massachusetts institute of technology — artificial intelligence laboratory

# Exploring Vision-Based Interfaces: How to Use Your Head in Dual Pointing Tasks

Trevor Darrell, Neal Checka, Alice Oh and Louis-Philippe Morency

AI Memo 2002-001 January 2002

© 2002 massachusetts institute of technology, cambridge, ma 02139 usa — www.ai.mit.edu

## ABSTRACT

The utility of vision-based face tracking for dual pointing tasks is evaluated. We first describe a 3-D face tracking technique based on real-time parametric motion-stereo, which is non-invasive, robust, and self-initialized. The tracker provides a real-time estimate of a "frontal face ray" whose intersection with the display surface plane is used as a second stream of input for scrolling or pointing, in parallel with hand input. We evaluated the performance of combined head/hand input on a box selection and coloring task: users selected boxes with one pointer and colors with a second pointer, or performed both tasks with a single pointer. We found that performance with head and one hand was intermediate between single hand performance and dual hand performance. Our results are consistent with previously reported dual hand conflict in symmetric pointing tasks, and suggest that a head-based input stream should be used for asymmetric control.

### Keywords

Head tracking, vision-based interface, perceptual user interface, dual pointing tasks

## INTRODUCTION

New "perceptive" user interfaces have been developed that allow direct control of pointers or other user interface degrees of freedom using body gestures [29, 11, 14, 35]. In contrast to previously available commercial technologies for users with disabilities (e.g., [13, 27]), these systems are designed for uncalibrated, non-contact interaction with naive users. In this paper, we present a state-of-the-art perceptive interface for head-based pointing and evaluate its utility on a dual selection task.

Gaze or head pose is a potentially powerful and intuitive pointing cue if it can be obtained accurately and unobtrusively. Eye movement is an important topic of study in evaluating user interface performance and potentially as an input cue [21]. However, the dynamics of involuntary human eye movements limit its accuracy for independent fine pointing control, although it can provide a coarse reference frame [41]. In this paper, we focus on head pose tracking, which has been a topic of active interest in the computer vision community [12] and appears to be less affected by involuntary control movements than eye-gaze. Several authors have built prototypes which incorporate head pose tracking into user interface controls [2,22,24,26]. These systems generally propose that head pose be used to control a second pointer or scrolling parameter in drawing, selection, or navigation applications.

Many human-computer interface tasks could benefit from the simultaneous control of multiple input devices. In our everyday experience, we interact with other people and devices with multiple hands and information streams (speech, gesture, gaze), but we can generally only type and click on personal computers. Ideally, head pose input can provide a second pointing or focus of attention channel and potentially make tasks faster and more intuitive.

There is substantial literature exploring the possible benefits of interaction with dual input channels for human-computer interface. Buxton [9] demonstrated that the use of both hands improved performance for common CAD tasks. More recent work has shown that dual hand-based input many not independently increase performance for all task types. Increased performance was found on asymmetric tasks [15] where the non-dominant hand is used to control a general reference frame and handle coarse tasks and the dominant hand handles precise tasks [18]. A comparative study of dual hand control for compound drawing and selection tasks showed that an asymmetric interface style with overlaid transparent menus [3] increased performance, but that for symmetric pointing tasks "two hands can be worse than one" [23].

We investigated whether head-pose could effectively be used as a parallel input stream. Does a head-pose input channel enhance or conflict with a dominant hand input channel? How does it compare with a second hand in symmetric pointing tasks? We investigated these questions on a task that required dual selection of objects on a computer screen. We first review the details of our head tracking system, and then describe the experiment performed. We conclude with a discussion of how our observations and results should influence the design of future perceptive user interfaces.

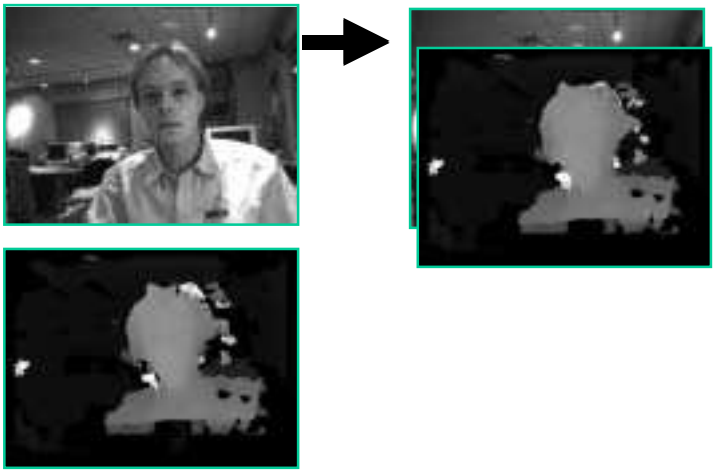


**Figure 1: Intensity and range images of a user taken at two time instants. Top row shows intensity images, and bottom row depth images obtained from a real-time stereo range camera [37]. Rigid-body motion techniques can recover the 3-D motion of a user's face accurately and unambiguously when both intensity and depth data are available [16].**

