PRE-PRINT OF PAPER IN ACM INTERNATIONAL CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS (CHI), PP 1275-1278, 2009.

Evaluation of a Tool-Mounted Guidance Display for Computer-Assisted Surgery

Kevin Kassil
School of Computing
Queen's University

A. James Stewart
School of Computing
Queen's University

ABSTRACT

We attached a small LCD display and video camera to a surgical drill. The LCD shows the tool position with respect to a planned trajectory, overlaid on video captured by the camera. We performed a user study to determine whether such a tool-mounted guidance display yields faster and more accurate tool placement than the conventional guidance display on a separate computer monitor. Our study showed that the tool-mounted display provides better positional and angular accuracy than the conventional display but that the video camera provides no significant improvement in error.

Author Keywords

surgical guidance interfaces, hand-held displays, tool guidance

ACM Classification Keywords

H.5.2 Information Interfaces and Presentation

INTRODUCTION

In computer-assisted surgery, the patient typically undergoes pre-operative computed tomography or magnetic resonance imaging to build a 3D computer model of the anatomy of interest. A surgical plan is made on the computer model; the plan specifies the positions and orientations of holes to be drilled (in orthopaedics) or probes to be placed (in neurosurgery).

In the operating room, an optical or magnetic device tracks the 3D position and orientation of the surgical tool. The current tool position is shown on a monitor, superimposed on the planned trajectory of the computer model. When it comes time to drill or probe, the surgeon watches the monitor to guide the tool along the planned trajectory. This method of tool guidance has a number of drawbacks. To focus upon the display, the surgeon must turn his or her head away from the surgical field. As a result, the surgeon moves the tool into the patient without continuously watching the tool or the patient. The surgeon must also make a mental transformation between the coordinate system of the image on the monitor and that of his or her own viewpoint.

We built a tool-mounted guidance display consisting of an LCD screen and video camera mounted on a surgical drill, as shown in Figure 1. The LCD screen shows the guidance information (consisting of the computer model of the patient and the planned trajectories) from the real-world viewpoint of the tool. The guidance information can be overlaid onto live video from the camera to approximate a "see-through screen" as shown in Figure 2.

We performed a user study to determine whether the tool-mounted guidance display improves the accuracy and speed with which the drill can be positioned and oriented according to a surgical plan. The study compared several guidance interfaces on the LCD display against a conventional guidance interface on a fixed computer monitor.

RELATED WORK

Many authors have presented related work on displays that are hand-held or instrument-mounted. The novelty of our work is our user study to determine whether a tool-mounted display can, in fact, improve hand-eye coordination for tool manipulation. Weber et al. [5] presented the NaviView system for maxillofacial surgery, consisting of a tracked, free-hand 16 cm LCD screen which displays patient CT data from the viewpoint of the device and uses a CCD camera mounted on the back to superimpose the CT data on the image of the surgical field. Hayashibe et al. [2] described a ceiling-mounted articulated 15 inch LCD monitor which shows a volumetric rendering of a patient's CT scan from a viewpoint above the monitor. They also placed a camera on the back of the display to show the surgical field under the monitor. Other authors have attached the display directly to the tool, as we did. Stetten et al. [3] presented the "sonic flashlight," a modified ultrasound probe incorporating a screen and a half-silvered mirror, which reflects the ultrasound image from the screen so that the operator sees the image superimposed on the patient in the right location. Azar et al. [1] studied performance in a needle placement task, comparing a computer monitor to a head mounted display, and considering 2D and 3D views. They found that avoidance of dangerous areas was better using 2D and the head mounted display. Traub et al. [4] studied performance in a drilling task using a head mounted display, comparing orthographic, augmented reality, and a combination of the two for guidance visualization. They found the combined visualization to be best.

CHI 2009, April 4–9, 2009, Boston, Massachusetts, USA.

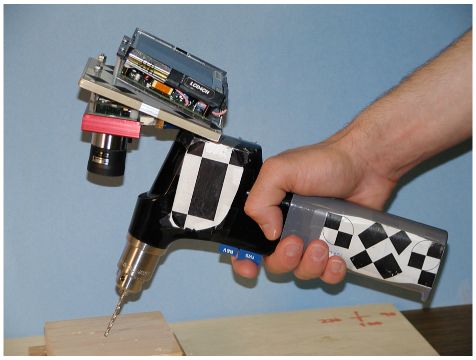


Figure 1. The drill with attached LCD display and camera.

