# A Tele-operation System for Collaborative Works with Vision-Guided Autonomous Robot

Shiqi Li, Mingfu Li, Di Zhao, and Wenge Zhu

School of Mechanical Science and Engineering Huazhong Universty of Science and Technology, Wuhan, Hubei Province, China
sqli@mail.hust.edu.cn, Lmfwolmf2000@gmail.com, vbscripzt@126.com, wgzhu2000@mail.hust.edu.cn

**Abstract.** Fully autonomous robot is limited by level of nowadays artificial intelligence. The research of interface between human and autonomous robot is a burgeoning area in robotics research. User Interface for dynamic autonomy is proposed in this paper for the sake of switch between different control modes to cope with dynamic verification in remote environment. Vision guided fully autonomous act as the basic function for the system. Shared autonomy based on natural image user interface and remote program act as default control mode to combine merits both of human and autonomous robot. The experiments show that the dynamic autonomy tele-operation system proposed in this paper is friendly for human operator and efficient.

**Keywords:** Dynamic autonomy; Human-robot interaction; User interface; Tele-operation.

## 1 Introduction

Remote environment often verify dynamically in tele-operation, timedelay is existed in the process as well. Traditional bilateral tele-operation can't handle with timedelay and dynamic verification at the same time. To cope with dynamic environment, telerobots required to be endowed with capability of local autonomy. However, fully autonomous robots are not yet competent for complex tasks operation in remote environment independently because nowadays autonomous systems are unapt at task definition and object recognition. Human operator is with stronger capability of task definition and object recognition than robot, but robot outshines human being in precision, and nowadays computer is not worse, even better than human being in task planning and path planning [1]. Therefore, the problem is expected to be solved by combining merits both of human being and autonomous robot to construct a collaborative tele-operation system.

However, the research of interface between human and autonomous (or semi-autonomous) robot is arisen only recently few years [2]. Researchers in CMU have the pioneer researches, Fong et al. [3] [4] present collaborative control which based on the interface of operator-robot ask, answer question and dialog, in the system humans and robots work as partners. Furthermore, human and autonomous robot exchange information and assist one another to achieve common goals. The researches

C. Xiong et al. (Eds.): ICIRA 2008, Part II, LNAI 5315, pp. 111–120, 2008.

revealed that interface of dialogue are valuable for tele-operation and significantly helped novices understand problems encountered by the robot.

Birk et al. [5] presented an adaptive human machine interface for rescue robots which supports adjustable autonomy by automatically changing its display and control functions based on relevant measures. Bruemmer et al. [6] developed a novel human-robot interaction concepts and interfaces for robust, mixed-initiative interaction between robots and humans. Those interfaces utilize simultaneous localization and mapping techniques that capture an abstracted representation of the robot's experience and exploit sensor-suites and fusion algorithms that enhance capabilities for sensing, interpreting, and "understanding" environmental features. These studies of interaction between humans and autonomous robots present a fascinating and crucial new area of research.

In this paper, user interface based on vision guided autonomous robot is developed to construct a dynamic autonomy tele-operation system. This paper is organized as follows. The framework and overview of the system is described in section 2. Fully autonomy mode of vision guided robotic tracking, reaching and grasping is presented in section 3. Section 4 discussed the natural image user interface and remote programming based shared autonomy tele-operation. Dynamic autonomy of switching between four control modes is discussed in section 5. Experiments of the system are presented in section 6. Conclusions and possible future directions are shown in section 7.

## 2 System Overview

Accessible and fluent interaction is vital for the sake of combining of human and autonomous robot to the same system. As dynamic verification in remote environment, relevant different control mode is proposed in this paper to suit different situation. Control mode of direct tele-operation, augmented reality and predictive display, shared autonomy and fully autonomy are developed in this paper. A controller clutch is presented to switch between these four control modes, feedback information filtrated according to selected control mode, and display by relevant channel for the operator. Remote robot is endowed with autonomous capability by vision guided reaching and grasping. The framework of multi-control-mode dynamic tele-operation as showed in Fig.1. The details of different modes will describe in following sections.

## 3 Fully Autonomous: Vision Guided Autonomous Robotic System

Fully autonomous is presented as a basic capability of telerobot in this paper. Telerobot can handle with certain tasks by close loop control, visual information which captured by vision sensors guide the robotic actuator execute expected tasks. The close loop control contains three aspects: object searching, visual tracking and vision guided reaching.

