Medical Informatics and the Internet in Medicine
September 2006; 31(3): 205–218

Efficacy of stereoscopic visualization and six degrees of freedom interaction in preoperative planning of total hip replacement

DEBORA TESTI[1], RICCARDO LATTANZI[1], MARCO BENVEGNÙ[1], MARCO PETRONE[2], CINZIA ZANNONI[2], MARCO VICECONTI[1] & ALDO TONI[1,3]

[1]*Laboratorio di Tecnologia Medica, Istituti Ortopedici Rizzoli, Bologna, Italy* [2]*CINECA, Supercomputing Centre, Bologna, Italy and* [3]*1st Division, Istituti Ortopedici Rizzoli, Bologna, Italy*

(Received September 2005; accepted June 2006)

Abstract
The aim of this study was to assess the accuracy of a six-degrees-of-freedom application for pre-operative planning of total hip replacement in a virtual reality (VR) environment. A test was performed estimating the system inherent accuracy. The users can move objects in the VR environment with an intrinsic accuracy almost four orders of magnitude greater than the object dimension. A second unambiguous and relevant task was defined to assess the accuracy achievable with the interface in a specific planning task. The results were compared with those obtained with 2D interfaces for both the stem and the cup component. The RMSE was assumed as an indicator of the achievable accuracy. The accuracy of the immersive interface was comparable with that achievable with a standard mouse – monitor interface. The users were consistent using the VR interface, confirming the high usability of the new interface and the steep learning curve of users unfamiliar with the new environment. This study has demonstrated that the application of VR environment for pre-operative planning of total hip replacement may help to shorten the duration of the positioning and to yield consistent results even with first-time users.

Keywords: *Virtual reality environments, pre-operative planning, total hip replacement*

1. Introduction

Owing to great improvements in computer science and decreasing computational costs, in recent years several specialized virtual reality (VR) approaches have been proposed in medicine. The use of computational methods to provide a multimedia environment that simulates reality has become of great benefit in the training of clinical personnel with virtual patients rather than with real patients [1]. VR techniques are becoming increasingly common in neurosurgery [2], craniofacial surgery [3], and orthopaedic surgery [4], since they help in precise preoperative planning and with minimally invasive surgical procedures.

Preoperative planning is a fundamental phase in total hip replacement (THR) surgery [5], and the advantage of using a three-dimensional (3D) environment to plan the operation is

---

Correspondence: D. Testi, Laboratorio di Tecnologia Medica, Istituti Ortopedici Rizzoli, Via di Barbiano, 1/10, 40136 Bologna, Italy. Tel: +39-051-6366864. Fax: +39-051-6366863. E-mail: testi@tecno.ior.it

ISSN 1463-9238 print/ISSN 1464-5238 online
DOI: 10.1080/14639230600886900

demonstrated [6]. Many systems for 3D preoperative planning of THR have been developed worldwide in the last few years [4,7,8]. These systems allow surgeons to position prosthetic models within a 3D navigation environment, combining 2D and 3D graphical representation of the patient anatomical data.

Although it is more intuitive to grasp and examine a virtual object working with an immersive VR system, the navigation through the virtual 3D environment is obtained using a mouse with two degrees of freedom and a flat screen for pseudo-3D interaction. This strategy is questionable, however, since the accuracy in positioning the implant using a 3D preoperative planning software has been shown to be strongly affected by the graphical user interface [9], and to the authors' knowledge, no published studies have investigated the accuracy of such a 3D monomodal interface in preoperative planning of THR surgery.

The present work is aimed at comparing the positioning accuracy achieved with conventional mouse–monitor interfaces with that obtained with a stereoscopic display and a six-degrees-of-freedom (6DOF) tracker.

2. Materials and methods

The test application was developed using the Multimod Application Framework (MAF) (http://www.openmaf.org). MAF is based on the Visualization Tool Kit (VTK), a graphical library for visualization, 3D graphics, and image processing, [10–12] and wraps functionalities from other Open Source libraries such as the Insight Tool Kit. To develop the immersive interface, new classes that handle tracking and interaction in 6DOF environments were added to the VTK distribution.

The application was executed on the Intersense tracking system IS-900 VWT, which is an immersive system for workbench. The projection table used for display is a Baron workbench (BARCO Projection Systems Inc., Kennesaw, GA), which is a motorized table that can be oriented at any angle between 0° and 90° (Figure 1). The IS-900 tracking station also includes

Figure 1. Baron workbench.

a lightweight 6DOF stylus with two buttons (Figure 2) and supports StereoGraphics CrystalEyes 3D shutter glasses.

