"Toward remote teleoperation with eye and hand: A first experimental study"

Despinoy, Fabien ; Leon Torres, Jonathan Roberto ; Vitrani, Marie-Aude ; Herman, Benoît

Abstract

Despite the continuous improvement of master interfaces, distant robot teleoperation remains a challenging task. In many applications, including robot-assisted single port laparoscopic surgery and NOTES, the operator has only an indirect vision of the remote environment, provided by a video camera usually mounted on the robot end-effector itself and displayed on a monitor. Driving such an eye-in-hand robot smoothly in real time is known to be difficult. Even a skilled user will likely generate a succession of independent translations and rotations punctuated by frequent stops, in order to keep the target inside the camera field of view. Since one can assume that she/he will look most of the time at the target on the video monitor, our idea is to gather additional information on the remote target with an eye-tracking device, and to use it to facilitate teleoperation. This paper introduces a novel remote teleoperation system that comprises such a device, as well as an 6-axis joystick. ...

Document type : *Communication à un colloque (Conference Paper)*

Référence bibliographique

Despinoy, Fabien ; Leon Torres, Jonathan Roberto ; Vitrani, Marie-Aude ; Herman, Benoît. *Toward remote teleoperation with eye and hand: A first experimental study.*3rd Joint Workshop on New Technologies for Computer/Robot Assisted Surgery (CRAS2013) (Verona, Italy, du 11/09/2013 au 13/09/2013).

Toward Remote Teleoperation with Eye and Hand: A First Experimental Study

Fabien Despinoy, Jonathan Leon-Torres, Marie-Aude Vitrani, and Benoît Herman

***Abstract*— Despite the continuous improvement of master interfaces, distant robot teleoperation remains a challenging task. In many applications, including robot-assisted single port laparoscopic surgery and NOTES, the operator has only an indirect vision of the remote environment, provided by a video camera usually mounted on the robot end-effector itself and displayed on a monitor. Driving such an *eye-in-hand* robot smoothly in real time is known to be difficult. Even a skilled user will likely generate a succession of independent translations and rotations punctuated by frequent stops, in order to keep the target inside the camera field of view. Since one can assume that she/he will look most of the time at the target on the video monitor, our idea is to gather additional information on the remote target with an eye-tracking device, and to use it to facilitate teleoperation. This paper introduces a novel remote teleoperation system that comprises such a device, as well as an 6-axis joystick. Several eye-hand couplings are compared in a first experimental trial. It is shown that the system is effective and intuitive.**

I. Context, Motivation, and Objectives

When operating an interactive robotic device in real time, both user and robot sensorimotor loops act in synergy to perform a desired task under the user's supervision. Implementation of such a system depends on the chosen interaction between user and robot. With most master-slave systems used for teleoperation and virtual reality, interaction is limited to the exchange of kinematic information (e.g. position and/or velocity) with or without haptic feedback, and the slave arm simply copies the operator's motion. In many applications where robotic arms operate in inaccessible locations (e.g. space, nuclear plants, or during minimally access surgery), the robotic end-effector and its environment are captured by a scene camera and displayed on a monitor in front of the operator. The da Vinci teleoperation system from Intuitive Surgical [1] and the DLR MiroSurge [2] are typical examples of such an *hand-in-eye* configuration in surgery, and are known to be easy to drive with their master arms.

This work was supported by Fondation Belge Contre le Cancer (Grant No. SCIE 2006/20), and by own funds from UCL, ISIR, and F.R.S.-FNRS.

F. Despinoy, J. Leon-Torres, and B. Herman are with the Center for Reserach in Energy and Mechatronics, Université catholique de Louvain, Louvain-la-Neuve, Belgium. J. Leon-Torres is also with Université Pierre et Marie Curie, Paris, France. B. Herman is also with Fonds de la Recherche Scientifique – FNRS benoit.herman@uclouvain.be.