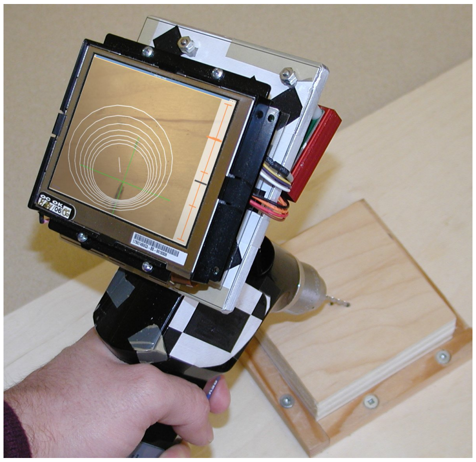


Figure 2. The drill's guidance display (cables removed for clarity).

APPARATUS

The drill was a Smith & Nephew Dyonics 450. It was tracked with a Claron Technologies Micron Tracker 2 H40 stereo optical tracker using a black and white pattern on the drill. The LCD screen was an AcceleVision LCD4CH, attached to the back of the drill and tilted 10 degrees backward of perpendicular from the drill axis to be approximately perpendicular to the subject's line of sight when the drill was held naturally. The camera was a Videre Design DCSG with 640×480 resolution, attached to the back of the LCD and above the drill body, and tilted 15 degrees forward of perpendicular from the drill axis to be aligned as closely as physically possible to the drill axis while still showing the tip of the drill near the middle of the screen. The screen and camera made up 35% of the total tool weight. This additional weight would not be onerous in real use because the tool is held for only a minute or two at a time during an operation.

The drill tip was calibrated in the drill coordinate system by moving the drill around with the tip fixed in place, then fitting a sphere to the tracked origins of the coordinate system. The drill axis was calibrated in a similar manner by moving the drill around with the entire drill bit held rigid. The LCD was registered with a separate tracked probe touched to the four corners of the display.

Four interfaces were used in the experiment:

The **Hand-held Axial (HA)** interface, shown in Figure 3(a), used the drill-mounted LCD display, which showed the target point, the trajectory line, a wireframe representation of the wooden block, and a set of concentric rings around the trajectory line (as used, for example, by BrainLab (BrainLAB AG, Feldkirchen, Germany)). The drill was correctly aligned when the trajectory line appeared as a point at the centre of the crosshairs, and all the rings appeared concentric on the display. A vertical ruled scale was drawn on the right margin of the display to show the depth of the drill tip. The goal depth was highlighted on the scale. The video camera was turned off.

The **Hand-held Axial with Video (HAV)** interface, shown in Figure 3(b), was identical to the Hand-held Axial interface, except that the video camera was turned on and no wireframe representation of the block was drawn. The subject could see the real block with guidance information superimposed.

The **Monitor Ortho (MO)** interface, shown in Figure 3(c), consisted of a computer monitor one meter to the right of the task location, where the subject could easily look at it. The monitor showed three orthographic views from the front, top, and right of the task area, and one axial view. Each view showed the target point, the trajectory line, a wireframe representation of the wooden block, and a set of concentric rings around the trajectory line. The drill-mounted LCD and video camera were turned off.

The **Hand-held Ortho (HO)** interface used the drill-mounted LCD to display the same four views as the Monitor Ortho interface. The video camera and monitor were turned off.

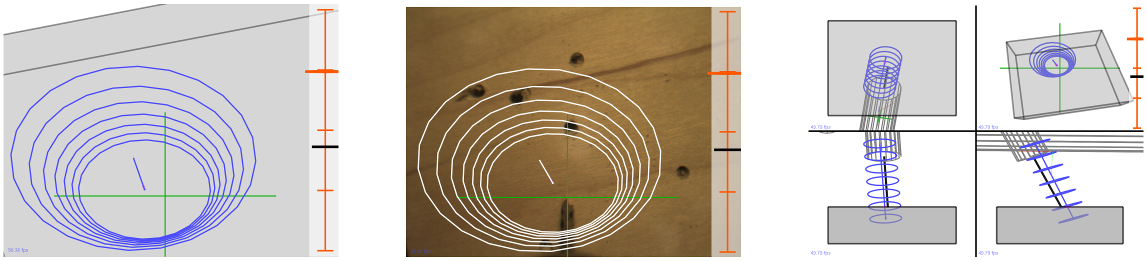
For the drilling, a plywood block of size $9.5 \times 9.5 \times 5$ cm was secured in a fixed holder. The block hid the target point and enforced the constraint that orientation be achieved before the drill was inserted. The fixed holder permitted us to replace the block with an identical, unused block as necessary. The block was registered to the coordinate system of the tracker by touching the four inside corners of the holder with the tracked drill tip. Re-registration with each new trial was not required because the holder remained fixed.

Six targets were set up within the wooden block. Each target consisted of a target point and trajectory line. Targets were selected so that (a) the drill paths did not intersect and (b) the drill paths faced generally toward the subject. The block was replaced whenever the next hole to be drilled had already been drilled in a previous trial on the same block.