2.1. Inherent accuracy

The first part of the study was aimed at evaluating the inherent accuracy of the VR environment. For this purpose, it was necessary to define an unambiguous task, independent of the visualization technique, and future clinical use. Two three-dimensional objects were created, and polygonal surfaces of the two objects were obtained from the NURBS models using a triangulation algorithm (Unigraphics v.16, UGS, USA).

Each of the two objects consisted of three conical surfaces assembled together and having different colours. When the two objects, hereinafter referred as target geometry (Figure 3a) and movable geometry (Figure 3b), were loaded in the same position, the resulting geometry was a single block of three pairs of cones facing each other along three orthogonal axes, with the apex of each cone in one geometry touching the apex of the cone with the same colour in the other geometry (Figure 3c).

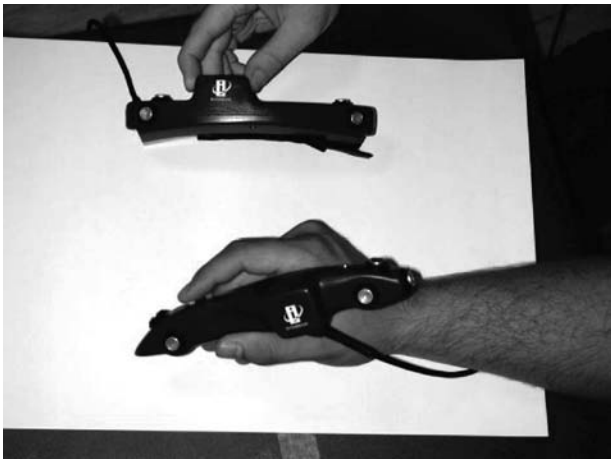

Figure 2. 6DOF stylus.

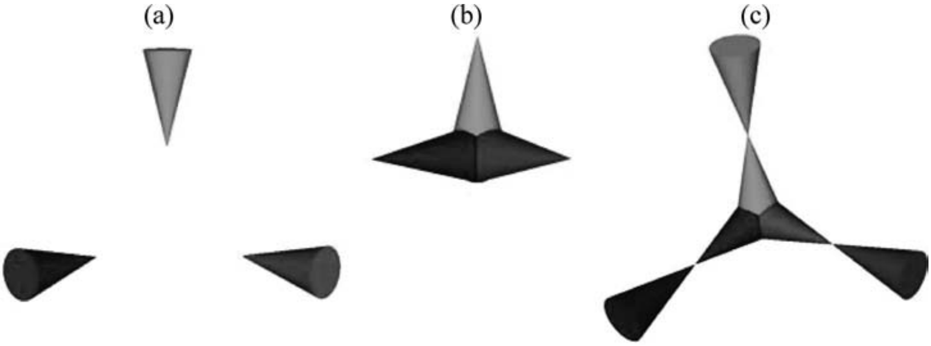

Figure 3. (a) Target geometry. (b) Movable geometry. (c) Two objects placed in the same position, resulting in a single-block perfectly aligned.

Five users were enrolled for the inherent accuracy study and were asked to perform a series of five tests. None of them had ever used a 6DOF tracker in a VR environment. Before the testing started, the users were first introduced to the VR environment and to the different interaction modalities on the tracker tool, and they were allowed to use the interface in order to ensure they had understood the capabilities of the system.

For the testing procedures, the target geometry was loaded in a known position, while the movable geometry was loaded in the scene in a randomized initial position in order to avoid the user remembering the task execution. The user was then asked to use the stylus pen to position the movable geometry as close as possible to the reference position, so that the tips of the cones in the datasets fitted exactly. The starting position was chosen in a way that forced the user to make use of all 6DOF in order to move the object toward the correct position. When the user thought they had done their best, the position of the movable geometry thus obtained was recorded and compared with the known position of the target geometry.

The error in the final pose was decomposed in the three translations and three rotations that needed to be applied to the position achieved by the user in order to reach the correct position. Since the geometries used were symmetric along the main axes, a mean error was considered for position and orientation. The average position error was normalized by the largest linear dimension of the target geometry (distance between the tips of the target, 84.85 mm) and was thus expressed in %, while the average rotation error was reported in degrees. The results were analysed with a repeated-measurement ANOVA test.

2.2. Accuracy of the interface in a pre-operative planning application

The second part of this study was aimed to assess the accuracy achievable with a 6DOF interface in an application for pre-operative surgical planning for both the stem and cup components of total hip prosthesis.

A task was defined that was complex enough to account for most of the specificity imposed by the chosen application but without the intrinsic uncertainty that surgical planning usually implies. Therefore, it was used as an unambiguous and relevant task, described in a similar study [9]. The task aim was to position the prosthesis exactly over a reference surface, which represents the correct position inside the host bone (Figure 4).

