Proceedings of the 2nd International Conference on Human-Computer Interaction
Prague, Czech Republic, August 14-15, 2014
Paper No. 125

New Technique to Estimate Positions of 3-D Images by Fitting Velocity of Reaching Movements into Function of *Chi*-Squared Distribution

Piyarat Silapasuphakornwong, Masahiro Suzuki, Hiroshi Unno, Kazuki Matushima, Kazutake Uehira, and Yuichi Takeda
Kanagawa Institute of Technology
1030 Shimo-ogino, Atusgi, Kanagawa, Japan 243-0292
silpiyarat@gmail.com; msuzuki@ctr.kanagawa-it.ac.jp; unno@nw.kanagawa-it.ac.jp;
122062.mk@gmail.com; uehira@nw.kanagawa-it.ac.jp; yuichi@gen.kanagawa-it.ac.jp

Abstract **-** This paper presents an improvement to the technique that we previously proposed to estimate the positions of 3-D images that users see. Applications in which users' bodies interact with 3-D images require such estimates to execute interactions when bodies are at the positions of 3-D images that users see. Although the previously proposed technique, which estimated positions by fitting the velocity of reaching movements into a Gaussian function, was able to meet this requirement, the accuracy and precision of estimates were insufficient in practical use. The improvement that is proposed in this study is to use a function of the *chi*-squared distribution instead of a Gaussian function. The improved technique was evaluated by comparing its accuracy and precision of estimates with those of the previously proposed technique. The results obtained from the evaluation demonstrated that the improved technique could more accurately and precisely estimate positions than the previously proposed technique.

Keywords**:** 3-D displays, 3-D images, Augmented reality, Mixed reality, Interactive applications, Interaction, Human body movements, Reaching.

1. Introduction

Applications in which users' bodies interact with 3-D images are attractive applications of 3-D displays. For example, users in interactive applications of on-line shopping can touch and take hold of 3-D images of products, which are seen as floating above the screens, with their own hands, and can feel as if they were touching and grasping actual products. Users in interactive applications in training for precision work can pick or cut 3-D images of objects with actual tools that they are handling, and can feel as if they were picking or cutting actual objects. In fact, many researchers have studied such interactive applications. Thus, interactive applications need to be provided to enable practical use in the near future.

Interactive applications require the positions of 3-D images that users see to be estimated and interactions to be executed when their bodies are there because they naturally expect that. Such phenomena that are inconsistent with users' expectations strongly reduce the usability of interactive applications. Thus, techniques to meet this requirement are crucial to provide interactive applications for practical use.

We previously proposed and evaluated a new technique that met this requirement (Suzuki et al. 2009; Uehira et al., 2010); however, the accuracy and precision of estimating the positions of 3-D images that users saw were insufficient for practical use. The acceptable error range that we considered to be the goal for practical use was between about –1 cm and +1 cm. The results obtained from the previous evaluation indicated that the mean error range was between about –2 cm and +2 cm (Suzuki et al. 2009, 2012; Takazawa et al., 2011; Uehira et al., 2010; Unno et al., 2012). Thus, improvement to the proposed technique was necessary to provide interactive applications for practical use.

We propose and evaluated an improvement to the previously proposed technique to more accurately and precisely estimate the positions of 3-D images that uses see. The previously proposed technique is described in Section 2. The improved technique is proposed in Section 3, and its evaluation is described in Sections 4 and 5. The conclusion is presented in Section 6.

2. Previously Proposed Technique

We previously proposed a technique that executed interactions when users' bodies were at the positions of 3-D images that were estimated based on human body movements (Suzuki et al. 2009; Uehira et al., 2010). The underlying principle of the proposed technique is as follows. According to previous research on human body movements, the velocity of a reaching movement, which is the basic movement when users interact with objects, as a function of time approximates to a bell curve. Therefore, the velocity of a reaching movement as a function of time in the proposed technique is fitted into a Gaussian function, which is the well-known bell curve (Uehira et al., 2010):

$$v = (p_2 / p_1 \sqrt{\rho/2}) e^{-2(t-p_4)^2 / p_3^2} \quad (1)$$

where v is the velocity of the reaching movement, t is time, and p_1, p_2, p_3, and p_4 are parameters for fitting. The p_2 is the area of the fitted Gaussian function, i.e., the distance between the starting and finishing positions of the reaching movement. The object that a user is reaching out to exists at the finishing position of the reaching movement; therefore, the positions of 3-D images that users see can be estimated by using this parameter when users reach out to the images. Our previous evaluations demonstrated that such estimates could be completed before reaching movements had finished, and that interactions could be executed at exactly the time when users' hands were at the estimated positions. Thus, the proposed technique can meet the requirement that was previously described.