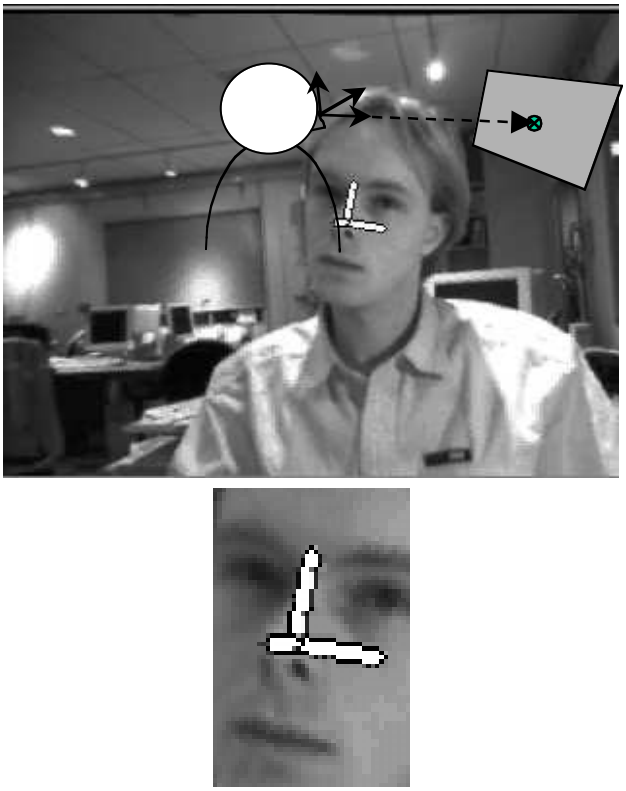


**Figure 2: Example of pointer control using a head tracking system. The 3-D position and orientation of the user's head are tracked in real-time. In the top image, a set of axes is overlaid on the user's face to show the estimated pose. A close-up of the face is shown in the lower-left image. The intersection of the frontal face ray with a display screen is used to control a pointer, as shown in the lower-right figure.**

## HEAD-POSE TRACKING USING STEREO VISION

Many techniques have been proposed for tracking a user's head based on passive visual observation. To be useful for pointing or focus of attention tasks, tracking performance must be accurate enough to localize a desired region, robust enough to ignore illumination and scene variation, and fast enough to serve as an interactive control. Examples of 2-D approaches to face tracking include color-based [28,40,4], template-based [24] and eigenface-based [36] techniques. These techniques are fast but capture only gross head motion and may require initialization or calibration.

Techniques using 3-D models have greater potential for accurate tracking but require knowledge of the shape of the face. Early work presumed simple shape models (e.g., planar [5], cylindrical [25], or ellipsoidal [2]). More recently, tracking has been performed with a 3-D face texture mesh [32] or 3-D face feature mesh [39]. Toyama [33] developed a robust head tracking system that used a hierarchy of tracking approaches that ranged from 3-D rigid features to coarse color blobs and could automatically reset and initialize more accurate layers when appropriate. Applied to cursor control, accuracy on the order of 1cm was reported. This is comparable to the eye-gaze tracking uncertainty reported in [41].

Very accurate shape models are possible using the "active appearance model" methodology [10,22], especially when applied to 3-D head data as in [6]. However, tracking 3-D active appearance models to monocular intensity images is currently a time-consuming process that is not well suited for interactive systems.

Our approach has been to track using 3-D observations and direct rigid motion models. This has the advantage of being accurate and fast, as well as robust to substantial illumination variation. We observe a user's face using a real-time optical stereo system [30,37] and compute the rigid motion which best accounts for the intensity and shape variation seen from frame to frame. The core of our approach is a system for "stereo parametric motion", using a combination of an instantaneous rigid motion model and traditional optic flow equations extended to include a depth term [16]. Figure 1 shows an example of the intensity and depth images used by our system for tracking.