F. Despinoy is now with the Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier, Université Montpellier 2 – CNRS (UMR5506), Montpellier, France fabien.despinoy@lirmm.fr.

M.-A. Vitrani is with UPMC Univ. Paris 06, UMR 7222, ISIR, Paris, France marie-aude.vitrani@upmc.fr.

The ongoing trend to reduce invasiveness leads to the design of poly-articulated devices (e.g. iSnake form Imperial College [3]) or continuum robots using concentric tubes (e.g. robots from Dupont at Boston University [4] and Webster at Vanderbilt University [5]). When the camera is embedded in such a snake-like system, creating an *eye-in-hand* configuration as with a flexible endoscope, teleoperation difficulty can increase rapidly. Even a skilled user will not make the most of a 6-axis master interface that could normally induce a smooth and continuous motion. Instead, she/he will likely generate a succession of basic motions in Cartesian or even joint space, punctuated by frequent stops [6]. The main reason is that the target (e.g. tissue to grasp or to cut) must be kept inside the camera field of view during robot motion. A smoother motion would require to combine translations and rotations of the end-effector, with a ratio between linear and angular velocities that depends on the distance to the target, the latter being usually unknown.

The development of numerous eye- and gaze-tracking devices allows to add a new channel for human-robot interaction. Several systems have been developed for providing assistance to disabled people. In [7], authors present an actuated wheelchair with eye-control, using the eye as a computer mouse. The same idea is used in [8] to drive a mobile robot remotely. Yet, the gaze is generally not combined with any other interactive information. Furthermore, user needs to perform a specific eye gesture to send an order to the robot, using his/her eyes in an unnatural way. In this paper, we address the following question: Can natural gaze tracking help to teleoperate a robot? Interestingly, Pinpin *et al.* [9] predict the motion of a robot based on user's natural gaze. The main limitation of their setup is that gaze is only reliable for predicting which actions from a predefined list will occur. To overcome this predefined list, we propose to drive a robot using the user's eye movement, without any pattern recognition. Spindler and Chaumette already implemented a visual servoing of the orientation of a robot-mounted camera, using human eye-movement [10]. This device is however limited to a 2D task. We propose to extend this work to a general 3D task.

The paper introduces a novel teleoperation system that combines joystick and natural gaze inputs in several ways to teleoperate an *eye-in-hand* robot. Section II presents the robotic platform, and 3 control schemes are described in Section III. A first experimental performance assessment is then reported and discussed in Section IV.

II. Remote Teleoperation Platform

A proof-of-concept prototype was built around a 6 degrees-of-freedom (DOF) Adept Viper s650 robot with a MotionBlox 60R real-time controller, as depicted in Fig. 1. The teleoperation system comprises a webcam mounted on the robot end-effector to capture the task environment, a head-mounted eye-tracking camera (EYE-TRAC H6 from Applied Science Laboratories, Bedford, MA), a stand to prevent any head motion with respect to a monitor where the webcam images are displayed, and a 6-axis joystick (SpaceNavigator from 3Dconnexion, Waltham, MA).

The platform specificity lies in the original combination of control interfaces. The 6-axis joystick is used to drive the robot in a pose/velocity scheme: Robot linear and angular Cartesian velocities are linked to the joystick position and orientation imposed by the user's hand. An eye-tracker device is also used to record user's pupils movements and to compute gaze direction with respect to the monitor. These two interfaces can then be combined in several ways to control the robot more easily and efficiently.

III. Control Schemes with Eye and Hand Inputs

When performing a task on an object—directly by hand or with any teleoperation means—one usually looks at this object or point of interest (POI): The brush tip while painting, the nail when holding a hammer, or the blood vessel the surgeon wants to clip. This basic assumption lead us to use a gaze-tracking device as a supplement to a joystick or master arm so as to make teleoperation faster, smoother, and easier. Knowing *where on the monitor* the operator looks thanks to a gaze-tracker, a couple of end-effector motions can be automated, or information about robot-target distance or relative orientation can be deduced, even *without any image processing* of what is displayed. We designed and implemented 3 control schemes that use in different ways the 2D-coordinates of the focus point on the screen.