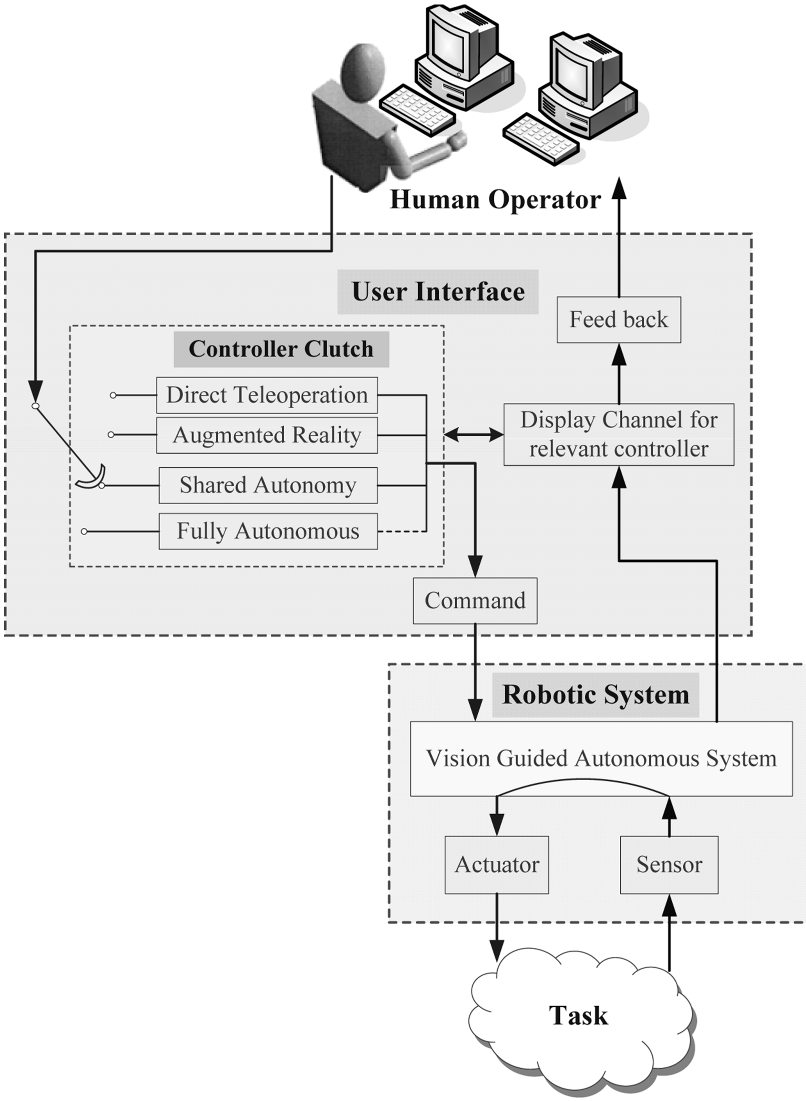



**Fig. 1.** System architecture of dynamic autonomy tele-operation system

**Object Searching & Visual Tracking.** A shape template can be constructed before hand by defining geometric parameter according to the method of template tracking [7]. And so, the position angle '$\theta$', distance from center to left-down point of image plane '$\gamma$', width of object '$\omega$', length of object '$\iota$' and slope angle '$\varphi$'. And so the position and status of object can be described by state vector '$S$':

$$S=[\theta,\gamma,\omega,\iota,\varphi] \tag{1}$$

Contour is divided by $J$ control points into $J$+1 segmented lines, visual tracking can be implemented by estimation of the observation likehood:

$$h(I_k|S_k) \equiv h(Z_k|S_k) = \prod_{i=1}^{I} h(z_{i,k}|s_{i,k}) \tag{2}$$

In Eq. (2), $h(z_{i,k} \mid s_{i,k}) \equiv \prod_{j=1}^{J} p(z_{j,i,k})$. CONDENSATION [8] tracking can be implemented by prediction, observation, re-sampling and estimation step. Therefore, object can be perceived and tracking in real time.

