# Prediction of Visually Perceived Depth of Virtual Objects from Observer's Actions Using Approximation Obtained by Gaussian Function

Kazutake Uehira, Minoru Yokono, and Masahiro Suzuki

Human Media Research Center, Kanagawa Institute of Technology

**ABSTRACT**

This paper presents a technique for predicting the visually perceived depth of virtual objects using an approximation of observer action obtained with the Gaussian function. It is very important for observers to accurately predict the visually perceived depth of virtual objects in augmented reality to enable the human hands and objects to smoothly interact. We evaluated the technique we propose in this study through subjective tests. The visually perceived depth was predicted by fitting the velocity of the hand as a function of action time into the Gaussian function. The results from the subjective tests demonstrated that the accuracy of prediction with the new method was sufficient for practical use.

**KEYWORDS:** Virtual objects, depth perception, stereoscopic.

## 1 INTRODUCTION

A great deal of research on augmented reality where users can interact with virtual objects that are seen in the real world has recently become popular. It is natural for observers who are interacting with virtual objects in augmented reality to treat virtual objects directly with their hands or other parts of their bodies and not manipulate virtual hands or levers as if they were treating real objects in the real world. For users to directly treat virtual objects with their bodies, the system needs to accurately know their position, especially the depth of the virtual objects that users visually perceive. This is because if the system does not know this depth, it cannot appropriately react with virtual objects or user hands; for example, move the virtual objects and provide the sense of touch when users see their hands reaching the surface of these objects.

When virtual objects are displayed on a stereoscopic display, the depth perceived by the observers varies depending on individuals [1]–[4], because the display only provides binocular disparity and convergence as a cue for perceiving the depth of the displayed objects. This has also often differed within the individuals themselves [2]. Therefore, it is crucial for a system to have the means to precisely know the depth of virtual objects perceived by observers before their bodies reach them to enable the human body and virtual object to smoothly interact.

We previously proposed a technique of predicting the visually perceived location by using an approximated function of an observer's hand movements when it was reaching for an object to touch it [5]. We evaluated the proposed technique by expressing the hand position as a function of time using a logistic function. The results indicated that the perceived depth roughly predicted using the logistic function was not sufficiently accurate for practical use.

K. Takashima ea al reported the technique to predict a target icon on a GUI from the peak velocity of the movement of a mouse cursor [6]. Our study is similar to their study in predicting the target position using the velocity of a user's hand, however, it differ from their study in predicting perceived depth of a virtual object using the movement of a user's hand in the real space.

This paper presents a new method that uses the Gaussian function to express the relation between the velocity of the hand and time to improve the accuracy of predicting the perceived depth of a virtual object This paper also describes experiments on evaluating the new technique and presents the results from these that demonstrate the feasibility of the technique we propose.

## 2 PREDICTION USING GAUSSIAN FUNCTION

There are some universal features in the movement of the hand when it is reaching for a real object to touch [7]–[9]]. One of these is the velocity profile of the observer's hand as a function of the time in bell form. We expected that the velocity profile would be the same for the virtual object displayed on a stereoscopic display. We then applied a Gaussian function that had a bell form.

Figure 1 indicates how the perceived depth can be predicted before the observer's hand reaches a virtual object. The bottom of Fig. 1 shows the relationship between time from when the observer's hand begins to move and its velocity. It reaches the virtual object at time $T_E$ when the velocity becomes zero. $P_E$ is the perceived depth of the virtual object. We assumed that the relation between the velocity of the hand and time in this technique could be expressed by the Gaussian function as

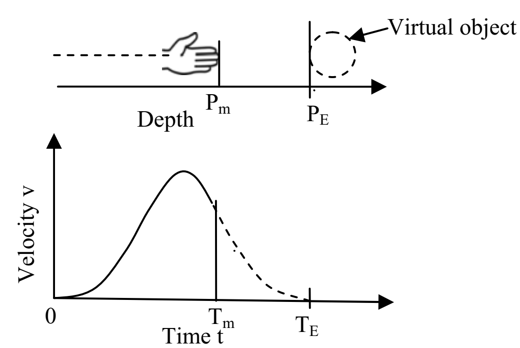


Figure 1 How perceived depth can be predicted before observer's hand reaches virtual object.