The two first possibilities are to maintain the target at the center of the screen, either by correcting automatically the end-effector orientation during a manual translation (*autonomous rotation*, see Fig. 2), or conversely, by shifting the robot when the user rotates it (*autonomous translation*). In these two similar scenarios, the user has to control only 4 DOF manually, the remaining two being handled by the controller thanks to the eye-tracker 2D signal. This signal can be seen as a measure of the on-screen distance between POI and monitor center. The control loop then simply tries to minimize this distance by using the 2 automatic rotational or translational DOF, respectively.

A third possibility is to estimate the distance between end-effector and target, so as to make the robot rotate around the target instead of around its end-effector. This can be achieved iteratively by measuring the user's eye motion induced by a robot displacement, in order to move the center of rotation from the end-effector to the target. In opposition to the first two above, this *distance estimation* control scheme does not induce any autonomous motion. All translations and rotations are then induced manually but the rotations are more natural. If this distance is underestimated, a manual rotation will induce a shift of the POI in the image in the opposite direction. Conversely, an overestimation of the distance will also translate the POI in the same direction. An iterative algorithm can use this error measured by the gaze-tracker to update the distance estimation.

All these modes, along with a fully manual mode without eye-tracking, were implemented on the robot controller. After initial calibration of the gaze direction with respect to the monitor center, signals form joystick and eye-tracker are acquired, filtered and scaled. They are then used and combined, depending on the scheme, to update the robot pose in real-time.

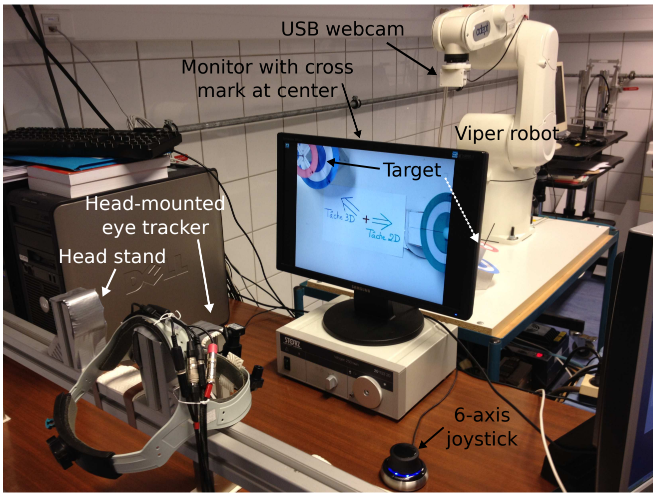

Fig. 1. Experimental setup built around the Viper robot.

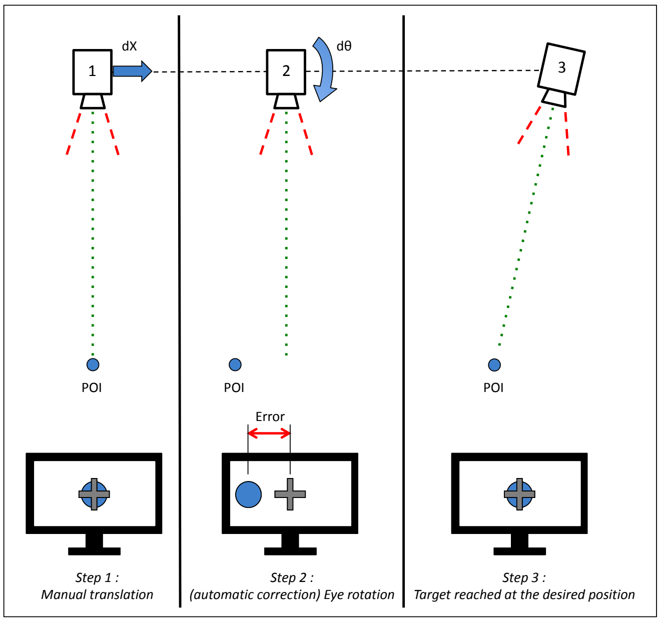

Fig. 2. Autonomous rotation to bring the target back to the center of the monitor after a manual translation.

