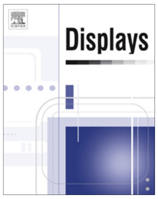
Displays 42 (2016) 19–24

Contents lists available at ScienceDirect

Displays

journal homepage: www.elsevier.com/locate/displa



New technique of obtaining visually perceived positions of 3-D images using movements of users' bodies ☆

Masahiro Suzuki *, Kazutake Uehira

Kanagawa Institute of Technology, Atsugi, Kanagawa, Japan

ARTICLE INFO

Article history:
Received 3 December 2014
Received in revised form 5 February 2016
Accepted 26 February 2016
Available online 2 March 2016

Keywords:
3-D display
3-D image
Interactive application
Interaction
Human body
Human body movement

ABSTRACT

Three-dimensional (3-D) images are perceived as images that float in front of the screens of 3-D displays. Users should be able to interact with these images instantaneously and accurately in applications where their bodies actually seen by them interact with the images. However, conventional techniques using just binocular disparity are too slow and inaccurate. Therefore, we propose a new technique where the visually perceived positions of images are obtained from the body movements of users. The feasibility of this technique was evaluated in an experiment using the positions obtained from users as they reached out to touch the images. These positions were closer to the visually perceived positions of the images than those calculated from binocular disparity. These findings demonstrate the feasibility of the proposed technique for 3-D interactive applications.

1. Introduction

Three-dimensional (3-D) displays have recently seen widespread use among consumers. Currently, they are primarily used to enable consumers to watch 3-D images, e.g., 3-D films, 3-D TV programs, and 3-D games, which are seen as floating images in front of the screens of 3-D displays. Applications where users' own bodies, which are actually seen by the users unlike images captured using video cameras, interact with 3-D images will become very important in the future. In fact, many researchers have studied such interactive applications [1–3]. The underlying principle behind interactive applications is to carry out interactions when the users' bodies are exactly at the positions of the floating images in front of the screens. Users typically expect to interact with 3-D images when they see their bodies touch the floating images. Therefore, interactive applications require interactions to take place when the users' bodies are exactly at the visually perceived positions of the floating images. Although this requirement seems easy to fulfill, conventional techniques find it difficult to do so with the 3-D displays that are used in practice.

Three-dimensional displays in current practical use employ binocular disparity as depth information for 3-D images [4–6]. However, the visually perceived positions of 3-D images often differ from the positions calculated from binocular disparity. Therefore, it is difficult to satisfy the aforementioned requirement if the interactions are carried out when the users' bodies are exactly at the positions calculated from binocular disparity.

Pre-experiments can be conducted before trials of interactions to measure the visually perceived positions of floating images in front of 3-D display screens, and interactions can be implemented when the users' bodies are exactly at the positions measured in the pre-experiments. This technique can satisfy the requirement for interactive applications if the visually perceived positions during the trials of interactions match those during the pre-experiments; however, these two sets of visually perceived positions often differ, which makes it difficult to fulfill the requirement.

Because conventional techniques do not meet the requirement for interactive applications, phenomena that are inconsistent with the users' typical expectations often occur. For example, users often find themselves not able to interact with the floating images in front of the screens although they see their bodies touch the floating images. Conversely, they can often interact with the floating images although they do not see their bodies touch them. Such phenomena adversely impact the usability of interactive applications, especially those that involve precise work.

Conventional techniques suffer from not only the problems described previously but also problems with their specific required preparation. One method, which implements interactions when the users' bodies are exactly at the positions calculated from binocular disparity, requires the interocular distances of all users to be

---

☆ This paper was recommended for publication by Richard H.Y. So.

* Corresponding author.

E-mail address: msuzuki@ctr.kanagawa-it.ac.jp (M. Suzuki).

http://dx.doi.org/10.1016/j.displa.2016.02.003

measured to calculate the exact positions from binocular disparity. Another method, which implements interactions when the users' bodies are exactly at the positions measured in pre-experiments, requires pre-experiments to be conducted for every instance of the use of the application. Thus, conventional techniques reduce the usability of interactive applications because users cannot readily begin to use them.

This paper presents a new technique that can satisfy the requirement for interactive applications. The visually perceived positions of floating images in front of screens are obtained from human body movements with this technique, and interactions are executed when the users' bodies, which they can actually see, are exactly at the obtained positions. This paper also demonstrates the feasibility of the proposed technique, which was evaluated in an experiment. In the evaluation, the proposed technique was compared with the first conventional technique (see the second paragraph) with which usability decreased less than with the second method (see the third paragraph) because the preparation for it was easier.