Classically, optic flow equations relate the image gradient to local motion under the assumption that image variation is due to scene motion [20], and these local motion estimates can be aggregated into a single global motion estimate (rigid [19] or affine [5]). When tracking faces with intensity alone, it can be difficult to discriminate small translations and rotations, and the systems of equations that

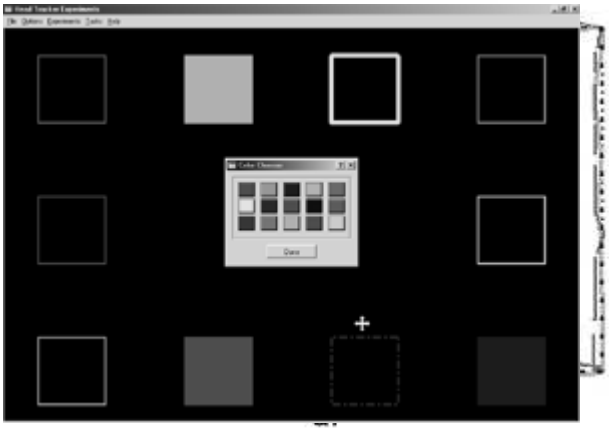
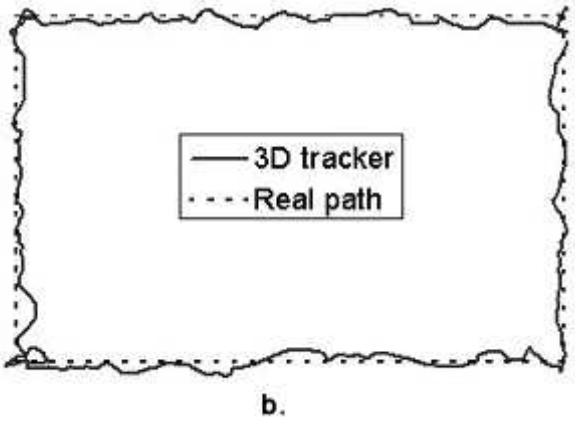


**Figure 4: A screen show of our box-coloring task. Users selected a box from the outer ring, then a color from the dialog box.**



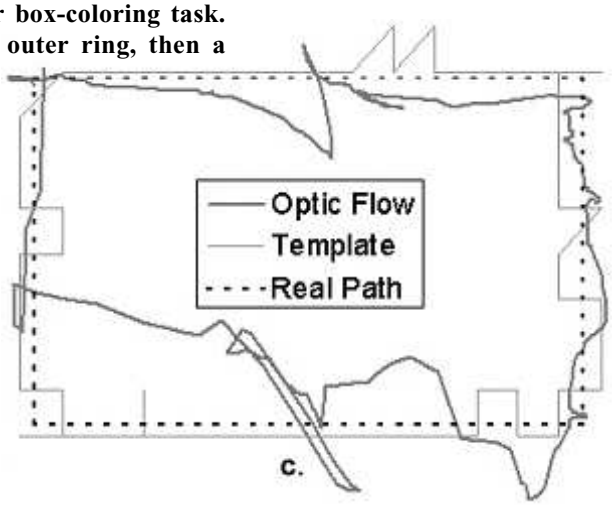




**Figure 3: Examples of trajectories where the user was attempting to follow a rectangular shape. A) shows trackball and mouse performance for comparison. B) shows performance using 3-D head tracker described in this paper. C) shows performance with 2-D face tracking techniques based on optic flow or template tracking.**

have to be solved are therefore often numerically ill-posed.[1] However, with stereo range data, small rotations and translations become easy to distinguish. For mathematical details of our system see [16].

However robust, rigid motion tracking alone is insufficient to build a head tracker; the system must be initialized and reset if tracking is lost. Rather than implement the "hard reset" that many systems use, we have developed a "soft refinement" approach to error stabilization. Our algorithm automatically tracks pose change using multiple base frames determined at run time. We have shown how this system can run for very long sequences without needing to be re-initialized [31]. To find the origin of the face coordinate system, e.g., the front of the face, we applied the face detector described in [38]. Stereo data was obtained with the real-time stereo camera system developed at SRI Labs [37].

[1] If an accurate segmentation of the object boundary is available, then this ambiguity is not present. However, this is been notoriously difficulty to achieve in practice.

