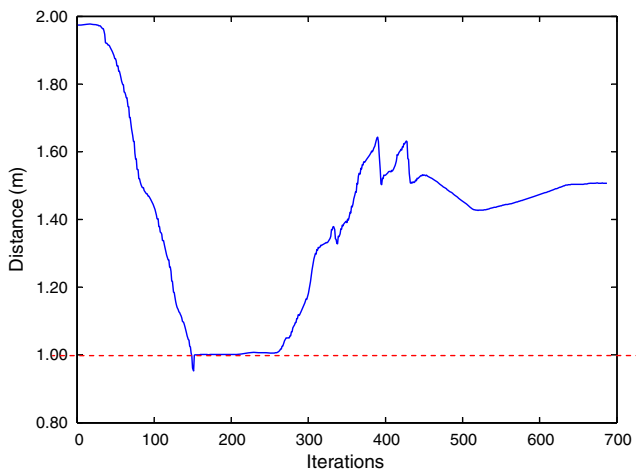


Fig. 16 Evolution of the distance between the closest node of the human skeleton and the end-effector of the robot. The safety threshold (1 m) is marked with a *dashed line*

collaborate with the robots during the task. In order to avoid possible collisions between the human operator and the robots, a new human localisation system is proposed. This localisation system is composed of an inertial motion capture system and a UWB localisation system. Although most previous motion capture systems can register precisely human movements, not all motion capture technologies are suitable for human–robot interaction tasks in industrial applications. Vision-based tracking systems determine the position of the human operator by processing the images of the environment captured by one or several cameras. In many cases, a complex installation of multiple cameras is needed in order to avoid occlusions. However, these types of installations are not possible in cluttered industrial environments. Magnetic-based trackers are also used for human tracking in human–robot interaction applications. Nevertheless, these systems are not suitable for industrial

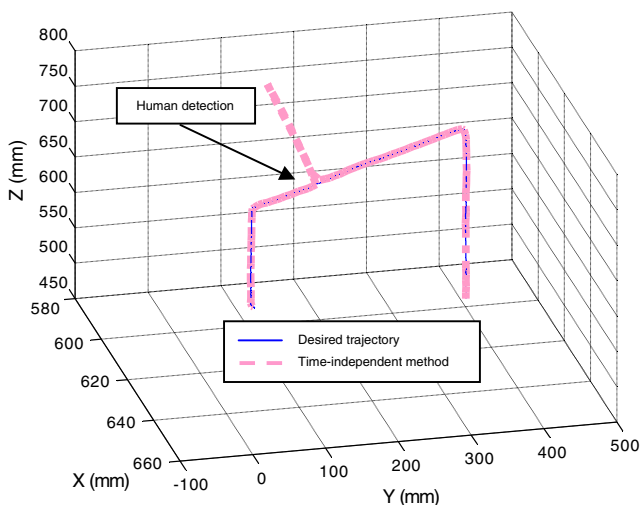


Fig. 17 3D trajectory of the end-effector of the robot obtained with the proposed time-independent method