**Vision Guided Reaching: Visual Servoing.** For the sake of providing more visual information for visual feedback control and decreasing the adverse effects brought by contour overlay in monocular camera, stereo vision system constructed by two cameras is applied. Visual servoing control is implemented based on binocular disparity. According to dynamic contours of object and robotic hand (as previous described), the geometric feature (such as area center and center line) can be positioned. And the binocular disparity between two cameras and object can be got dynamically. Accurate visual alignment is achieved by controlling binocular disparity to zero (more details see [9]).

# 4 Shared Autonomy

Remote fully autonomous system can only handle with tasks which very specific such as object with special shape, operation with special flow and so on. Shared autonomy is proposed in this paper to combine merits of autonomous robot and human operator.

In order to improve the efficiency of communication between human and vision guided autonomous robot, natural image user interface is proposed in this paper based on active contour, by which operator can select the objects to be operated in natural video stream, then program the task based on shared autonomy remote program. The autonomous telerobot execute the tasks automatically which defined by remote program. In the process of tele-operation, human operator can intervene to cancel the current task, send command of exigency stop and override the commands at any moment.

## 4.1 Objects Pick-Up: Natural Image User Interface

On the one hand, natural image of video stream is directviewing for human operator to understand the situation in the remote environment. On the other hand, vision guide autonomous robot is constructed based on natural image feedback. Therefore, natural image can act as intermedium between human operator and telerobot for interacting and exchanging information each other.

One of the most important purposes of interaction between human operator and telerobot is to provide the objects to be operated. The objects are viewed as dynamic contour for telerobot, human operator pick-up contours in the natural image of feedback video stream by B-spline method. The position information of control points of

B-spline curve is sent to remote telerobot serve. Telerobot approximating and tracking the real contour by B-Snake [10].

According to B-spline, the curve $\boldsymbol{v(s)}$ ( $\boldsymbol{r(s)}=(x(s),y(s))$ ) in the image plane can be defined by control points as follow:

$$\boldsymbol{v(s)}=\sum_{n=0}^{N-1} B_n(s)p_n \qquad 0\le s\le L \tag{3}$$

$B_n(s)$ are basis functions, $p_n$ ( $n=0,1,\cdots N-1$ ) are control points which picked by mouse or keyboard input, the initial curve $\boldsymbol{v}_0\boldsymbol{(s)}$ is got by applying B-Snake to approximating the real contour of the object. In order to track the contour at real time, the B-spline curve required plane affine transformation, affine space of $\boldsymbol{v}_0\boldsymbol{(s)}$ can be viewed as a set of linear transformation:

$$\boldsymbol{v(s)} = \boldsymbol{A}\boldsymbol{v_0(s)} + \boldsymbol{C} \tag{4}$$

In which $\boldsymbol{A}=\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$, $\boldsymbol{C}=\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$, and so:

$$\begin{aligned} \boldsymbol{v(s)} - \boldsymbol{v_0(s)} &= \boldsymbol{C} + (\boldsymbol{A} - \boldsymbol{I})\boldsymbol{v_0(s)} \\ &= \begin{pmatrix} c_1B^T(s) \\ c_2B^T(s) \end{pmatrix} + (\boldsymbol{A} - \boldsymbol{I})\begin{pmatrix} B^T(s) & 0 \\ 0 & B^T(s) \end{pmatrix}\begin{pmatrix} \boldsymbol{P_0}^{\boldsymbol{x}} \\ \boldsymbol{P_0}^{y} \end{pmatrix} \end{aligned} \tag{5}$$

Let $\boldsymbol{\beta(s)}=\begin{pmatrix} B^T(s) & 0 \\ 0 & B^T(s) \end{pmatrix}$, then:

$$\begin{aligned} \boldsymbol{v(s)} - \boldsymbol{v_0(s)} = c_1\boldsymbol{\beta(s)}\begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_2\boldsymbol{\beta(s)}\begin{pmatrix} 0 \\ 1 \end{pmatrix} + (A_{11} - 1)\boldsymbol{\beta(s)}\begin{pmatrix} \boldsymbol{P_0}^{x} \\ 0 \end{pmatrix} + \\ A_{21}\boldsymbol{\beta(s)}\begin{pmatrix} 0 \\ \boldsymbol{P_0}^{x} \end{pmatrix} + A_{12}\boldsymbol{\beta(s)}\begin{pmatrix} \boldsymbol{P_0}^{y} \\ 0 \end{pmatrix} + (A_{22} - 1)\boldsymbol{\beta(s)}\begin{pmatrix} 0 \\ \boldsymbol{P_0}^{y} \end{pmatrix} \end{aligned} \tag{6}$$