2. Proposal

The critical process in fulfilling the requirement for interactive applications is to obtain the visually perceived positions of floating images in front of screens before the users' bodies reach these positions. The proposed technique achieves this process by using the characteristics of human body movements. The velocity of reaching movements, which are basic movements by which users interact with objects, follows a bell curve as a function of time according to previous studies on human body movements [7–11]. With the proposed technique, data on velocity are fitted into a bell-shaped function, e.g., a Gaussian function, before the reaching movements have finished, and the fitted function is used to obtain the positions where the users will finish their reaching movements, namely, the positions where the velocity of reaching movements will reach zero (Fig. 1). Because users will stop their reaching movements when they see their hands touch the floating images, the positions obtained with the fitting function are the visually perceived positions of the floating images. Thus, the proposed technique can be used to obtain the visually perceived positions of floating images before the users' bodies arrive there, and can fulfill the requirement for interactive applications.

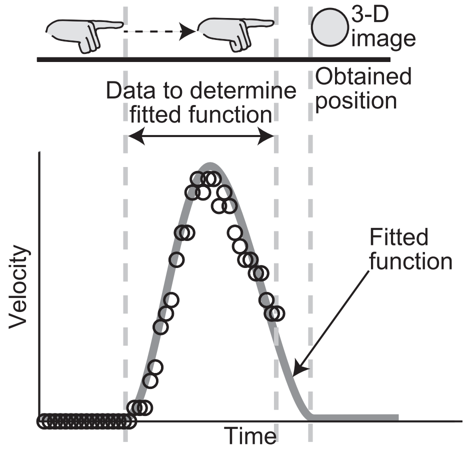


Fig. 1. Principle underlying the proposed technique. The fitted function was determined on the basis of the data on velocity collected before a reaching movement was finished, and the position where the velocity would decrease to zero was obtained from it. The obtained position was also the position where a reaching movement would have finished; therefore, a 3-D image was seen at the obtained positions.

Fig. 2 outlines the four steps of the algorithm used by the proposed technique. In the first step, the current velocity of reaching movements is compared with $V_{\mathrm{beginning}}$, which is the threshold for detecting the beginning of reaching movements; if the current velocity of reaching movements is faster than $V_{\mathrm{beginning}}$, then reaching movements are considered to have started (Fig. 2A). After the beginning of reaching movements is detected, the peak velocity of reaching movements is detected in the second step (Fig. 2B). The velocity V_{obtain}, which is the trigger to obtain the visually perceived positions of floating images, is determined on the basis of the peak velocity, e.g., half the peak velocity or a quarter of the peak velocity. The current velocity of reaching movements is compared with V_{obtain} in the third step (Fig. 2C); if it is lower than V_{obtain}, the data on velocity collected until this moment are fitted to a Gaussian function:

$$V = \frac{p_2}{p_1\sqrt{\frac{\pi}{2}}} e^{-2\frac{(t-p_4)^2}{p_3^2}}, \tag{1}$$

where V is the velocity of reaching movements, t is time, and p_1, p_2, p_3, and p_4 are fitting parameters. p_2 is the area of the fitted function, namely, the positions obtained in the third step. After the visually perceived positions of the floating images are obtained, interaction is achieved if the users' bodies are exactly at the obtained positions (Fig. 2D). Thus, the proposed technique can satisfy the requirement for interactive applications.

3. Evaluation methodology

The feasibility of the proposed technique was evaluated by assessing the accuracy and precision of the obtained positions. The absolute value of the difference between the position obtained with the proposed technique and the position where a floating image in front of the screen was actually perceived visually was calculated as

$$D_{\mathrm{absolute}} = \sqrt{(P_{\mathrm{obtained}} - P_{\mathrm{perceived}})^2}, \tag{2}$$

where D_{absolute} is the absolute difference, P_{obtained} is the position obtained with the proposed technique, and $P_{\mathrm{perceived}}$ is the visually perceived position of the floating image. The absolute difference decreases when the obtained position is closer to the visually perceived position. Thus, the accuracy and precision of the obtained positions in this evaluation were assessed using the absolute difference.

An experiment was performed to collect the required data. Participants in this experiment reached out to a floating image with their hands that they could actually see, and their reaching movements were measured. After the experiment, the position with the proposed technique, the visually perceived position, and the absolute difference between the two were calculated using the collected data. Details of the experiment and computation are described in the following subsections.

3.1. Experimental apparatus

The experiment was conducted in a well-lit room. Fig. 3 shows a schematic of the apparatus. Three-dimensional images were generated with a workstation (Precision 390, Dell, Round Rock, TX, USA) and a cathode-ray tube (CRT) display (A201H, Iiyama, Tokyo, Japan). The CRT display was located 70 cm from the participants, and the center of the screen was on their median plane and at eye level. The participants wore a pair of liquid crystal shutter glasses for stereoscopic vision (NuVision 60GX, MacNaughton, Beaverton, OR, USA), and their heads were stabilized using a chinrest and a forehead rest. The shutter glasses were synchronized with the CRT display. Videos of the participants' reaching