$$v(t) = \frac{p_E}{p_2\sqrt{\frac{\pi}{2}}} e^{-2\frac{(t-p_1)^2}{p_2^2}} \quad (1)$$

1030 Shimo-ogino, Atsugi-shi, Kanagawa, 243-0292, Japan
E mail: uehira@nw.kanagawa-it.ac.jp

20th International Conference on Artificial Reality and Telexistence (ICAT2010), 1-3 December 2010, Adelaide, Australia
ISBN: 978-4-904490-03-7 C3450 ©2010 VRSJ

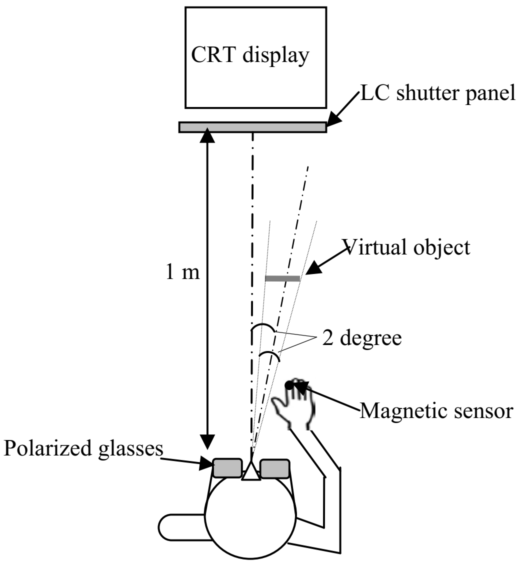


Figure 2 Layouts for evaluating perceived depth of virtual object.

where $P_1$ and $P_2$ are constant and $P_E$ is the perceived depth. Assuming that the hand is at a depth of Pm ($<P_E$) at time $T_m$ ($<T_E$), the data on velocity can be obtained before $T_M$ (solid line in graph in Fig. 1). The constants including $P_E$ in Eq. 1 can be determined by using these data and the least squares method. The system can then know the perceived depth of the virtual object, $P_E$, before the hand reaches it.

## 3 Evaluation

### 3.1 Method of evaluation

We conducted an experiment in this study to evaluate the technique we propose by measuring the perceived depth and predicted perceived depth. Figure 2 shows the layout for the experiment. The virtual object was generated with a workstation and displayed on a 17-inch CRT display. To enable the virtual object to be seen in three dimensions, a liquid crystal shutter panel that permitted dichoptic presentation by polarization was placed directly in front of the screen and the observer wore a pair of polarizing glasses. The screen was located 1 m from the observer's head. The center of the screen was on the observer's median plane and at the observer's eye level.

The virtual object was a square whose size was 2-degrees of width and 2-degrees of height. The center of the virtual square was located 2-degrees to the right from the observer's median plane and at the observer's eye level. The virtual square was red. The depth of the virtual object was adjusted by each observer so that they could see it at the same depth as the real paper object that was 39.6, 44.6, and 49.6 cm in depth and had the same size as the virtual object. We call these depths set depths in this paper.

The observer was asked to reach for the virtual square with his/her right hand and to touch it with his/her index finger. Before beginning the task, the observer was instructed to determine whether his/her index finger touched the virtual square when they overlapped. We defined this depth as the perceived depth in this study. Although the set depth mentioned above is also the perceived depth, these two perceived depths do not necessarily agree. Considering the interaction with virtual objects, the depth at which the observer visually perceives that his/her finger reaches the object is more important as the perceived depth. This is why we defined the depth at which the observer perceived that his/her index finger came to touch the object and stopped moving it as the perceived depth.

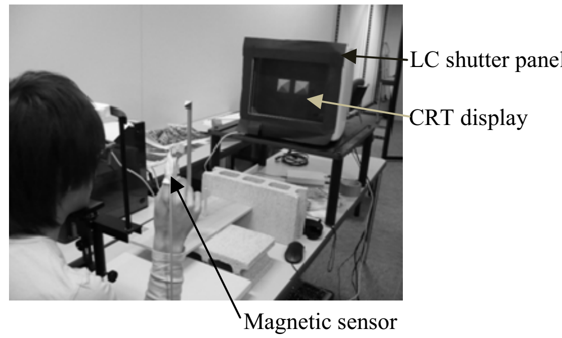


Figure 3 Scene of the experiments.