In the VR environment, the same femur and ileum were visualized as a polygonal surface. The 3D rendering of the surface representing the reference position was placed in its correct location inside the femur and given a colour different from that of the femur. The opacity of the surface representing the femur was reduced so that the reference surface was clearly visible. Thereafter, the surface of the prosthesis, obtained from NURBS model, was visualized with another colour (Figure 5).

2.2.1. Stem component positioning. Five users were asked to perform the task five times. The test was unconstrained: the user was asked to reach, in the shortest amount of time and as accurately as possible, the position believed to represent the correct position.

Each repetition produced an error vector composed of three translations and three rotations. However, for the statistical treatment it was necessary to have a single error indicator. The geometry of the prosthesis was thus described by a polyhedral surface, i.e. a collection of triangles, so that it was possible to compute the distance of each vertex from its nominal position, which it would have in the reference position. The root mean square average of the resulting distance error (RMSE), computed over the 3357 vertices describing the implant surface, was assumed as a synthetic error indicator for the consequent statistical analysis.

Figure 4. Test performed in a previous study by [9] and used as term of comparison for the accuracy in the pre-operative planning of the stem component positioning.

The results were compared with those obtained performing the same task with different kinds of interfaces based on a standard mouse–monitor interaction as already published [9]. The effects of interface, operator, or testing session were investigated using a two-way factorial ANOVA. Single effects were analysed using Scheffé post-hoc test. One user was asked to perform the test 10 times, and for each test session the time necessary to perform the task was also recorded, in order to define the learning curve of the VR application. Furthermore, the system overall usability was evaluated by interviewing the users at the end of the test sessions.

2.2.2. Cup component positioning. Five users were asked to perform the task for ten times. The tests were unconstrained in the sense that the user was asked to reach, in the shortest amount of time and as accurately as possible, the position that they believed to represent the correct position. In this positioning evaluation, the number of test repetitions was increased as a preliminary study revealed a higher variance in the cup component positioning with respect to the stem positioning.

Each repetition produced an error vector composed of three translations and three rotations. Owing to the symmetry of the component, the rotations were expressed as the resulting direction of the axis of the cup. As before, the geometry of the prosthesis was described by a polyhedral surface, i.e. a collection of triangles. For each vertex, the distance of such vertex from its nominal position that it would have in the reference position was computed. The root mean square average of the resulting distance error (RMSE), computed over the 3330 vertices describing the implant surface, was assumed as a synthetic error indicator for the consequent statistical analysis.

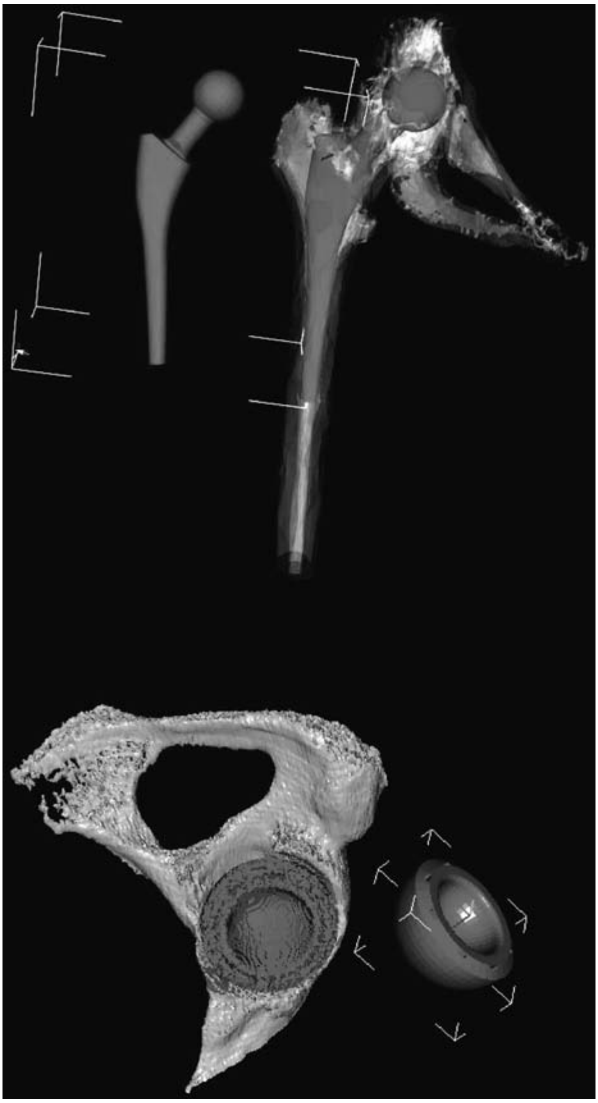

Figure 5. Users had to place the surface of the prosthesis over the reference surface, positioned inside the surface of the femur or of the ileum.