However, the previously proposed technique was insufficient for practical use because of its error range in estimating the positions of 3-D images that users see. We considered an error range between about –1 cm and +1 cm to be acceptable for practical use. Our previous evaluations indicated that the mean error range was between about –2 cm and +2 cm (Suzuki et al. 2009, 2012; Takazawa et al., 2011; Uehira et al., 2010; Unno et al., 2012). Although this range was close to the range that we considered as a goal in practical use, the difference between the ranges is too large to disregard it. Thus, the previously proposed technique is unacceptable for practical use without any improvements being made to it.

3. Improved Technique

According to previous studies on human body movements, reaching movements consist of a bullet movement and a corrective movement. The bullet movement occurs first on the basis of a feedforward mechanism, and most reaching movements are this kind of movement. The corrective movement occurs when the hands are close to objects. A visual feedback mechanism underlies this movement. These two movements are crucial to reach out to objects as rapidly and correctly as possible.

We assumed that the unacceptable error range in the previously proposed technique was due to the corrective movement. As previously described, the previously proposed technique estimated positions by fitting the velocity of a reaching movement into a Gaussian function. Such estimates were only based on the bullet movement because a symmetrical bell curve such as a Gaussian function is only able to describe the velocity of movement that occurred on the basis of a feedforward mechanism. Therefore, it is no wonder that the corrective movement caused the unacceptable error range because a Gaussian function was not able to describe the velocity of movement that occurred on the basis of a visual feedback mechanism.

We also assumed that we could achieve an acceptable error range by using estimates not only based on the bullet movement but also the corrective movement. According to previous studies on reaching movements, the velocity of reaching movements in which the corrective movement clearly appears follows a bell curve with a long tail. Therefore, we can estimate positions based on both the bullet and corrective movements by fitting the velocity of reaching movements into a long-tailed bell curve.

Based on these assumptions, we propose an improvement to the previously proposed technique by using a function of the *chi*-squared distribution, which is a long-tailed bell curve, instead of a Gaussian function:

$$v = \left(p_2 \Big/ \left(2^{p_1/2} \times \Gamma(p_1/2)\right)\right) \times (t \times p_3 + p_4)^{p_1/2-1} \times e^{-(t \times p_3 + p_4)/2} + p_5 \qquad (2)$$

where v is the velocity of the reaching movement, t is time, p_1, p_2, p_3, p_4, and p_5 are parameters for fitting, and Γ is the gamma function. The length of the tail depends on p_1. The integrated value for the fitted function is the distance between the starting and finishing positions of the reaching movement; therefore, the positions of 3-D images that users see can be estimated by using this value. Thus, this improvement enables us to estimate positions based on both the bullet and corrective movements.

We also propose using the improved technique only under conditions where the velocity of reaching movements fits into a function of the *chi*-squared distribution better than a Gaussian function. When a Gaussian function is better than a function of the *chi*-squared distribution, the previously proposed technique is used. Thus, either the improved technique or the previously proposed technique is used under conditions that are most suitable for either of them.

4. Methodology of Evaluation

The improved technique that was previously described was evaluated by comparing its error range with that of the previously proposed technique under conditions where the improved technique was used. The absolute value of error, which indicates the error range, is computed as:

$$E_{\mathrm{absolute}} = \sqrt{(P_{\mathrm{estimated}} - P_{\mathrm{actual}})^2} \qquad (3)$$

where E_{absolute} is the absolute value of error, $P_{\mathrm{estimated}}$ is the position of a 3-D image that is estimated with the improved technique or the previously proposed technique, and P_{actual} is the position of a 3-D image that a user sees. When the absolute value of error with the improved technique is smaller than that with the previously proposed technique, we can conclude that the improved technique is better than the previously proposed technique.