In contrast to previous head tracking systems for user interface, our system is robust to strong illumination changes, automatically initializes without user intervention, and can reinitialize automatically if tracking is lost (which is rare.) Related, but independent work on stereo face tracking [17] has applied a face feature mesh model to tracking with 3-D features. This system is also fast and robust but requires manual initialization of some of the tracking features or prior experience with the people being tracked.

To guide a pointer on a screen, we adopt the paradigm of a "nose pointer" (Figure 2) and compute the 3-D intersection of the frontal face ray with the display screen [33]. We presume the position and extent of the display surface relative to the stereo camera location is known a priori.

We ran experiments to test the raw accuracy of our system before starting the user studies described below. Figure 3b shows examples where the user used his or her face to guide a pointer along a fixed trajectory. Figure 3c shows a comparison with several other face tracking methods (template[2,24], color[8], optic flow[26]) on the same task.

Figure 3a shows the trackball and mouse accuracy on the task.

Based on the tracing experiments in Figure 3, our system has tracking accuracy equivalent to 5.3 pixels standard deviation on a 72 dpi screen with a user 50cm from the display. This compares favorably with the accuracy in [33], which reported a pointer position uncertainty region of approximately 1cm. Table 1 summarizes the tracking accuracy for our implementation of several systems: normalized correlation template tracking [2,24], color tracking[8], 2D optic flow [26], as well as our 3D head tracker, reports that eye gaze accuracy is one-half to one degree of arc [41] (without mouse for fine control), as well as empirical results with trackball and mouse (Figure 3). When run on a 1.5GHz Pentium IV workstation, our system runs at approximately 15Hz with 60ms lag.

**Table 1: Tracking accuracy of several systems**

| Method | Std. dev. (pixels) |
|---|---|
| Correlation head tracker [2] | 17.6 |
| 2D Optical Flow head tracker [26] | 22.9 |
| Hybrid [33] | 30 |
| 3D head tracker | 5.3 |
| Eye gaze | 18-36 |
| Trackball | 3.7 |
| Mouse | 1.9 |

## IS HEAD BETTER THAN HAND AS A SECOND INPUT STREAM?

Using the head tracker described above we sought to test whether passive head input was a useful input stream compared to a second hand input. Since it is still an open question how to sense "clicks" from a face (some have suggested dwell time or eye blinks), we focus on the use of face input as a second stream. We use a task based on dual selection since existing tools for scrolling (wheelmouse, scrolling-at-window-edge) seem highly effective.

### Task

Our experimental paradigm was a box-coloring task designed to represent an object selection and attribute setting interaction present in many complex graphical interfaces. Users were presented with two sets of boxes, one indicating regions to be filled and other indicating possible colors. The boxes to be filled were larger than the color specification boxes.

Users first selected boxes and then selected a color with which to fill it. In a dual pointer regime, clicking on a color with the first pointer set the color of a box selected with the second pointer. In a single pointer regime, users first clicked on a box to select it and then clicked on a color to perform the color fill operation. The boxes were arranged in a rectangular grid on the outside of the display screen, and the color section window was placed at the center of the screen (see Figure 4). (This arrangement was chosen to minimize the overall length of pointer travel in both single pointer and dual-pointer conditions.)

We experimented with two ordering conditions to evaluate the effect of pointer trajectory length on performance:

A. Boxes were selected in a user-specified order and had to be filled with a particular color (as shown by the color of the border of the box).

B. Boxes had to be selected in a random order specified by the computer (as shown by highlighting the next box's border) and had to be filled with a particular color (as shown by the color of the border of the box).

**Figure 5: Our experimental setup. Note dual mice, and stereo camera positioned above display screen.**

These two conditions imply different average trajectory lengths when two pointers are used since in (A) the user can move one around the ring of boxes in a systematic fashion with short movements while in (B), the user has to move on average approximately half a screen away to select the next box.

We tested each of the above ordering conditions in each of three input device conditions:

1. Single mouse only - users selected both box and color using one mouse in their dominant hand (if any).

2. Dual mouse input - users used one mouse for selecting the box and the second mouse for selecting the color. Either mouse could be used for either task.

3. Head-pose input with second mouse input - users selected the box to be colored using a head-driven pointer and the color using a mouse pointer in their dominant hand (if any).