The system overall usability was evaluated by interviewing the users at the end of the test sessions. The results were compared with those obtained asking the users to perform the same task ten times with the standard mouse – monitor interface of the HipOp$^{\copyright}$ pre-operative planning software [9,13] (Figure 6) and with a standard mouse – monitor surface rendering view (Figure 7).

For each test session, the time necessary to perform the task was also recorded for both user interfaces. The effects of system and operator were investigated using a two-way factorial ANOVA. Single effects were analysed using the Scheffé post-hoc test.

Figure 6. Test performed with the standard mouse – monitor interface of the HipOp software for the pre-operative planning.

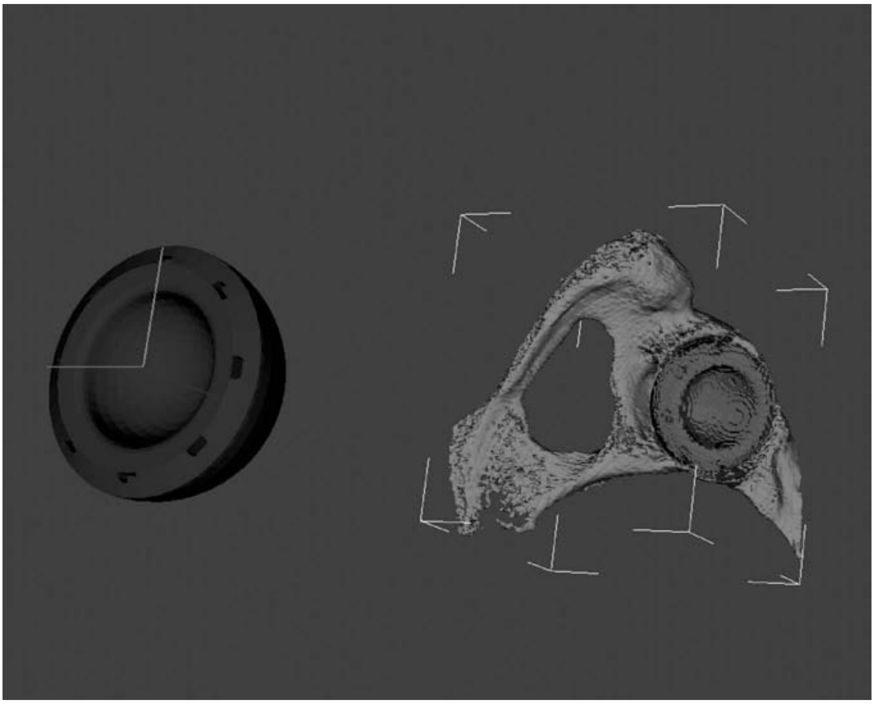

Figure 7. Test performed with the surface view with a mouse – monitor interface.

3. Results

3.1. Inherent accuracy

The mean error for position was always below 0.015 mm, about 0.02% of the greatest dimension involved (Figure 8). For rotations, the mean error was less than 1.5° for every user (Figure 9).

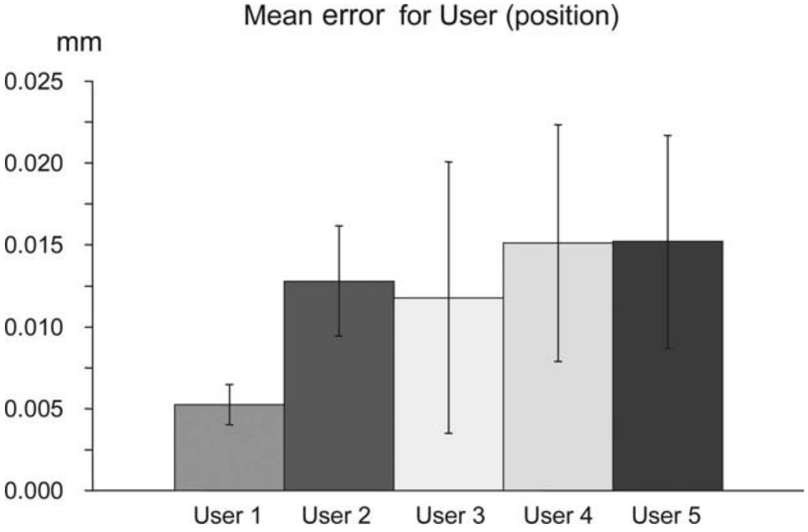

Figure 8. Translation error for each user averaged over the five test sessions.