The data that were collected in our previous experiment (Suzuki et al., 2012; Takazawa et al., 2011) were used to compute the absolute value of error. Five participants in this experiment reached out to 3-D images whose positions were defined by binocular disparity and were 40 cm from them, and the positions of their hands were measured during reaching movements. The data from the beginning of a reaching movement to the time when the velocity of the reaching movement decreased into a quarter of the peak velocity were used to compute the statistic that indicated the degree of fitting into a function of the *chi*-squared distribution or a Gaussian function. When the statistic of a function of the *chi*-squared distribution was better than that of a Gaussian function, the data were also used to compute the positions of 3-D images that were estimated using the improved technique or the previously proposed technique. Equation (2) was used to estimate positions with the improved technique, and Eq. (1) was used to estimate them with the previously proposed technique. The data when participants finished making reaching movements were used to compute the positions of 3-D images that the participants saw. The absolute value of error was computed for each trial.

5. Results and Discussion

We computed the mean value and standard error of the absolute value of error and have summarized them in Fig. 1. We also conducted a *t*-test, and found significant differences at the 5% level ($t_{(5)} = 2.669$, $p < .05$). These results indicate that the absolute value of error for the improved technique was smaller than that for the previously proposed technique.

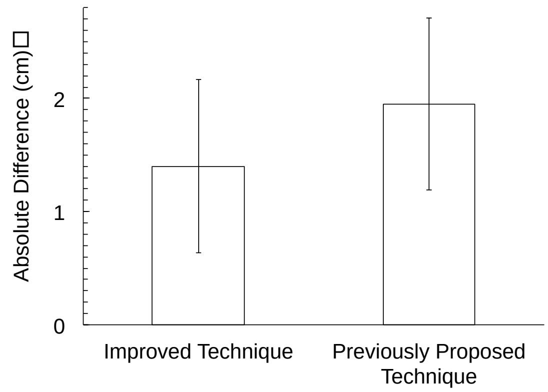

Fig. 1. Mean value and standard error of absolute value of error.

We demonstrated that we could reduce the error range by using the improved technique. However, we need to reduce it more for practical use because the error range of the improved technique, which was between about –1.4 cm and +1.4 cm, was larger than the acceptable error range. We assume that we can reduce the error range by using another function that is better than a function of the *chi*-distribution to fit the velocity of reaching movements into. We will examine this assumption in the near future.

6. Conclusion

We proposed and evaluated an improvement to the technique that we previously proposed to estimate the positions of 3-D images that users see. The positions of 3-D images that users see in the improved technique were estimated by fitting the velocity of a reaching movement into a function of the *chi*-squared distribution instead of a Gaussian function in the previously proposed technique. The error range of the improved technique was compared with that of the previously proposed technique, and the former was smaller than the latter under conditions that a function of the *chi*-squared distribution was better than a Gaussian function to fit the velocity of the reaching movement into. We demonstrated that the improved technique could more accurately and precisely estimate positions than the previously proposed technique.

Acknowledgements

This study was supported by JSPS KAKENHI Grant Number 24500155.

References

Suzuki M., Takazawa K., Uehira K., Unno H. (2012). Prediction of visually perceived location using reaching action and effect of reaching distance on it "Proceedings of SPIE on CD-ROM: Electronic Imaging 2012: Volumes 8288–8305," Burlingame, CA, Jan. 22–26, pp. 82890N-1–82890N-6.

Suzuki M., Yokono M., Uehira K. (2009). New technique for prediction of visually perceived location of virtual object in mixed/augmented reality using observer's action "The Eleventh IEEE International Symposium on Multimedia," San Diego, CA, Dec. 14–16, pp. 418-424.

Takazawa K., Suzuki M., Uehira K. (2011). Prediction of visually perceived location of virtual objects using observer's reaching motion: Influence of reaching distance on precision and accuracy "International Conference on Imaging and Printing Technology," Bangkok, Thailand, Aug. 17–20, pp. 106–110.

Uehira K., Yokono M., Suzuki M. (2010). Prediction of visually perceived depth of virtual objects from observer's actions using approximation obtained by Gaussian function "Proceedings of International Conference on Artificial Reality and Telexistence," Adelaide, Australia, Dec. 1–3, pp. 189–192.

Unno H., Takazawa K., Suzuki M., Uehira K. (2012). Interaction technique between virtual objects and user body by predicting visually perceived locations "Proceedings of the IIEEJ Image Electronics and Visual Computing Workshop," Kuching, Malaysia, Nov. 21–24, 4C-3.