To the extent that the task is separable, we would expect condition 1 to be slower than condition 2 and 3. If it is

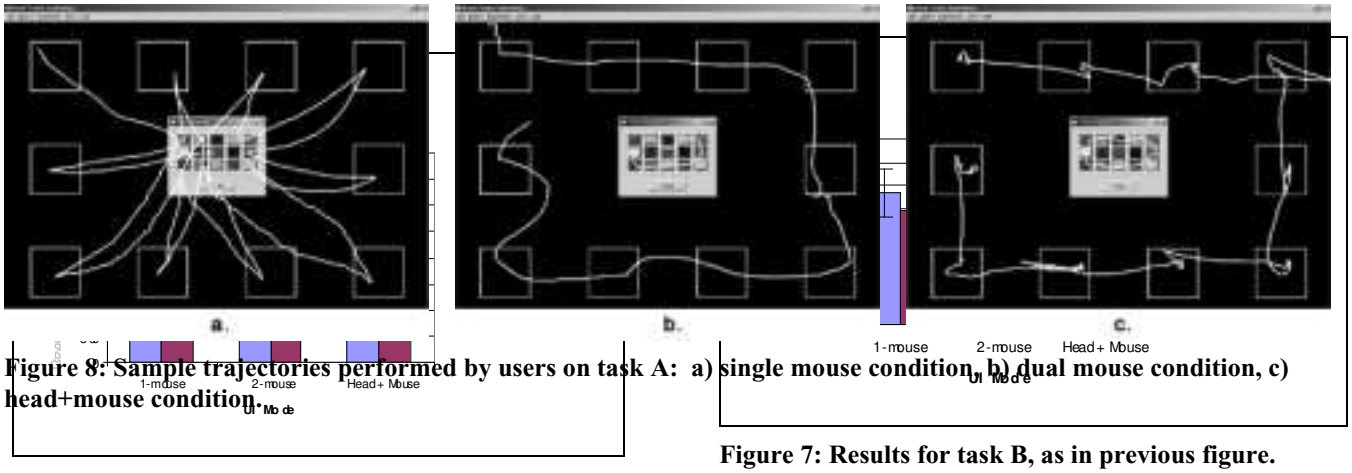



**Figure 8: Sample trajectories performed by users on task A: a) single mouse condition, b) dual mouse condition, c) head+mouse condition.**

**Figure 6: Results on task A. Execution time in milliseconds is shown for single mouse, dual mouse, and head+mouse conditions, for novice and advanced users. Error bars show one standard deviation.**

**Figure 7: Results for task B, as in previous figure.**

separable but there is conflict at a hand motor control level, we could expect condition 3 to be faster than 2. We would expect the converse result if there is bilateral motor enhancement. If conditions 2 and 3 were equally slow, it would indicate an attentional or cognitive conflict.

### Apparatus

Users sat in front of a 17" LCD display screen connected to a conventional personal computer running custom software to implement the experiment described above. (Figure 5) A mouse was connected to the first computer system via standard interface and controlled the first pointer on the application described above. A second computer system was used to control the second pointer, via custom network software. The second computer either relayed a pointer signal from a conventional mouse or sent the signal from face tracking software running on that computer. Latency due to network transduction of the second mouse was designed to be less than 10ms. Mouse speed was set such that the user could move the head pointer at approximately the same speed as the mouse pointer.

### Procedure

Fifteen subjects successfully participated in the experiment. Subjects were recruited from our research lab and were graduate students in computer science who are expert users of personal computers. None of the subjects had any prior experience with a head-tracking system. Nine of the subjects were in the "novice" group and the other six were in the "advanced" group. We divided the subjects this way to study how naive subjects perform using the input modes, and how "advanced" subjects, or subjects with more practice, compare with the naive subjects.

The experiment included two stages: a training stage and testing stage. During the training stage, subjects were told to practice selecting boxes around the screen using the three input device conditions until they felt comfortable. This stage lasted approximately 5 minutes. The "advanced" group had additional practice of 10 - 15 minutes using the actual tasks and the input modes.

After the training and practice stages, the subjects completed the tasks using the three input device conditions in counterbalanced order. Subjects in the "novice" group completed each of the tasks once, and subjects in the "advanced" group completed each of the tasks 5 or more times. The execution times for coloring each of the boxes was recorded as well as the total times for each task. After the experiment, we asked the subjects for comments

### Results

Figure 6 shows the average time required to complete task A, for each of the input conditions and for different levels of user experience. In task A, the user chooses the order of boxes to be colored. The lighter bars indicate performance for naive users, who had little experience with the task, apparatus, or the head tracking system. Darker bars indicate performance after approximately one hour's experience with the system.