IV. Performance Assessment

An experimental study was conducted in order to assess the feasibility and usability of the proposed teleoperation system, and to compare the performance of the proposed control schemes. This section describes the teleoperation task proposed, the metrics used for performance comparison, the experimental protocol and the results.

A. Teleoperation Task

A wide variety of tasks can be performed with a teleoperated robot, depending on the context and the tool held by the robot. However, most tasks can be divided in two phases, that can occur at the same time but are more frequently successive: First, the robot is driven to approach a target, and then an action is performed with the tool.

Since our teleoperation system only intends to increase the performance of robot motion, the task is a simple point-to-point displacement combined with an effector reorientation so as to reach a target under a prescribed pose. A specific target (see Fig 3) was designed to ensure that all 6 DOF must be used. It also gives to the operator visual clues on the position and orientation of the end-effector.

B. Metrics

Task duration is the main criterion to compare the schemes. Duration corresponds to the time elapsed between the first operator manual instruction with the joystick and the arrival next to the target. A tolerance zone was programmed to ensure the experiment stop and each linear or angular tolerance was tuned experimentally to define the task difficulty. Moreover, during the experiments, articular robot positions and input signals were also recorded. Graphical analysis of the kinematic data showed that the task could be decomposed into two successive phases: A first approach phase where the error decreases rapidly, followed by an adjustment phase where motions are much shorter and slower. It was then decided to compute the duration of these two phases for each trials, as secondary metrics.

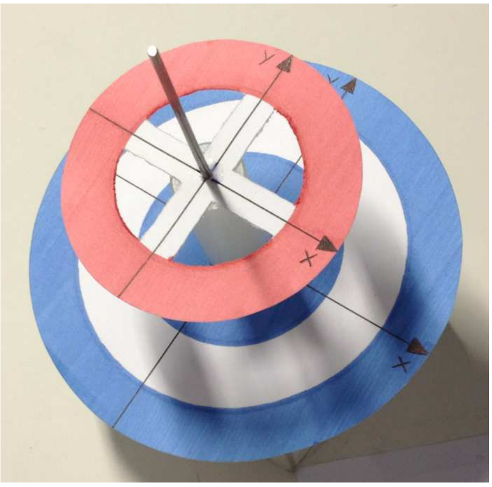

Fig. 3. Target that requires 6 DOF motions of the end-effector.

Subjects were also asked to estimate, on an integer scale from 0 to 10, the performance, precision and intuitiveness of each mode. Performance indicates speed perception to perform the task. Precision reflects the possibility to drive the robot exactly where the operator wants to. Finally, intuitiveness reveals the operator mental workload to control the system correctly.

C. Experimental Protocol

Six researchers from the CEREM were enrolled in the experimental campaign and repeated the same target-reaching task with the various control schemes, being asked to be as quick and precise as possible. Each subject started the protocol with the fully manual mode so as to learn to use the system. He/she repeated the task until the reach of a plateau on the learning curve fitted and refreshed after each trial. The last 3 trials (on the plateau) were taken into account for statistical analysis. Afterwards, the subject tried the 3 assisted modes in randomized order. For each assisted mode, 8 repetitions were performed. The 4 best out of the 5 last trials were taken for analysis.