We also measured the predicted perceived depth on real objects that had the same depth（39.6, 44.6, and 49.6 cm）and the same size as the virtual object as a reference.

The measurements were conducted in a lit room. The observer's head was fixed with a chin rest and a forehead rest. The depth of the index finger of his/her right hand was measured with a magnetic sensor. The sampling rate was 50 Hz. The data for the velocity of the observer’s hand were calculated by using these depth data for the index finger. The observers were four 21 or 22 year old students who had normal eyesight. Figure 3 shows one of scenes for the experiments.

The velocity of the index finger as a function of action time was fitted into Eq. 1. We used two conditions that we called the 1/2 and 3/4 conditions. Under the 1/2 condition, data in the time range from the beginning of action to the time the velocity of the index finger reached a peak were used for fitting and under the 3/4 condition, data in the time range from the beginning of action to the time when the velocity of the index finger decreased into half of the peak were used.

### 3.2 Results and Discussion

Figure 4 shows some of the experimental results. These data are for the virtual object that was 49.6 cm in depth. Although we measured the velocity of the index finger and the Gaussian function was obtained by fitting the velocity data, we used the integrated Gaussian function by time in Fig. 4 so that we could see the predicted perceived depth more easily; therefore, these data express the depth as a function of time. The dashed lines in Fig. 4 indicate the perceived depth defined as previously mentioned. The solid line indicates the integrated Gaussian function obtained using the data for the 3/4 condition and dotted line indicates that using those for the 1/2 condition. The saturated values for these curves indicate the predicted perceived depth. It can be seen from Fig. 4 that the perceived depth of the virtual object differed depending on individuals. These data on perceived depth that depended excessively on individuals support the need for our new technology. We can see from Fig. 4 that the predicted data for the 1/2 condition are not accurate; however, those for the 3/4 condition are very accurate and the predicted depth was almost the same as the visually perceived depth. These results indicate that the data around the midpoint or right after the midpoint fit the Gaussian function better than those in the first half of the action time.