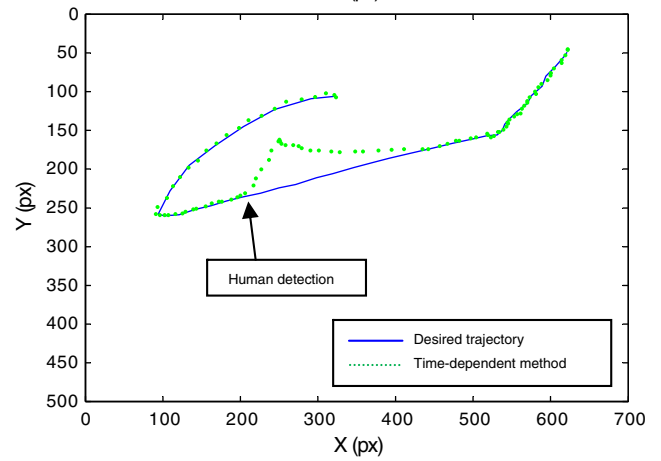
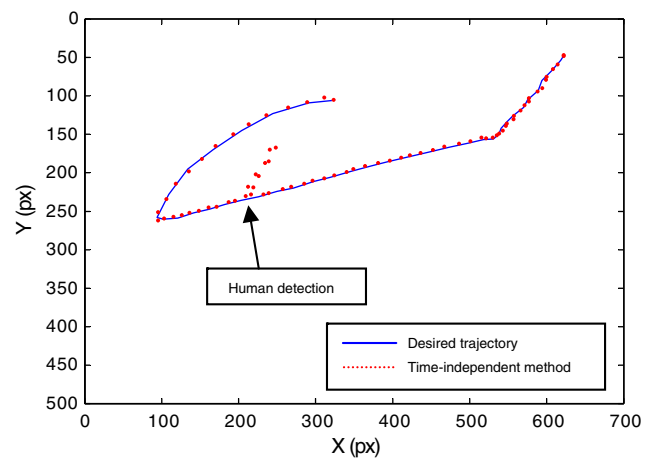


Fig. 18 Evolution of one of the features in the image space using the time-independent tracking (*top*) and using a time-dependent method (*bottom*)

workplaces because heavy industrial machineries generate intense dynamic magnetic fields which distort the magnetic field of the motion capture source. Due to these drawbacks, the proposed localisation system is used to precisely locate the human in industrial environments.

Different improvements have been developed in the proposed sensor control systems with respect to the previous ones. A new time-independent visual force control system has been proposed in order to guarantee the correct robot guidance in situations in which the robot interacts with the workspace. Our time-independent visual force control is robust against interruptions in its trajectory. In addition, another contribution is a visual–force interaction matrix which guarantees the coherence between the two types of information: visual and force data.

In this work, the use of other sensors like flight 3D camera have allowed inspecting better pieces and parts of them when the light conditions and visibility are difficult for sensorial systems based on CCD or CMOS cameras. In particular, a range camera has been used to detect the holes of small dimensions and to determine the location approached of the pieces in the workspace.

Acknowledgements This work was funded by the Spanish MEC project “Design, Implementation and Experimentation of Intelligent Manipulation Scenarios for Automatic Assembly and Disassembly Applications (DPI2005-06222)” and the pre-doctoral grant AP2005-1458.