D. Statistical Analysis

Data were analyzed using General Linear Model (GLM) with JMP 9.0.3 and SAS 9.3 statistical software. The mixed model contained the following effects: subject and control scheme, and the two-factors interaction. Control scheme was defined as fixed factor for the ANOVA (analysis of variance). Subject and its interaction with the fixed factor were defined as random. The model was solved using REML (Restricted Maximum Likelihood). The Tukey HSD test of multiple comparisons was used to compare modalities of significant factors.

E. Results

Figure 4 depicts the average value and standard deviation of approach, adjustment and total duration for all trials of all subjects, for each control mode. Standard deviation for each mode is similar, ensuring that the ANOVA can model the data properly. One can observe that the autonomous translation mode is the slowest, especially during the adjustment

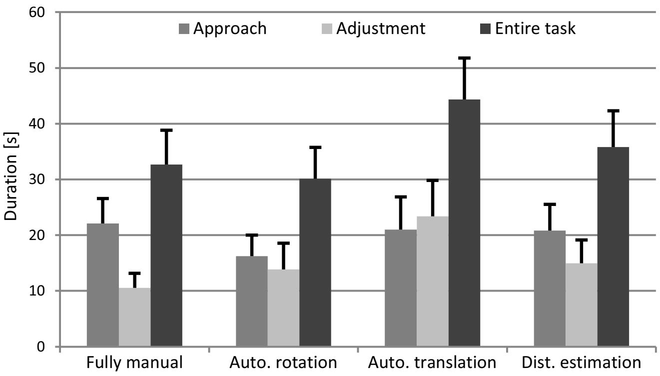

Fig. 4. Average value and standard deviation of task duration for all subjects and repetitions.

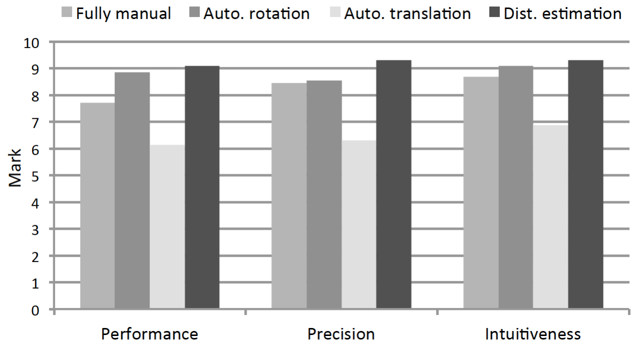

Fig. 5. Average value of subjective indicators normalized between subjects.

phase. This is confirmed by the statistical analysis, since this mode is significantly different from all others regarding the adjustment duration. For the approach duration, the only significant difference is between the fully manual (slowest) and the autonomous rotation (fastest). Finally, analysis of the overall task duration shows that the autonomous translation (slowest) is significantly different form the autonomous rotation (fastest).

Average values of subjective indicators are presented on Fig. 5, after normalization. It appears that the autonomous translation mode is believed to be the less efficient, precise and intuitive mode. In opposition, the distance estimation is perceived as the best mode, followed closely by the autonomous rotation mode.

F. Discussion

This first experimental session demonstrates the feasibility and efficiency of our teleoperation approach using eye and hand. Although the statistical analysis does not highlight many large differences, results are promising. In particular, it seems that the approach phase can be shortened with a gaze-tracker, since all 3 assisted modes are faster that the fully manual one. This might be thanks to the fact that, during large motions, it is easier to keep the target close to the center of the monitor with assistance.

However, during the adjustment phase, a fully manual control seems faster. Subjects declared that this was because it was difficult to keep the focus on the center of the target while perform fine adjustments, since several parts of the targets were to be used to check alignment. Therefore, autonomous motions were sometimes felt like a small disturbance during the final adjustment phase. This suggests that the eye-tracking system should be switched off after the initial approach phase.

Additionally, subjective indicators show the same trend than the objective measurements, with the autonomous translation mode that seems less efficient, precise and intuitive. This suggests to abandon this mode and to keep improving the two other assisted schemes.