Novice users were faster using the head and one hand condition than using two hands, but were fastest with a single mouse. Experienced users, in contrast, were reversed: they were fastest using two mice, second fastest on head and one mouse, and slowest on a single mouse.

A similar trend was seen in task B, where the order of boxes to be colored is predetermined by the computer. Results for task B are shown in Figure 7.

Surprisingly, the results for task A and B are qualitatively similar, despite different trajectory styles performed by users for each task.

Figure 8a shows a typical pointer trajectory for a user performing task A using a single mouse. The user first guides the pointer to select a box, then the specified color, then the next box of the user's choosing, and so on. Figure

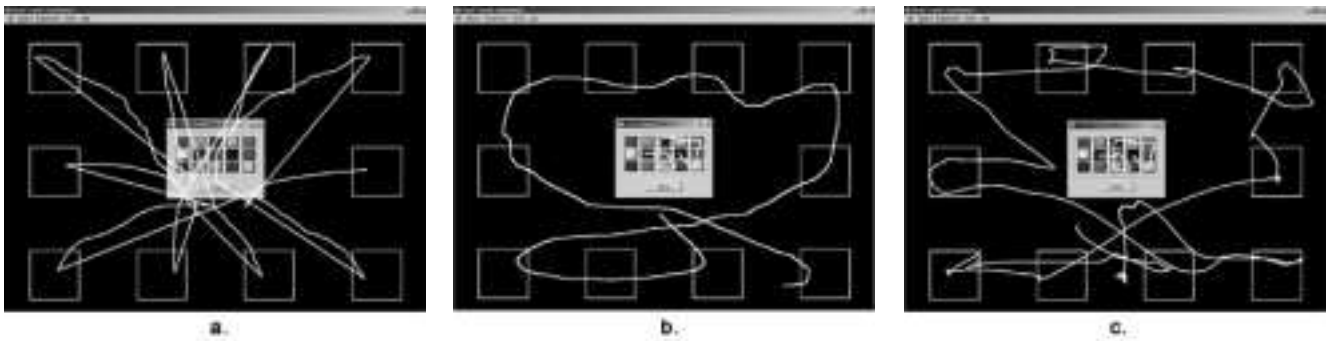



**Figure 9:** Sample trajectories performed by users on task A: a) single mouse condition, b) dual mouse condition, c) head+mouse condition.

8b shows the trajectories on this task for two mice, and Figures 8c show the trajectory using the head and one hand.

Since task A allows the user to select the order of boxes, a compact trajectory is possible for each mouse in the dual pointer conditions. However, task B fixes the order of boxes, and therefore requires the box selection pointer to traverse on average half the screen width per box. Sample trajectories for task B are shown in Figure 9

The trajectory lengths (in mm) corresponding to Figure 8 and 9 are shown in Table 2. As expected, in task A trajectory lengths for dual pointer input were lower than for single pointer input. In task B the lengths are more similar. Intuitively, the box selection pointer in task B can be used to also select the color without adding significantly to the overall trajectory length, since it will be passing by the color selection window often during a typical run.

The comments from the "novice" group on the dual mouse condition included "2 mice is hard, confusing" and "two mice was most difficult because I almost never use my left hand as a pointing or gesturing hand". One positive remark about the head-pose input mode was that "head tracker [is especially easy to use] for coarse movements, mouse for fine movements". A comment that "head tracker was not well adjusted for my usual head movement" suggests that we may need to design more natural mapping between head movements and input actions. Many subjects commented that the single mouse condition was the easiest because they were very used to it.

## DISCUSSION

The results presented above show that head tracking can be approximately as accurate as a second mouse in dual selection tasks. Interestingly, they show that naïve users perform better on head and hand tracking than on dual hand tracking, even when the hand input device is one that the users are extremely familiar with (a mouse). Our results support the hypothesis that naïve users have some motor coordination conflict that makes symmetric dual manipulation difficult [15,23], and that this conflict is lessened with the use of the head and one hand.

However, naïve users performed best on the device they use every day, a single mouse. An additional input stream did not help performance, lending support to the hypothesis that in addition to the motor conflict, naïve users have an attentional or cognitive conflict in driving two pointers.