References

- Torres F, Puente S, Aracil R (2003) Disassembly planning based on precedence relations among assemblies. *Int J Adv Manuf Technol* 21(5):317–327, doi:10.1007/s001700300037
- Bai YW, Chen ZN, Bin HZ, Hun J (2005) An effective integration approach towards assembly sequence planning and evaluation. *Int J Adv Manuf Technol* 27(1–2):96–105, doi:10.1007/s00170-004-2155-y
- Pauli J, Schmidt A, Sommer G (2001) Vision-based integrated system for object inspection and handling. *Robot Auton Syst* 37:297–309, doi:10.1016/S0921-8890(01)00161-0
- Dongming Z, Songtao L (2005) A 3D image processing method for manufacturing process automation. *Comput Ind* 56:975–985, doi:10.1016/j.compind.2005.05.021
- Kosmopoulos D, Varvarigou T (2001) Automated inspection of gaps on the automobile production line through stereo vision and specular reflection. *Comput Ind* 46:49–63, doi:10.1016/S0166-3615(01)00113-0
- Gil P, Pomares J, Puente ST, Diaz C, Candelas F, Torres F (2007) Flexible multi-sensorial system for automatic disassembly using cooperative robots. *Int J Comput Integr Manuf* 20(8):757–772, doi:10.1080/09511920601143169
- Onishi M, Odashima T, Luo Z, Hosoe S (2006) An immersion type 3D dynamic simulation environment for developing human interactive robot systems. *Syst Comput Jpn* 37(10):368–377
- Son C (2002) Optimal control planning strategies with fuzzy entropy and sensor fusion for robotic part assembly tasks. *Int J Mach Tools Manuf* 42:1335–1344, doi:10.1016/S0890-6955(02)00063-9
- Torres F, Puente S, Diaz C (2008) Automatic cooperative disassembly robotic system: task planner to distribute tasks among robots. *Control Eng Pract* 17(1):112–121, doi:10.1016/j.conengprac.2008.05.013
- Hutchinson S, Hager GD, Corke PI (1996) A tutorial on visual servo control. *IEEE Trans Robot Autom* 12(5):651–670, doi:10.1109/70.538972
- Chaumette F (1998) Potential problems of stability and convergence in image-based and position-based visual servoing. In: Kriegman DJ, Hager GD, Morse AS (eds) *The confluence of vision and control*. Springer, New York, pp 66–78
- Kim DH, Ryu YG, Chung MJ (2007) Stability analysis of image-based visual servoing regarding with camera calibration errors. *Proceedings of International Conference on Advanced Robotics (ICAR 2007)*, pp 176–181
- Malis E, Rives P (2003) Robustness of image-based visual servoing with respect to depth distribution errors. *Proceedings of IEEE International Conference on Robotics and Automation (ICRA 2003)*, pp 1056–1061
- Mezouar Y, Chaumette F (2002) Path planning for robust image-based control. *IEEE Trans Robot Autom* 18(4):534–549, doi:10.1109/TRA.2002.802218
- Pomares J, Torres F (2005) Movement-flow based visual servoing and force control fusion for manipulation tasks in unstructured environments. *IEEE Trans Syst Man Cybern. Part C Appl Rev* 35(1):4–15
- Schramm F, Morel G (2006) Ensuring visibility in calibration-free path planning for image-based visual servoing. *IEEE Trans Robot* 22(4):848–854, doi:10.1109/TRO.2006.878955
- Garcia GJ, Pomares J, Torres F (2007) A new time-independent image path tracker to guide robots using visual servoing. *Proceedings of IEEE International Conference on Emerging Technologies and Factory Automation*, pp 957–964
- Garcia GJ, Pomares J, Torres F (2007) Robot guidance by estimating the force-image interaction matrix. *Proceedings of IFAC International Workshop on Intelligent Manufacturing Systems (IMS 2007)*
- Puente ST, Torres F, Díaz C (2008) Assembly/disassembly strategies for service applications. *Proceedings of 17th IFAC World Congress*, pp 15805–15810
- Welch G, Foxlin E (2002) Motion tracking: no silver bullet, but a respectable arsenal. *IEEE Comput Graph Appl* 22(6):24–38, doi:10.1109/MCG.2002.1046626
- Foxlin E, Harrington M, Altshuler Y (1998) Miniature 6-DOF inertial system for tracking HMDs. *Proceedings of SPIE Helmet and Head-Mounted Displays III*
- Corrales JA, Candelas FA, Torres F (2008) Hybrid tracking of human operators using IMU/UWB data fusion by a Kalman filter. *Proceedings of ACM/IEEE International Conference on Human-Robot Interaction*, pp 193–200
- Ubisense (2008) System overview, white paper. Available at <http://www.ubisense.net>
- Gil P, Torres F, Reinoso O (2007) Estimation of camera 3D-position to minimize occlusions. *Proceedings of International Conference on Informatics in Control, Automation and Robotics (ICINCO 2007)*, pp 311–317
- Puente ST, Torres F (2004) Automatic screws removal in a disassembly process. *Proceedings of CLAWAR/EURON Workshop on Robots in Entertainment, Leisure and Hobby*
- Diaz C, Puente S, Torres F (2007) Grasping point for handle objects in a cooperative disassembly system. *Proceedings of IFAC International Workshop On Intelligent Assembly and Disassembly (IAD 2007)*