Hand Pointing Accuracy for Vision-Based Interactive Systems

Kelvin Cheng[1,2], Masahiro Takatsuka[2]

[1] HxI Initiative, NICTA, Australian Technology Park,
13 Garden St, Eveleigh, NSW 2015, Australia
[2] ViSLAB, School of IT, The University of Sydney, NSW 2006 Australia
kelvin.cheng@nicta.com.au, masa@vislab.usyd.edu.au

Abstract. Vision-based hand pointing interactive systems always assume implicitly that users' physical pointing accuracy is perfect. However, this may not be the case. We investigated the accuracy provided by users in three pointing strategies. Result showed that pointing inaccuracy can be as high as 239mm at 3 metres away and suggest that the line-up method provides the best accuracy overall.

Keywords: Hand Pointing, Pointing Accuracy, Computer Vision.

1 Introduction

In most vision-based interactive systems the accuracy of the estimated pointing direction is an essential element to their success. The focus is usually on finding new ways of improving the detection, estimation and tracking the pointing hand or finger, and deduces the pointing direction [1, 2, 4]. However, these systems do not account for inaccuracies made by the user and assumed implicitly that the user is always pointing to the desired location. In this paper, we report on a study that was conducted to investigate the inherent accuracy provided by three common pointing strategies, which were widely used in previous interactive systems:

1) Forearm – only the forearm is used to point, with limited upper arm movement.
2) Straight-arm – the whole arm is used to point, without bending.
3) Line-up – the fingertip is positioned collinearly between the eye and the target. Several names have been used in the literatures for this method such as "eye-fingertip"[3] or "head-hand line"[2], we call this the "line-up" method.

Despite advances in computer vision, there is still no consistent method to segment and extract the pointing direction of the arm. To minimize errors introduced by computer vision systems, we made use of the laser pointing. A laser pointer can be attached to the arm and gives us a direct and simple resultant point, making it possible to quantitatively compare the different targeting strategies. However, it would be difficult to fix a laser pointer at a specific position on the arm. Consequently subjects were asked to hold the laser pointer in their hand in a way that best represent the extension of their arm direction, in a consistent manner across all three methods (Figure 1). Although physically pointing with hand or arm compared to holding the laser in the palm are slightly different, we believe that this difference would be

consistent enough so that it would still be comparable relatively between the strategies. Therefore until a more suitable method is found, we felt that it was appropriate in this case to use the laser pointer. In addition, users were not provided feedback from the laser pointer to adjust their accuracy.

The effect of distance on the accuracy of pointing – whether pointing deteriorates as user moves away from the target – was also investigated in this experiment.

2 Experiment

Fifteen volunteers (3 female, 12 male) participated in this study. Their age ranged from 20 to 31 (average of 24.9). One is left handed while 8 are right eye dominant. A 5x5mm black target was constructed on a wall at a height of 1.5 metres. The laser pointer used was a pen (1x14.7cm) with a red laser pointer fitted at the back end. In order to detect the target and the red laser dot, we used a webcam (Logitech Quickcam Pro 4000) positioned 2 metres away from the target, together with a custom image processing software (Visual C++ and OpenCV). The lighting in the room was dimmed to allow improved detection of the laser dot.

The study was a within-subject study where each subject performed pointing tasks with all three pointing styles from three distances: 1, 2 and 3 metres from the target. Three blocks of trails were completed for each of the 9 combinations and the mean position for each combination was determined. The order of pointing styles and distances were counter-balanced. Without turning on the laser, subjects were asked to aim as accurately as possible, and hold the laser pointer in their dominant hand, in a way that best represents the direction of their pointing arm (for straight arm and forearm pointing, illustrated in Fig 1a). For the line-up method, users were asked to place the laser pointer so that both ends of the laser pointer are collinear with the target and the eye (Fig 1b). Subjects were free to use one or both eyes for this method.
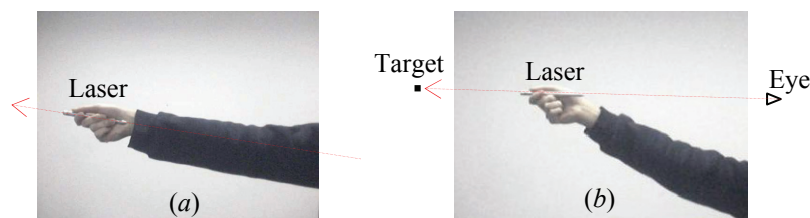

Fig. 1. (*a*) A laser pointer was used to represent the arm's direction (*b*) The laser pointer used with the line-up method, where the target, two ends of the laser pointer and eye are collinear.

To prevent subjects receiving feedback from the laser dot, the laser was only turned on after they have taken aim. Subjects were asked to keep the orientation of the pen consistent throughout the entire experiment, in order to minimize unwanted deviation from the button press. A snapshot was then taken by the webcam. Accuracy was measured in terms of the distance between the target and the laser dot produced on the wall. The positions of the laser dots were determined off-line. In summary, the experimental design was: 15 subjects x 3 pointing techniques x 3 distances x 3 blocks = 405 pointing trials.