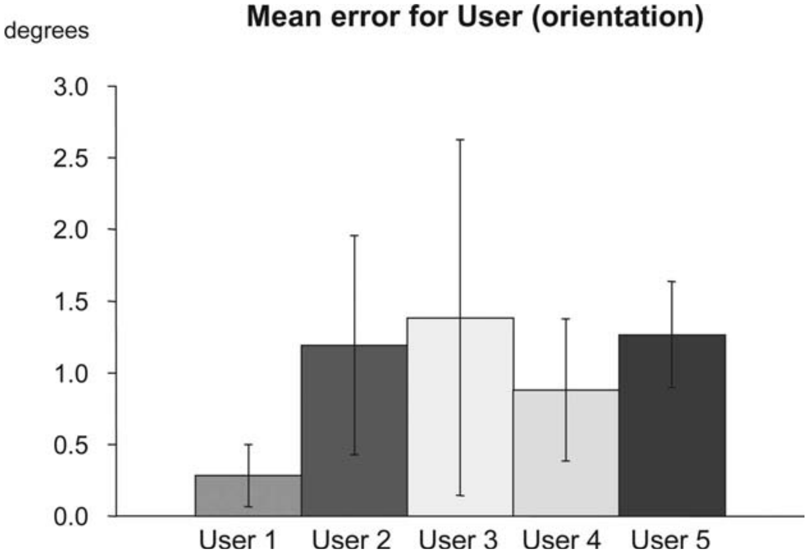

Figure 9. Rotation error for each user averaged over the five test sessions.

One of the users was more accurate than the others, but this difference was not statistically significant (ANOVA, $p > 0.05$). Moreover, the ANOVA analysis showed that there was no statistically significant effect of the user or of the testing session for either translation or rotation errors (ANOVA, $p > 0.05$).

3.2. Accuracy of the interface in a pre-operative planning application

3.2.1. Stem component positioning. For the 6DOF system, the peak RMSE error was 2.25 mm was computed and the users were consistent along the test sessions (Figure 10).

A statistically significant difference among users was reported (ANOVA, $p < 0.05$). In particular, User 4, whose error was always less than 1 mm, was significantly more accurate than User 5 (Scheffé post-hoc test, $p < 0.05$). The other differences between users were not statistically significant.

No learning curve was found for the application in terms of accuracy, since the RMSE decreased and increased randomly along the test sessions. Moreover, the statistical analysis showed that the RMSE did not change significantly among the testing sessions (ANOVA and Scheffé post-hoc test, $p > 0.05$). On the contrary, the time needed by the user to perform the task decreased with the number of testing sessions increasing (Figure 11).

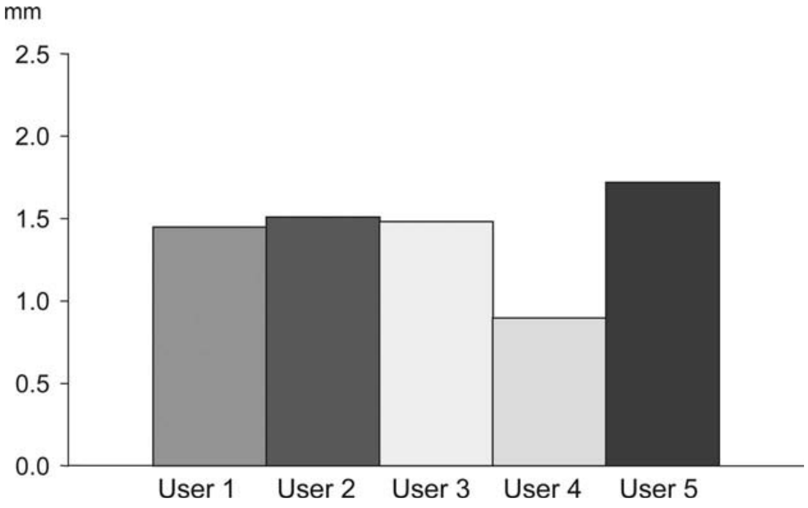

Figure 10. RMSE error for each user averaged over the five test sessions with the stereoscopic vision.

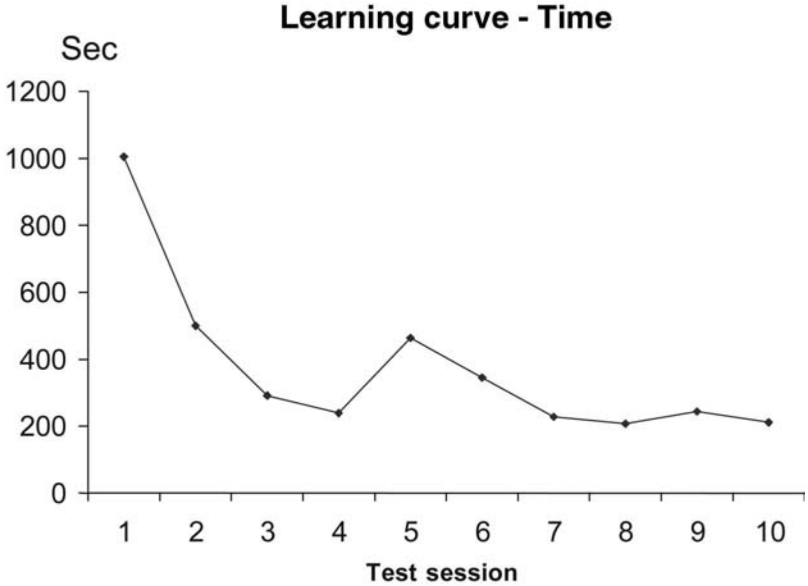

Figure 11. Time necessary to perform the task for the 10 test sessions for one user.