Affine space of $\boldsymbol{v}_0\boldsymbol{(s)}$ can also defined by control points:

$$\boldsymbol{P}=\boldsymbol{W}\boldsymbol{X}+\boldsymbol{P_0} \tag{7}$$

In which $\boldsymbol{W}=\begin{pmatrix} 1 & 0 & \boldsymbol{P_0}^{x} & 0 & \boldsymbol{P_0}^{y} & 0 \\ 0 & 1 & 0 & \boldsymbol{P_0}^{x} & 0 & \boldsymbol{P_0}^{y} \end{pmatrix}$ represent 6 DOF shape matrix, $\boldsymbol{P_0} = \begin{pmatrix} \boldsymbol{P_0}^{\boldsymbol{x}} \\ \boldsymbol{P_0}^{y} \end{pmatrix} = \begin{pmatrix} p_0{}^x, p_1{}^x, \cdots p_{N-1}{}^x \\ p_0{}^y, p_1{}^y, \cdots p_{N-1}{}^y \end{pmatrix}$ is initial curve, according to Eq.(6) and Eq.(7), the shape control vector $X$ is:

$$\boldsymbol{X}=(c_1,c_2,A_{11}-1,A_{22}-1,A_{21},A_{12})^T \tag{8}$$

The real contour can be tracked at real time by control the parameter $c_1, c_2, A_{11} - 1, A_{22} - 1, A_{21}$ and $A_{12}$.

### 4.2 Remote Programming for Shared Autonomy

The objects contours can be picked in natural user interface based on B-spline curve, the objects have their index such as: object ‘1’, object ‘2’, …object ‘N’, user can name each object manually.

As the objects are defined by natural image interface, the operation task needs to determine by operator. As for tele-operation, the operation tasks include four parts: end effector or tool selection, operational action, status checking and speed setting. And the four parts constructs a tele-autonomy command to be sent to telerobot. Telerobot capture and track relevant objects, and execute the tasks according to program explain module.

Operator can programming in the user interface freely, for the users who unfamiliar with the system, the interface provide step by step textbox filling and generate tele-autonomy command finally. A shared autonomy command can be programmed as follow:

```
{Start:
 object enumeration (object ‘1’,object ‘2’,…object ‘N’)
////***For object ‘n’ (1•n•N) ***////////
object ‘n’:{
tool: (?dextrous hand  ?gripper  ?driller)
action:(?reaching  ?grasping  ?drilling  ?following
?releasing)
velocity: (?high  ? moderate  ?slow  ?following)
is action finished ? next object operation : repeat
current object operation
}
Wait: XXX(ms)
////***For object ‘i’ (1•i•N) ***////////
object ‘i’:{
 //include four parts: tool, action, speed and status
checking
 }
Is task finished? End: next object operation
}
```

# 5 Interface for Dynamic Autonomy

Dynamic autonomy is proposed for the sake of cope with dynamic verification in remote environment. Therefore, the interface must provide channel for smooth switching and convenient human intervention. Interface proposed in this paper contain a control mode clutch, fully autonomous mode, shared autonomy mode, augmented reality and predictive display mode and direct bilateral tele-operation.

**Clutch for Control Mode Switching.** A clutch is designed for control mode selection and switching. Operator can make switch from one control mode to another, furthermore, display channel also switch to suit relevant control mode. As show in Fig.1, the clutch also provide filter for information of feedback data and send out command to reduce human and robot error.

**Fully Autonomy.** Vision guided fully autonomy is basic capability of telerobot, however, as described in section 3, application of fully autonomous limited to special tasks and special object. Therefore, fully autonomous mode suggested to be selected in the situation of special tasks or special sub-task.