Expert users were able to partially overcome both the attentional and motor conflicts, since their performance in the dual pointer conditions was faster than in the single pointer condition. Familiarity may still have an effect, as the two mouse condition was faster than the head and mouse condition.

Counter to our intuition, trajectory length did not seem to play a large role in determining task completion time. The relative trajectory lengths between single and dual pointer conditions in task A vary by a factor of 2 (Table 1), yet the completion times at most vary by 20%! We expected to see a dramatic improvement with dual pointers in task A relative to task B, but both tasks showed qualitatively similar results (Figures 6 and 7)

The improvement with dual pointers (hand or face) is modest. Our result casts doubt on the utility of a face-driven pointer to provide position context for actions selected by a primary pointer. However, the face tracker was nearly as accurate as a single mouse, suggesting that if the problem of “clicking” with one’s face or head can be overcome it may be useful in hands-free environments, such as interactive walls or kiosks.

Head tracking may be most appropriate for selection on a larger display surface, such as in an interactive room or wall display. We believe that integration of head tracking with natural language may provide the most intuitive and useful pointing system, allowing head gaze to provide linguistic reference to objects and people in the real (or virtual) world [7].

The technology for perceptive user interfaces based on head-tracking appears mature enough for evaluation and possible integration with other user interface modalities. Our experiments are the first in a series to explore how direct sensing can best be applied to interaction tasks. While head tracking is the most mature perceptive technology in engineering terms, there is still much work to be done to find the appropriate interaction paradigm.

**Table 2: Average trajectory lengths**

| Mode | Task A | Task B |
|---|---|---|
| Single mouse | 1902 | 2030 |
| Dual mouse | 854 | 1864 |
| 3D-Tracker | 1212 | 1773 |

We also expect that as other perceptive technologies mature, direct interfaces will move beyond head tracking to include full body gestures, interaction across large spaces, and interaction between multiple people. The results in this paper show that symmetric pointing tasks may be difficult no matter which body gesture is used, due to the attentional or cognitive load presented by such a tasks.

## CONCLUSIONS

Head tracking based on rigid motion stereo can be accurate, robust, automatically initialized, and fast. A head-driven pointer based on this system is approximately half as accurate as a trackball or mouse in terms of position accuracy. On a dual selection task naïve users perform slightly better using head-pose as a second input stream than with dual hand inputs. Attentional or cognitive conflict is still an issue with symmetric dual pointing tasks, even with a second head-based input stream. Future systems for head-based input should focus on asymmetric dual control tasks, as suggested in the dual hand manipulation literature.

## REFERENCES


1. Basu. S., Essa, I., and Pentland, A., "Motion Regularization for Model-based Head Tracking". Proc. ICPR, 1996, Vienna, Austria.
2. Berard, F. (1999) The Perceptual Window: Head Motion as a new Input Stream, in IFIP conference on Human-Computer Interaction (INTERACT)
3. Bier, E. A., Stone, M. C., Pier, K, Buxton, W., & DeRose, T. D. (1993). Toolglass and magic lenses: The see-through interface. Proceedings of SIGGRAPH '93. (pp. 73-80). New York: ACM
4. Birchfield, S., Elliptical Head Tracking Using Intensity Gradients and Color Histograms, Proc. CVPR 1998, pp. 232-237, IEEE Press, 1998.
5. Black, M., and Yacoob, Y., Tracking and recognizing rigid and non-rigid facial motions using local parametric models of image motion. In Proc. 5th ICCV, pages 374--381. IEEE, 1995
6. Blanz, V., and Vetter, T., A Morphable Model for the Synthesis of 3D Faces. SIGGRAPH'99 Conference Proceedings (pp 187-194), 1999.
7. Bolt, R. A. (1980) "Put-That-There: Voice and Gesture at the Graphics Interface." Proceedings of SIGGRAPH '80. ACM Press, New York.
8. Bradsky, G., "Computer Vision Face Tracking For Use in a Perceptual User Interface", Intel Technology Journal, Spring 1998.
9. Buxton, W., and Myers, B., "A study in two-handed input", Proceedings CHI '86, pp. 321-326, 1996.
10. Cootes, T., Edwards, G., and Taylor, C.,. Active appearance models. In Burkhardt and Neumann, editors, Computer Vision -- ECCV'98 Vol. II, Freiburg, Germany, 1998. Springer, Lecture Notes in Computer Science 1407.
11. Crowley, J., Coutaz, J., Berard, F., Things That See. Communications of the ACM, March 2000, p. 54-64
12. Crowley, J., ed., Proceedings of the Fourth IEEE Conferece on Automatic Face and Gesture Recognition, Grenoble, France, 2000.
13. EyeTech Digital Systems, http://www.eyetechds.com/
14. Flickner, M., IBM BlueEyes Project, http://www.almaden.ibm.com/cs/blueeyes/index.html
15. Guiard, Y. Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. Journal of Motor Behavior, 19(4), 486-517, 1987.
16. Harville, M,. Rahimi, A., Darrell, T., Gordon, G., Woodfill, J., 3D Pose Tracking with Linear Depth