Table I compares the accuracy achieved with different interfaces as reported in a previous study [9] with those obtained in the present study.

3.2.2. Cup component positioning. For the 6DOF system, the peak RMSE error was 2.97 mm, and the users were consistent along the test sessions (Figure 12).

Table II compares the positioning accuracy obtained with the 6DOF system with stereoscopic vision, with a standard mouse–monitor surface rendering view (3DOF non-stereoscopic monomodal display) and the standard HipOp mouse–monitor interface (3DOF non-stereoscopic multimodal display).

The differences among the systems were statistically significant for RMSE, maximum error, and time necessary to perform the task (ANOVA and Scheffé post-hoc test, $p < 0.05$). All indicators also vary significantly among users (ANOVA and Scheffé post-hoc test, $p < 0.05$).

Table I. RMSE error for different interfaces.

Interface	Window	Mean	SD	Peak
Orthogonal slices	Three-pane	1.0	0.6	2.1
3D rendering	Single	1.5	1.7	6.6
6DOF stereo vision	Single	1.4	0.3	2.3
Multimodal display	RX	0.6	0.3	1.8
Multimodal display	RX + CT	0.3	0.2	0.8
Multimodal display	RX + CT + AS	0.3	0.2	0.8
Multimodal display	RX + CT + AS + 3D	0.2	0.1	0.5

The results obtained without the 6DOF were reported in a previous study [9] together with the standard deviation and the peak error. All results are expressed in millimetres.

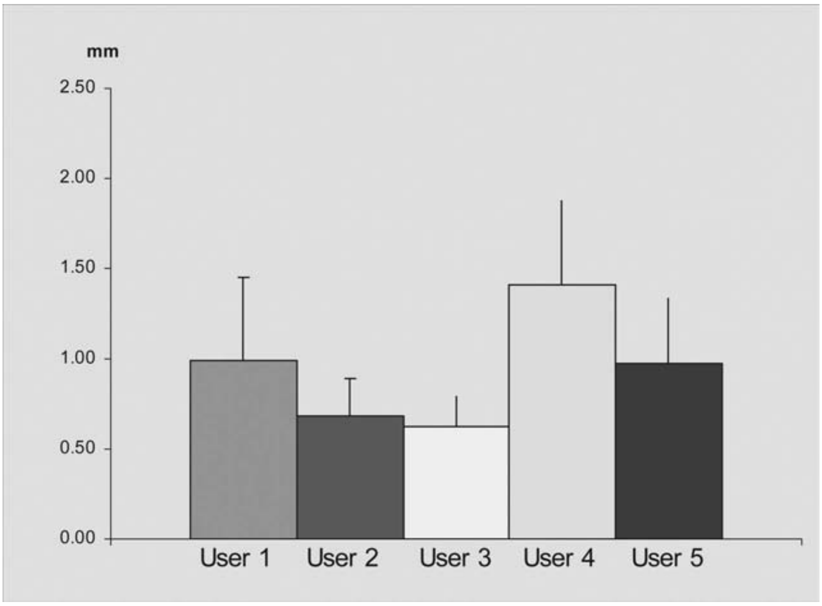

Figure 12. RMSE error for each user averaged over the five test sessions with the stereoscopic vision.

In particular, the multimodal environment showed significantly higher maximum errors and time necessary to carry out the task with respect to the VR environment and the standard monomodal surface view (ANOVA and Scheffé post-hoc test, $p < 0.05$). On the contrary, the monomodal surface view had a significantly better accuracy with respect to the other two interfaces (ANOVA and Scheffé post-hoc test, $p < 0.05$).

Observing the evolution of the accuracy with the testing sessions, it is evident that there is not a progressive decrease in the RMSE error, but the time necessary to perform the task tends to decrease (Figure 13). In all cases, there was no statistically significant difference among the testing sessions (ANOVA and post-hoc Scheffé test, $p > 0.05$).

4. Discussion

The present study aimed to assess the accuracy of a 6DOF application for pre-operative planning of total hip replacement surgery in virtual environments. A first test was performed

Table II. RMSE error for two different interfaces and average time needed to perform the task.