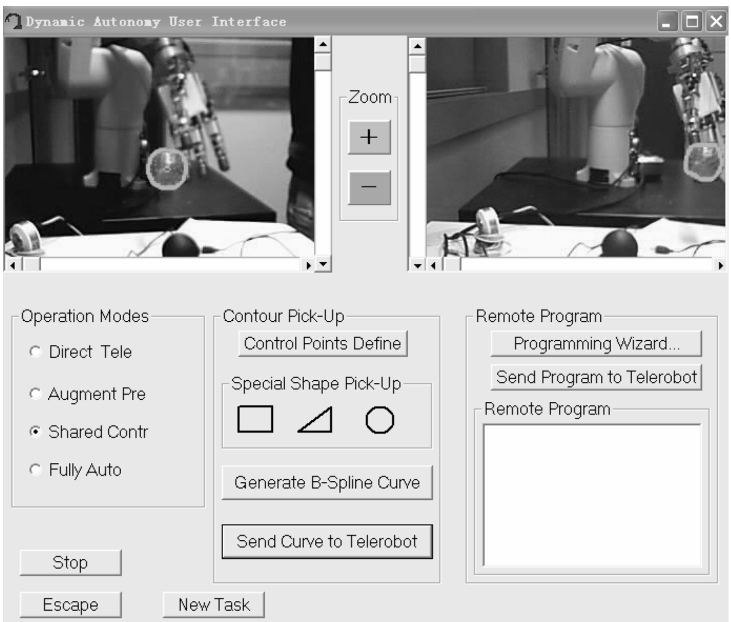
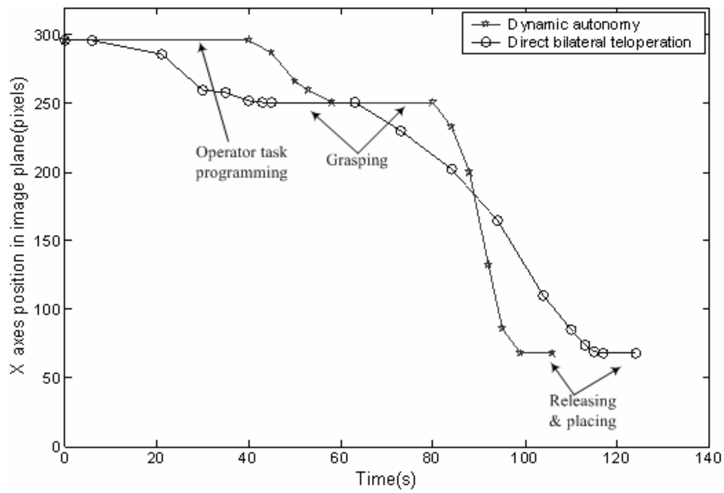



**Fig. 2.** Flow of shared autonomy remote programming

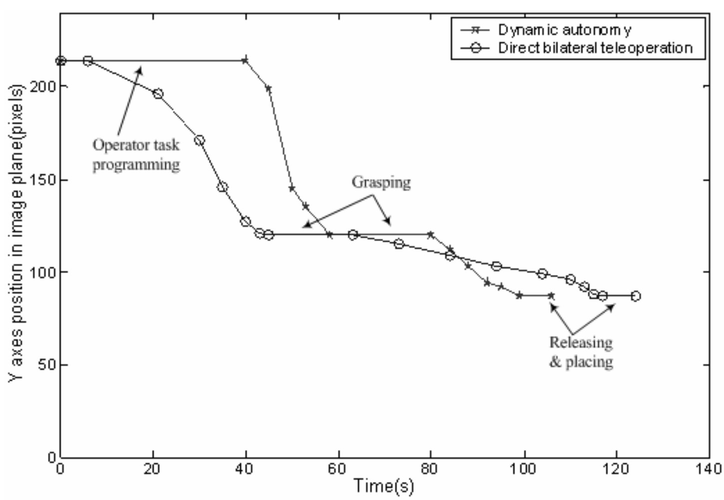
**Augmented Reality Predictive Display.** A virtual robot is registered to the real video image and becomes an interface between the operator and the telerobot, thus the time-delay is separated from interactive of the human-virtual robot (more details see [11]).

When shared autonomy encounter emergencies, by human operator intervene, this mode has the chance to be selected, especially when remote environment is unstructured.

**Direct/bilateral Tele-operation.** Direct bilateral tele-operation consists of transmitting video to local station. The operator can operate the telerobot by using control device such as space ball, data glove or joy stick. Since timedelay affect the quality of bilateral mode seriously, this mode only has the chance to be selected when without large timedelay.