Each target point was simply a position inside the block, without any corresponding physical target. The target point was represented in the Micron Tracker's coordinate system; this eliminated registration error as a factor in task performance.

SUBJECTS

Seventeen subjects aged 25 to 55 were tested. Of those, seven had a medical computing background and had expe-



(a) Hand-held Axial interface (HA) (b) Hand-held Axial with Video interface (HAV) (c) Ortho interface (MO and HO)

Figure 3. The interface displays used in the study. The Ortho interface was tested both on a separate monitor and on the drill-mounted LCD display.

rience with 3D spatial displays. Subjects were unpaid and spent about 45 minutes each.

We chose not to use surgeons and surgical residents as subjects because those people are highly trained in the Monitor Ortho interface and would very likely perform much better with that interface than with the other, novel interfaces, biasing our results. However, before a production-quality version of our device could be used in the operating room, we would require another user study with these expert subjects using the new device, and would require a substantially longer training period to overcome any pre-existing expertise with the Monitor Ortho interface.

EXPERIMENT

Subjects were asked to perform 24 trials. Each trial consisted of the subject guiding the drill into the wooden block along a trajectory line to a target point, using one of the four interfaces. The subject declared when he or she was done, at which time the subject's completion time was recorded. The positional and angular errors were recorded at the point of greatest depth so that, even if the subject would drill too deeply, pull back, then declare completion, we would still measure (as would be appropriate in a real surgical task) at the point of deepest drilling.

Subjects were initially trained with at least three trials using each interface. Each subject was then asked to perform 24 drilling tasks: For each of the four interfaces described above, the subject drilled to each of the six targets. The targets did not change between interfaces or subjects.

To minimize learning effects, the order of the six targets was randomized for each interface. Eight blocks of three trials each were made. In each block, a single interface was used. In the first four blocks, each of the four interfaces appeared. In the last four blocks, the four interfaces appeared in the same order as in the first four blocks. For the two (separated) blocks that used a particular interface, the 6 trials were randomized and three trials were placed in each block.

RESULTS

We tested the following hypotheses:

1. **Subjects perform better with the two hand-held axial interfaces (HA and HAV) than with the monitor interface (MO).** We expected the LCD display to reduce hand/eye coordination error because the LCD was, essentially, attached to the hand.
2. **Subjects perform better with the hand-held axial with video (HAV) interface than without the video (HA).** We expected the video to give subjects a better sense of their position on the wooden block.
3. **Subjects perform worse with the hand-held ortho (HO) interface than with the monitor ortho (MO) interface.** We expected that the reduction in size and resolution on the hand-held display would introduce positional and angular errors.

Figure 4 shows the experimental results for accuracy and completion times using each of the four interfaces. Student's paired t-test was used to test our hypotheses, with the following results:

1. Overall, both the HA and HAV interfaces provided significantly better angular accuracy than did the MO interface. But only the HA interface provided significantly better positional accuracy.

   For **positional error**, subjects performed significantly better with the HA interface (mean 2.41 mm) than with the MO interface (mean 2.81 mm, $p = 0.004$). But no conclusion could be drawn about the difference in positional error between the HAV interface (mean 2.54 mm) and the MO interface ($p = 0.27$). For **angular error**, subjects performed significantly better with the HA interface (mean 1.86 degrees) than with the MO interface (mean 2.32 degrees, $p = 0.004$) and significantly better with the HAV interface (mean 1.59 degrees) than with the MO interface ($p = 0.001$). For **completion times**, both the HA interface (mean 32.5 seconds) and the HAV interface (mean 31.1 seconds) had completion times longer than that of the MO interface (mean 28.6 seconds). The difference was significant for the HA interface ($p = 0.02$) but not for the HAV interface ($p = 0.10$).
2. Interestingly, there was no conclusive difference in positional error, angular error, or completion time between the HA and HAV interfaces ($p = 0.30$, 0.26, and 0.44 respectively).
3. As expected, the HO interface performed much more poorly than did the MO interface, with significantly more positional error (mean 3.65 mm, $p = 0.012$), and significantly more angular error (mean 2.92 degrees, $p = 0.004$).