Interface	Mean (mm)	SD (mm)	Max (mm)	Time (s)
6DOF stereoscopic Monomodal display	0.9	0.4	3.0	150
3DOF non-stereoscopic Multimodal display	1.1	0.6	5.3	199
3DOF non-stereoscopic Monomodal display	0.7	0.4	3.0	133

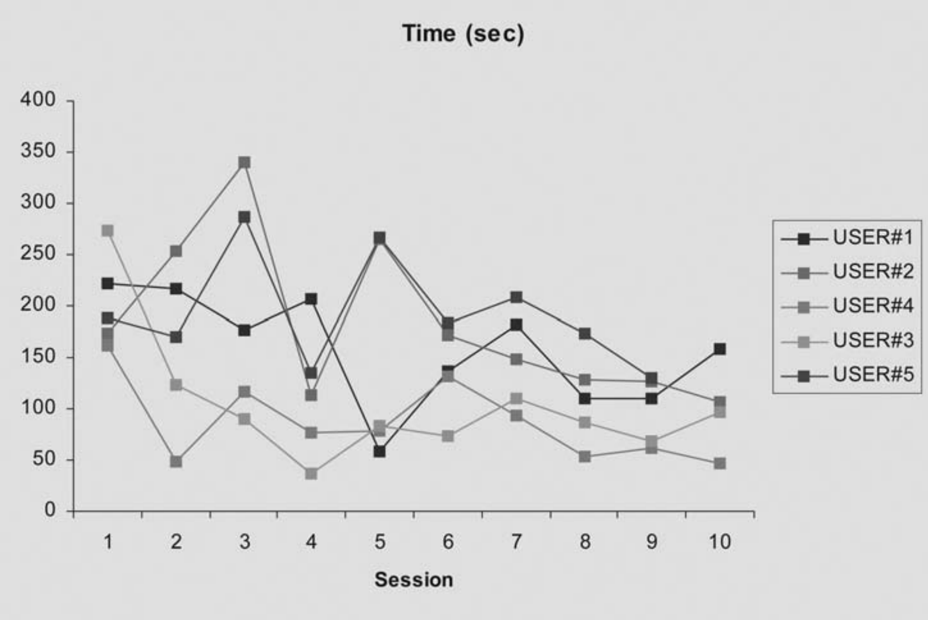

Figure 13. Time needed to perform the positioning with the testing sessions.

to estimate the inherent accuracy of the system. The chosen test was independent from the visualization technique and the geometry involved. A second validation test was carried out to determine the system efficacy in a specific application. An unambiguous and relevant task was defined to assess the accuracy achievable with the interface in pre-operative planning for both the stem and cup components of the total hip implant. The results were compared with those obtained with some 2D interfaces.

The system allows the users to move objects in the immersive environment with an intrinsic accuracy almost four orders of magnitude greater than the dimension of the objects.

RMSE was assumed as an indicator of the accuracy achievable in positioning prosthetic components in a pre-operative planning application. The accuracy obtainable with the immersive interface was lower than or comparable with that achievable with standard mouse–monitor interfaces. In particular, for the stem component, the VR environment yielded an accuracy comparable with that of the standard mouse–monitor interface but lower than that achievable with an interface specialized for pre-operative planning [9]. That suggests that the graphical user interface of any computer-aided surgery system should be highly specialized for the specific planning task, in order to be more accurate. In the case of the cup-positioning test, for which the symmetry of the component geometry and the particular positioning task required a different approach from that in the case of the stem, the single stereoscopic view with 6DOF interaction allowed users to obtain better positioning than the multimodal 2D interface.

However, the users were more consistent with the 6DOF interface, for which the standard deviation was lower than, and even comparable with, that of the 2D interface that showed the best results. This is further proof of the usability of the new interface and also of the easiness with which users accept the new immersive environment, since for most of the participants, it was the first experience with VR and 6DOF interactors.

The general opinion on the stylus pen was that it is easy to use. Some users wanted to hold the pen with both hands in order to stabilize it, a tendency that increased as fatigue started to show. This problem was probably due to the application being too sensitive to rotations.

One of the main limits of this study is the time between each test session. Each user went through the five test sessions in the same day. This caused fatigue, eye strain, and cybersickness. In some cases, the time it took to position the object decreased after one and two test sessions and then increased toward the final sessions. This was probably due to the users' bceoming tired and feeling weakness in their arm, since it is not possible to work with the arm at 90° for very long.

The user is much quicker in using the immersive interface, and the velocity could even be improved with further implementation. Most of the time was spent by users for rotating the scene in order to get the actual position of the objects. Furthermore, an undo function would even speed up a user's performance, since unwanted errors would be avoided. The lack of an undo function also influenced the test; in fact, users sometimes avoided moving the object further since that might yield a worse position.