V. Conclusion and Future Work

In this paper, we introduced an original system that uses both hand and eyes as master inputs to drive a robot in real time. Three control schemes were proposed that use the gaze information in different ways to reduce the number of DOF that must be driven with a joystick, or to facilitate rotation around the target. A first experimental session demonstrated that teleoperation with eye and hand is feasible and efficient for remote teleoperation with an *eye-in-hand* configuration.

An improved version of the system is currently being developed, using a 3D tracking system to allow head motion, in order to decrease mental and physical workload. As a consequence, this should probably increase the performance level of the whole system. In addition, the task difficulty should be adjusted, since duration decomposition showed that nearly half of the time was spent for small adjustments that might not be required in real applications. The system and control schemes should also be studied in case of a change of POI during the task, which is not unusual.

Acknowledgement

Authors would like to thank warmly all colleagues from the CEREM who took part in the experiments. Authors are also grateful to Catherine Rasse, from the Institute for Multidisciplinary Research in Quantitative Modelling and Analysis (UCL), for her support and advices regarding experimental methodology and statistical analysis.

References

[1] G. H. Ballantyne and F. Moll, "The da Vinci telerobotic surgical system: The virtual operative field and telepresence surgery," *Surgical Clinics of North America*, vol. 83, no. 6, pp. 1293–1304, Dec. 2003.

[2] U. Hagn, R. Konietschke, A. Tobergte, M. Nickl, S. Jrg, B. Kbler, G. Passig, M. Grger, F. Frhlich, U. Seibold, L. Le-Tien, A. Albu-Schffer, A. Nothhelfer, F. Hacker, M. Grebenstein, and G. Hirzinger, "DLR MiroSurge: A versatile system for research in endoscopic telesurgery," *International Journal of Computer Assisted Radiology and Surgery*, vol. 5, no. 2, pp. 183–193, Mar. 2010.

[3] R. C. Newton, D. P. Noonan, V. Vitiello, J. Clark, C. J. Payne, J. Shang, M. Sodergren, A. Darzi, and G. Z. Yang, "Robot-assisted transvaginal peritoneoscopy using confocal endomicroscopy: A feasibility study in a porcine model," *Surgical Endoscopy*, vol. 26, no. 9, pp. 2532–2540, Sep. 2012.

[4] P. E. Dupont, J. Lock, B. Itkowitz, and E. Butler, "Design and control of concentric-tube robots," *IEEE Trans. Robot.*, vol. 26, no. 2, pp. 209–225, Apr. 2010.

[5] D. C. Rucker, B. A. Jones, and R. J. Webster, "A geometrically exact model for externally loaded concentric-tube continuum robots," *IEEE Trans. Robot.*, vol. 26, no. 5, pp. 769–780, Oct. 2010.

[6] J. Cotti, A. Guillaume, N. Alahyane, D. Pelisson, and J. L. Vercher, "Adaptation of voluntary saccades, but not of reactive saccades, transfers to hand pointing movements," *Journal of Neurophysiology*, vol. 98, pp. 602–612, 2007.

[7] R. Barea, L. Boquete, M. Mazo, and E. Lopez, "System for assisted mobility using eye movements based on electrooculography," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 10, no. 4, pp. 209–218, Dec. 2002.

[8] H. O. Latif, N. Sherkat, and A. Lotfi, "Teleoperation through eye gaze (TeleGaze): A multimodal approach," in *Proc. IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Guilin, China, Dec. 19–23, 2009, pp. 711–716.

[9] L. K. Pinpin, R. S. Johansson, C. Laschi, and P. Dario, "Gaze interface: Utilizing human predictive gaze movements for controlling a hbs," in *Proc. IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, Scottsdale, AZ, Oct. 19–22, 2008, pp. 158–162.

[10] F. Spindler and F. Chaumette, "Gaze control using human eye movements," in *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, Albuquerque, NM, Apr. 20–25, 1997, pp. 2258–2263.