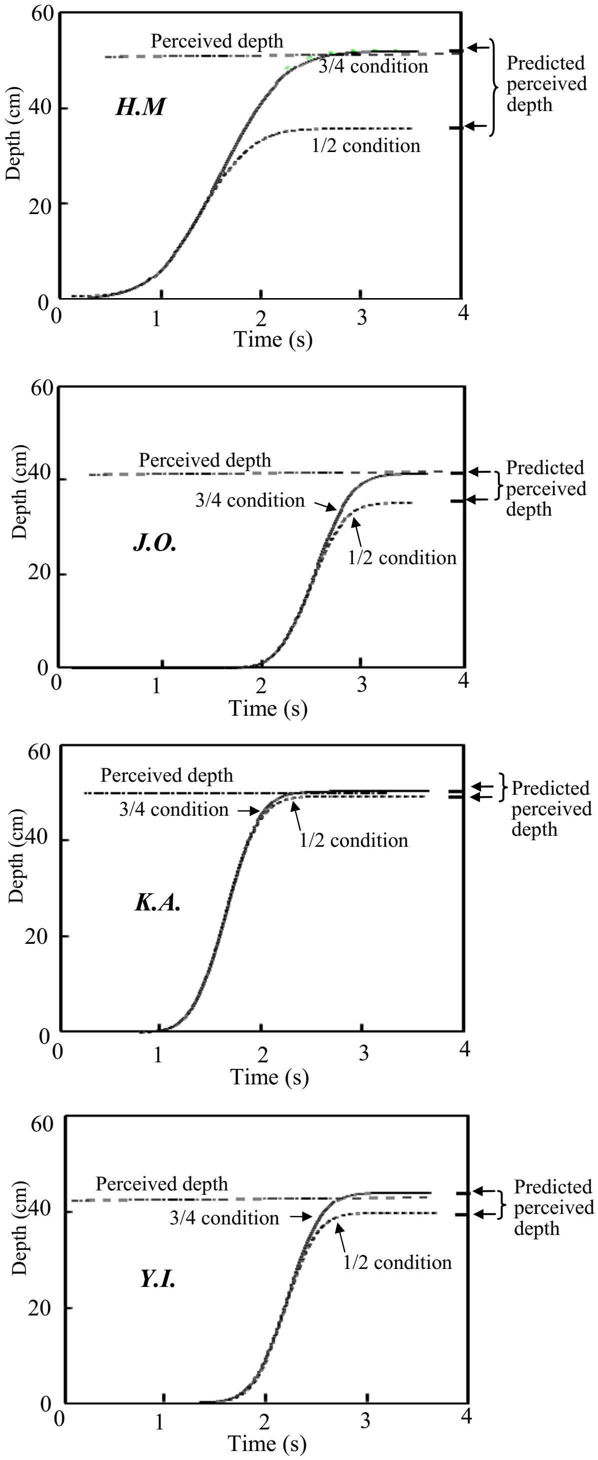


Figure 4 Experimental results for four observers for virtual object that was 49.6 cm in depth.

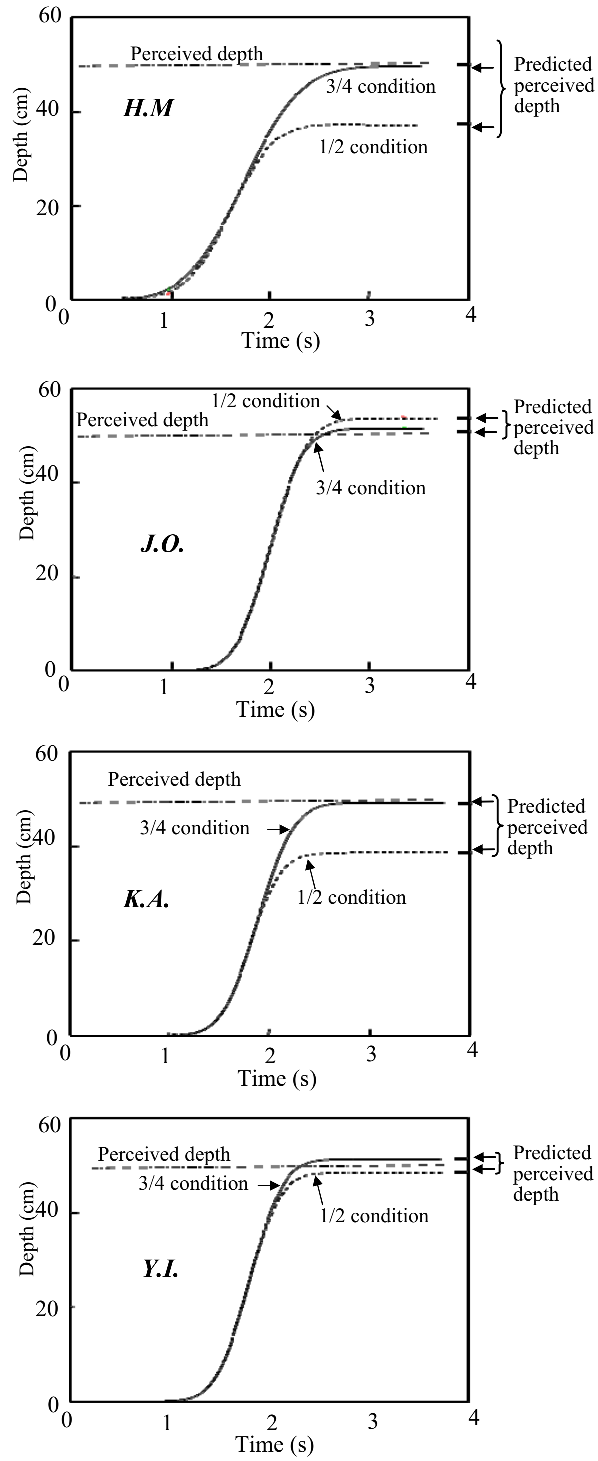


Figure 5 Experimental results for four observers for real object that was 49.6 cm in depth.

Figure 5 shows the results for a real object that was 49.6 cm in depth. For a real object, the perceived depth is the same as the depth of the real object. We can see from Fig. 5 that for also a real object, the predicted data for the 1/2 condition are not accurate; however, those for the 3/4 condition are accurate.

Comparing the results for virtual objects and a real object, it can be inferred that the velocity profiles of the hand for both cases are similar, as we expected, and this also makes it possible to use the bell form function for virtual objects.