The current version of the VR system allows the interaction modality to be changed by pressing one of the buttons on the stylus pen. This input strategy is not efficient, since users do not get an overview of the different modalities available. A possible solution could be the implementation of a graphical interface within the VR application where the user can switch between the modes. Instead of virtual menus, voice recognition could be adopted to make the system recognize certain voice commands from the user. The main issues to deal with in a voice-recognition system include the difficulty in users remembering the correct name for each command. There is also the problem of finding a system that can identify the voice of the user in a reliable way, disregarding pronunciations and accents. However, it is an approach that would be easy for the user to understand and easy to expand if further functionalities were needed.

One of the suggestions during the test sessions regards the viewpoint. Some users said that it could be helpful to have the window divided to show the scene from different angles simultaneously. However, instead of splitting the window, predefined viewpoints could be implemented, thus allowing users to quickly select a viewpoint using voice commands, for example. Visual feedback was also suggested so as to provide further information on the anatomical data, such as a frame displaying pre-selected CT slices and their intersections with the objects loaded in the environment.

Future releases of the application could include haptic feedback. The feeling of resistance in the interaction tool when being inside a CT dataset could improve the application further and make pre-operative planning even easier. This is important because, even if the general impression from the test results and the interviews is that it does not take a long time to become accustomed to the immersive environment, the users involved in the study were not the final users. Some surgeons might find it more difficult to adapt to a VR environment than test persons if they do not have the same technical background.

In conclusion, this study has demonstrated that a virtual reality environment is easy to use for new users and even reduces the time needed to perform the positioning. However, a virtual reality environment for pre-operative planning of total hip replacement is not necessarily more accurate than a standard mouse – monitor interface if not highly specialized for the specific application as it has already been done for 2D planning environments.

Acknowledgements

The authors would like to thank Luigi Lena for the illustrations and Mauro Ansaloni for the support during the experiments. This work was partially supported by the MULTIMOD (IST-2000-28377) and MULTISENSE (IST-2001-34121) European projects.

References

1. Dobson HD, Pearl RK, Orsay CP, Rasmussen M, Evenhouse R, Ai Z, Blew G, Dech F, Edison MI, Silverstein JC, Abcarian H. Virtual reality: new method of teaching anorectal and pelvic floor anatomy. Diseases of the Colon and Rectum 2003;46(3):349 – 352.
2. Spicer MA, Apuzzo ML. Virtual reality surgery: neurosurgery and the contemporary landscape. Neurosurgery 2003;52(3):489 – 497; discussion 496 – 487.
3. Hassfeld S, Muhling J. Computer assisted oral and maxillofacial surgery—a review and an assessment of technology. International Journal of Oral and Maxillofacial Surgery 2001;30(1):2 – 13.
4. Handels H, Ehrhardt J, Plotz W, Poppl SJ. Computer-assisted planning and simulation of hip operations using virtual three-dimensional models. Studies in Health and Technology Informatics 1999;68:686 – 689.
5. Schizas CG, Parker B, Leyvraz P-F. A study of pre-operative planning in CLS total hip arthroplasty. Hip International 1996;6:75 – 81.
6. Viceconti M, Lattanzi R, Antonietti B, Paderni S, Olmi R, Sudanese A, Toni A. CT-based surgical planning software improves the accuracy of THR preoperative planning. Medical Engineering & Physics 2003;25(5): 371 – 377.
7. Nishihara S, Sugano N, Nishii T, Tanaka H, Yoshikawa H, Ochi A. Comparison of the fit and fill between the Anatomic Hip femoral component and the VerSys Taper femoral component using virtual implantation on the ORTHODOC workstation. Journal of Orthopaedic Science 2003;8(3):352 – 360.
8. Jaramaz B, DiGioia AMr, Blackwell M, Nikou C. Computer assisted measurement of cup placement in total hip replacement. Clinical Orthopaedics 1998;354:70 – 81.
9. Viceconti M, Lattanzi R, Zannoni C, Cappello A. Effect of the display modality on the spatial accuracy of orthopaedic surgery pre-operative planning applications. Medical Informatics and the Internet in Medicine 2002;27(1):21 – 32.

10. Schroeder W, Martin K, Lorensen B. The visualization toolkit. Upper Saddle River, NJ: Prentice-Hall; 1998.
11. Schroeder W, Martin K, Avila LS, Law C. The visualization toolkit user's guide. Clifton Park, NY: Kitware; 2000.
12. Welch BB. Practical programming in Tcl and Tk. Upper Saddle River, NJ: Prentice-Hall; 1997.
13. Lattanzi R, Viceconti M, Zannoni C, Quadrani P, Toni A. Hip-Op: an innovative software to plan total hip replacement surgery. Medical Informatics and the Internet in Medicine 2002;27(2):71 – 83.