3 Results and Discussion

Figure 2 illustrates the mean distance from target for each pointing method at the three distances, and their interactions, for all trials.
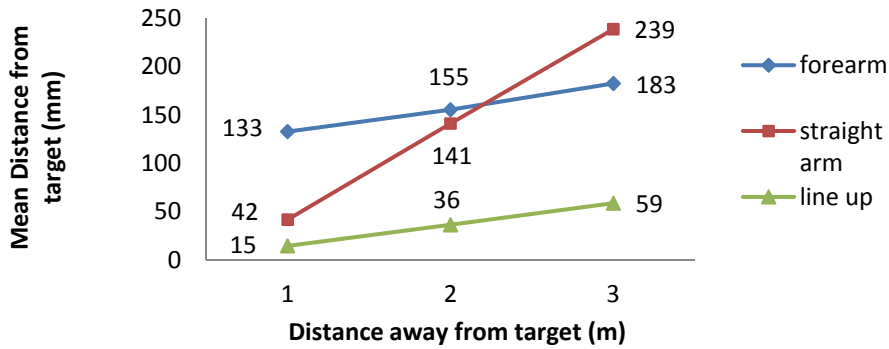

Fig. 2. Mean distance between target and the laser dot position.

A two-way analysis of variance (ANOVA) with repeated measures reveals a significant main effect for pointing technique on accuracy ($F[2,28]=37.97$, $p<0.001$), as well as for distance on accuracy ($F[2,28]=47.20$, $p<0.001$). We also observed a significant interaction between technique and distance ($F[4,56]=9.879$, $p<0.001$).

Multiple pairwise means comparisons were tested within each pointing technique. The p-values resulted with each comparison between the distances are shown in table 1 with Bonferroni correction. Trend analyses were also performed on each technique. Significance in the linear component for the particular technique signifies a linear increase in accuracy with increasing distance.

Table 1. The significance of multiple pairwise means comparisons within each pointing technique.

Technique	1m vs 2m	2m vs 3m	1m vs 3m	Linear component
Forearm	1.000	0.378	0.055	0.018*
Straight arm	<0.001*	0.002*	<0.001*	<0.001*
Line up	0.003*	0.116	0.014*	0.005*

*Denotes significance at the 0.05 level

Multiple pairwise means comparisons were also performed within each distance to investigate possible differences between each technique, and the p-values are shown in the table 2.

Table 2. The significance of multiple pairwise means comparisons within each distance.

Distance	Forearm vs straight arm	Straight arm vs line up	Forearm vs line up
1m	0.001*	<0.001*	<0.001*
2m	1.000	<0.001*	<0.001*
3m	0.182	<0.001*	<0.001*

*Denotes significance at the 0.05 level

Results suggest that the line-up method is the most accurate across all distances. A linear increase throughout the 3 distances at a rate of 14.7mm per meter can also be observed. The forearm pointing technique is consistently less accurate than the line-up method. Even though the linear increase for this technique (16.7mm per metre) is similar to the line-up method, the difference in accuracy between the two methods is at least 115mm at all three distances. It is interesting to note the insignificance between the three distances within the forearm pointing method. This may suggest a high tolerance with increasing distance from target for this pointing method.

On the other hand, the straight arm pointing method is highly affected by the increase in distance from target. This is illustrated by the significant difference across all distances and the relatively high linear increase (65.7mm per meter). Compared to forearm pointing, the accuracy at 2m and 3m are not significant. While the only difference between forearm and straight arm pointing is at 1m. This may be due to the higher level of feedback given by the longer arm extension, and that the straight arm pointing resembles the line-up method at close proximity to the target.

4 Conclusion

From this experiment, we have identified inaccuracies in users pointing performance, which varies depending on the strategy used. We observed that the line-up method is the most accurate pointing method, and that the straight arm method is more accurate than the forearm method only at a distance of one metre. These results provide us a basic understanding on the accuracy of each method. The question of whether these methods are sufficiently accurate depends largely on the specific task, and the decision rest on the individual interaction designer.

To complement this study, a follow-on experiment may be conducted to study the ergonomics of these pointing methods. Understanding the trade off between accuracy and ease-of-use can assist future vision-based hand pointing interaction researchers and practitioners to decide the input strategy that best suit their users in completing the required tasks.

References

1. Cipolla, R. and Hollinghurst, N. J., "A Human-Robot Interface using Pointing with Uncalibrated Stereo Vision," in *Computer Vision for Human-Machine Interaction*, R. Cipolla and A. Pentland, Eds.: Cambridge University Press, 1998, pp. 97-110.
2. Colombo, C., Bimbo, A. D., and Valli, A., "Visual Capture and Understanding of Hand Pointing Actions in a 3-D Environment," *IEEE Transactions on Systems, Man, and Cybernetics,* vol. 33, pp. 677-686, August 2003.
3. Hung, Y.-P., Yang, Y.-S., Chen, Y.-S., Hsieh, I.-B., and Fuh, C.-S., "Free-Hand Pointer by Use of an Active Stereo Vision System," in *Proceedings of 14th International Conference on Pattern Recognition (ICPR)*, Brisbane (1998), pp. 1244-1246.
4. Nickel, K. and Stiefelhagen, R., "Pointing Gesture Recognition based on 3D-Tracking of Face, Hands and Head-Orientation," in *Proceedings of ICMI '03 International Conference on Multimodal Interfaces*, Vancouver (2003), pp. 140-146.