and Brightness Constraints, Proceedings of ICCV99, Bombay, pp. 206-213, 1999.

17. Heinzmann, J., Zelinsky, A., Robust Real-Time Face Tracking and Gesture Recognition, Proc. IJCAI97(1525-1530), 1997.
18. Hinckley, K., Pausch, R., Proffitt, D., Patten, J., Kassell, N., "Cooperative Bimanual Action," ACM CHI'97
19. Horn, B., and Weldon, E., "Direct Methods for Recovering Motion," IJCV 2:1, 51--76, 1988.
20. Horn, B.. Computer Vision. MIT Press, Cambridge, Mass., 1986.
21. Jacob, Robert J.K. What you look at is what you get: Eye movement-based interaction techniques. Proceedings of CHI '90 (1990), 11-18.
22. Jones, M., and Poggio, T., Multidimensional Morphable Models, Proc. ICCV98, pp. 683-688, 1998.
23. Kabbash, P, Buxton, W., Sellen, A., "Two-handed input in a compund task," ACM CHI '94, pp. 417-423.
24. Kjeldsen, R., "Head Gestures for Computer Control", Proc. Second International Workshop on Recognition, Analysis and Tracking of Faces and Gestures in Real-time Systems, ICCV 2001, IEEE Press pp. 62-67
25. La Cascia, M., Sclaroff, S., and Athitsos, V., Fast, Reliable Head Tracking under Varying Illumination, IEEE Trans. PAMI, 22(4), April, 2000
26. MouseVision Inc., www.mousevision.com
27. Naturalpoint Inc., http://www.naturalpoint.com/prod/product.htm
28. Oliver, N. Pentland, A.P. and Bérard, F. (1997) LAFTER: Lips and Face Real Time Tracker, in IEEE Conference on Computer Vision and Pattern Recognition, Puerto Rico.
29. Pentland, A., "Perceptual Intelligence", Communications of the ACM, March 2000, pp. 35-44.
30. Point Grey Systems, www.ptgrey.com
31. Rahimi A., Morency, L., Darrell, T., Reducing Drift in Parametric Motion Tracking, Proc. ICCV 2001
32. Schold, A., Haro, A., Essa, I., Head Tracking Using a Textured Polygonal Model. Proceedings Workshop on Perceptual User Interfaces, San Francisco, November 1998.
33. Toyama, K., "'Look, Ma -- No Hands!' Hands-Free Cursor Control with Real-Time 3D Face Tracking", In Proc. Workshop on Perceptual User Interfaces (PUI'98), November 1998.
34. Turk, M, ed., Workshop on Perceptual User Interfaces (PUI'98), November 1998.
35. Turk, M., and G. Robertson., "Perceptual User Interfaces,", Communications of the ACM, March 2000, p. 32-34.
36. Turk, M., and Pentland. Eigenfaces for recognition. Journal of Cognitive Neuroscience, 3 (1):71--86, 1991.
37. Videre Design, www.videredesign.com
38. Viola, P., and Jones. M, Rapid Object Detection using a Boosted Cascade of Simple Features, To appear Proc Computer Vision and Pattern Recognition, 2001
39. Laurenz Wiskott, Jean-Marc Fellous, Norbert Krüger, and Christoph von der Malsburg, Face Recognition by Elastic Bunch Graph Matching, IEEE Trans. PAMI, Vol. 19, No. 7, July 1997. See also www.eyematic.com/tech_sensing.html
40. Wren, C., Azarbayejani, A., Darrell, T., and Pentland, P., "Pfinder: Real-Time Tracking of the Human Body ", IEEE Trans. PAMI, July 1997.
41. Zhai, S., Morimoto, C., and Ihde, S., Manual and gaze input cascaded (MAGIC) pointing, Proc. ACM CHI, pp. 246-253, 1999.