(a) Trajectory of X axes




(b) Trajectory of Y axes

**Fig. 3.** Trajectory of robot finger by in the right camera image plane

## 6 Experiments

As described in section 5, shared autonomy is common use control mode while augmented reality predictive display is a standby control mode. Therefore, experiment of object tracking, reaching and grasping is implemented by the dynamic autonomy system include these two modes. The task of the experiment is to put a bottle into a bucket which both in remote environment. The dynamic autonomy tele-operation system is constructed by autonomous telerobot (Staubli RX40), robot server and local user interface. The experiment is take in local area network, in order to simulate the large time delay, fix manmade 6 seconds round trip time delay is added in the process. As a comparison, experiment of bilateral tele-operation also took under the same condition. In the dynamic autonomy tele-operation, when the contours of bottle and bucket are picked-up, the remote program set as described in section 4.2.

The experimental results as show in Fig.3, total time of dynamic autonomy tele-operation is shorter than fully bilateral tele-operation although task programming in dynamic autonomy could spend longer time. Furthermore, for the operator, dynamic autonomy is easier and lower pressure than bilateral tele-operation

## 7 Conclusions

User interface for dynamic autonomy tele-operation is proposed in this paper. The interface provide smooth switch between four control modes. As default and common use mode, shared autonomy tele-operation is friendly and easy for an operator who unfamiliar with tele-operation. Furthermore, shared autonomy tele-operation can cope with large time delay with high operational efficiency. The augmented reality predictive display mode and fully autonomous mode act as safe supporter for dynamic autonomy system. However, the deficiency to be improved is that the process of contour pick-up in natural user interface is not so easy for operator and could spend long time to do it.

## Acknowledgments

The authors thank Dr. Ming Xie of Nanyang Technological University for providing useful insights contribution to this work.

## References

1. Hauck, A., Sorg, M., Farber, G., Schenk, T.: A Biologically Motivated Model for the Control of Visually Guided Reach-To-Grasp Movements. In: Proceedings of the 1998 IEEE ISIC/CIRA/ISJ, Gaithersburg, pp. 295–300. IEEE Press, New York (1998)
2. Steinfeld, A.: Interface Lessons for Fully and Dynamic Autonomy Mobile Robots. In: IEEE International Conference on Robotics and Automation (ICRA).IEEE Press, New York (2004)
3. Fong, T., Thorpe, C., Baur, C.: Collaboration, Dialogue, and Human-Robot Interaction. In: Robotics Research: The 10th International Symposium, vol. 6, pp. 255–270 (2003)

4. Fong, T., Thorpe, C., Baur, C.: Collaboration, Robot, asker of questions. Robotics and Autonomous Systems 42, 235–243 (2003)
5. Birk, A., Pfingsthorn, M.: A HMI Supporting Adjustable Autonomy of Rescue Robots. In: Bredenfeld, A., Jacoff, A., Noda, I., Takahashi, Y. (eds.) RoboCup 2005. LNCS (LNAI), vol. 4020, pp. 255–266. Springer, Heidelberg (2006)
6. Bruemmer, D.J., Dudenhoeffer, D.D., McKay, M.D., Anderson, M.O.: Dynamic-Autonomy for Remote Robotic Sensor Deployment. In: Proc. Spectrum 2002, Reno (2002)
7. Masahide, S., Norikazu, I., Motonori, D.: Tracking of Fingers in Dynamic Image of Omni-direction Camera by CONDENSATION Algorithm. In: Proc ISCIE Int Symp. Stoch Syst. Theory Appl (Inst. Syst. Control Inf. Eng.), vol. 36, pp. 14–19 (2005)
8. Isard, M., Blake, A.: CONDENSATION – Conditional Density Propagation for Visual Tracking. Int. J. of Computer Vision 29(1), 5–28 (1998)
9. Li, M., Fu, Y., Li, S., Zhu, W., Zhao, D.: Robotic Hand-Eye Coordination Control Based on Binocular Disparity and Active Contour. Robot 30(3), 248–253 (2008)
10. Menet, S., Saint, M.P., Medioni, G.: B-Snakes: Implementation and application to stereo. In: DARPA Image Understanding Workshop, pp. 720–726 (1990)
11. Xiong, Y., Li, S., Xie, M.: Predictive display and interaction of telerobots based on augmented reality. Robotica 24(1), 447–453 (2005)