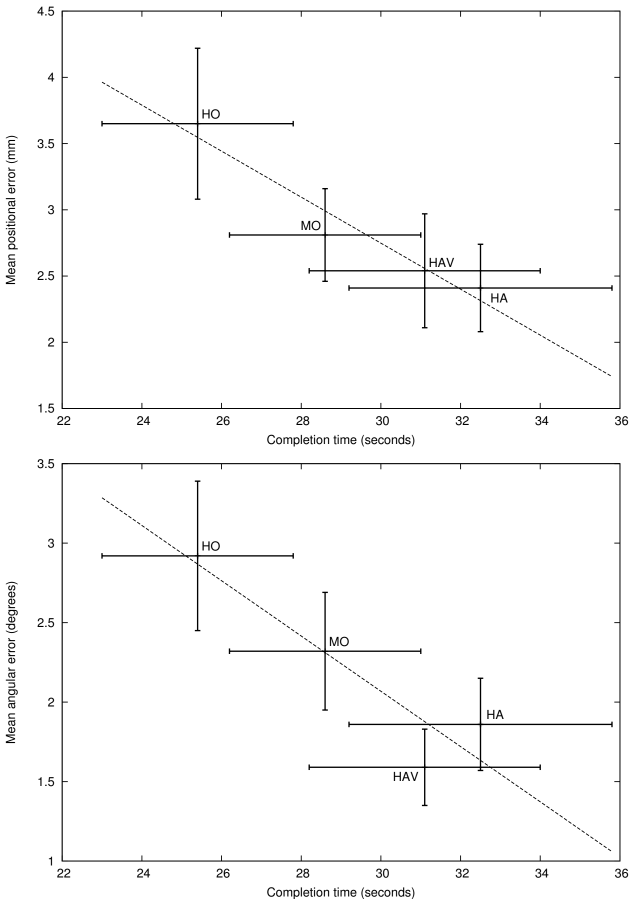


Figure 4. Error decreases as completion time increases. Shown is the 95% confidence interval for each mean (whiskers) and the least-squares linear fit (dashed).

Completion times were significantly *shorter* with the HO interface (mean 25.4 seconds, p = 0.001), suggesting that people could not see how poorly they were doing due to the low resolution. Interestingly, people reported a preference for the HO interface.

We observed a correlation between mean error and mean completion time. For mean positional error, err_{pos}, in mm, mean angular error, err_{ang}, in degrees, and completion time, t, in seconds, the least-squares linear fits were $err_{pos} = -0.17\,t + 7.9$ and $err_{ang} = -0.17\,t + 7.3$ (see Figure 4).

In post-experiment interviews, the subjects made several interesting observations:

- Six subjects expressed a strong preference for the MO interface, five for the poorly performing HO interface, and one subject for each of the HA and HAV interfaces. The remaining four subjects did not express a strong preference. The hand-held HA and HAV interfaces were usually ranked last in preference, although the experiment showed that they had superior performance. It may be that the subjects found the axial view harder to use because it showed misalignment more clearly than did the orthographic views.
- Most subjects said that the video (HAV) was distracting because it had surplus information. They qualified this by saying that, if there had been prominent landmarks on surface, it would be a more effective aid. Some said that the video provided a sense of scale to help judge distance on the surface.
- One subject suggested that the HA view, being axial and large, provided a more sensitive visualization of the error. Thus, subjects may have perceived that their error in the HA and HAV interfaces was larger than it actually was.
- One subject, who had done a lot of drilling into phantom bones as part of other work, said that the drill-mounted display significantly reduced the neck stiffness and fatigue that he had experienced using a Monitor Ortho display to guide drilling into the phantom bones.

CONCLUSION

We conclude that an LCD display mounted on a surgical drill can provide better positional and angular accuracy in a drilling task at the cost of slightly longer completion times.

It appears that a video camera mounted to the drill does not do much to improve accuracy. However, because the hand-held interfaces with and without the camera performed differently with respect to the conventional monitor interface, a further study with a longer training period should be conducted before ruling out the use of a camera. It could be that the camera is useful for setting the initial position of the drill tip and the orientation of the drill but that, after drilling commences, the camera serves no purpose.

REFERENCES

1. F. S. Azar, N. Perrin, A. Khamene, S. Vogt, and F. Sauer. User performance analysis of different image-based navigation systems for needle placement procedures. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 110–121, 2004.

2. M. Hayashibe, N. Suzuki, A. Hattori, Y. Otake, S. Suzuki, and N. Nakata. Data–fusion display system with volume rendering of intraoperatively scanned CT images. In *Medical Image Computing and Computer–Assisted Intervention*, pages 559–566, 2005.

3. G. Stetten, V. Chib, and R. Tamburo. Tomographic reflection to merge ultrasound images with direct vision. In *IEEE Proceedings of the Applied Imagery Pattern Recognition annual workshop*, pages 200–205, 2000.

4. J. Traub, P. Stefan, S. M. Heining, T. Sielhorst, C. Riquarts, E. Euler, and N. Navab. Hybrid navigation interface for orthopedic and trauma surgery. In *Medical Image Computing and Computer Assisted Intervention*, pages 373–380, 2006.

5. S. Weber, M. Klein, A. Hein, T. Krueger, T. C. Lüth, and J. Bier. The navigated image viewer — evaluation in maxillofacial surgery. In *Medical Image Computing and Computer–Assisted Intervention*, pages 762–769, 2003.