Table 1 lists the results for the perceived depth and predicted perceived depth for all set depths. We can see from Table 1 that

Table 1 Results for perceived and predicted perceived depth for virtual and real object. Data on left are for virtual object and those on right are for real object. Perceived depths are for virtual objects. Perceived depth for real object is the same as set depth.

| Observer | Set depth (cm) | Perceived depth (cm). | Predicted perceived depth (cm) | |
|---|---|---|---|---|
| | | | 1/2 condition | 3/4 condition |
| H.M | 39.6 | 44 | 37 / 33 | 44 / 38 |
| | 44.6 | 45 | 43 / 41 | 45 / 44 |
| | 49.6 | 52 | 38 / 38 | 53 / 50 |
| J.O | 39.6 | 41 | 37 / 29 | 41 / 39 |
| | 44.6 | 40 | 37 / 50 | 41 / 47 |
| | 49.6 | 41 | 36 / 54 | 41 / 51 |
| K.A | 39.6 | 43 | 40/ 34 | 42 / 41 |
| | 44.6 | 45 | 53 / 47 | 45 / 44 |
| | 49.6 | 50 | 50 / 39 | 50 / 50 |
| Y.I | 39.6 | 40 | 36 / 36 | 40 / 40 |
| | 44.6 | 45 | 61 / 47 | 50 / 45 |
| | 49.6 | 43 | 40/ 49 | 44 / 52 |

for all set depths, the predicted data for the 1/2 condition are not accurate; however, those for the 3/4 condition are accurate.

Comparing the results obtained from this study and those from a previous study [5], we can see that the data obtained with the Gaussian function are more accurate than those obtained with the logistic function where there was about 5% error for every observer. Therefore, the Gaussian function suits our technique better than the logistic function.

The perceived depth using the 3/4 condition can be predicted within some hundreds of milliseconds before the observer's hand reaches the virtual object. This is enough time to do the processing that is needed. Therefore, these results suggest that the accuracy of prediction with the new method is sufficient for practical use.

## 4 Conclusion

A technique to predict the visually perceived location of virtual objects using an approximation of the observer's action to the Gaussian function was studied to make virtual objects directly treated with users' hands more natural or smoother. The results of the experiments indicate that at 3/4 of the whole time to reach the virtual object, its depth could be predicted very accurately and this was sufficient to enable it to be implemented in practical use.

We intend to demonstrate quantitatively that humans can interact smoothly with virtual objects in future work.

## Acknowledgement

This study was supported by the Support Center for Advanced Telecommunications Technology Research in Japan.

## References

[1] M. F. Bradshaw, A. Glennerster, and B. J. Roggers. The effect of display size on disparity scaling from differential perspective and vergence cues. *Vision Research*, 36: 1255–1264, 1996.

[2] W. Richards. Anomalous stereoscopic depth perception. *Journal of the Optical Society of America*, 61: 410–414, Mar 1971.

[3] M. Ritter. Effect of disparity and viewing distance of perceived depth. *Perception and Psychophysics*, 22: 400–407, 1977.

[4] H. Wallach and C. Zuckerman. The constancy of stereoscopic depth. *American Journal of Psychology*, 76: 404–412, 1963.

[5] M. Suzuki, M. Yokono, and K. Uehira. New technique for prediction of visually perceived location of virtual object in mixed/augmented reality using observer's action. *IEEE International Symposium on Multimedia* (San Diego, CA, Dec 14–16, 2009), in press.

[6] K. Takahashi, T. Asano, E. Sharlin, Y. Kitamura, and F. Kishino. Proposing a Target Prediction Interface Using a Peak Velocity during Pointing Tasks (User Interfaces and Interactive System) [in Japanese]. *Transactions of Information Processing Society of Japan* 48 ( 2): 929–938, 2007

[7] C. G. Atkeson and J. M. Hollerbach. Kinematic feature of unrestrained vertical arm movements. *Journal of Neuroscience*, 5: 2318–2330, Sep. 1985.

[8] E. Bizzi and F. A. Mussa-Ivaldi. Geometrical and mechanical issues in movement planning and control. In M. I. Posner, ed., *Foundation of Cognitive Science*, MIT Press, Cambridge, Massachusetts, 1989.

[9] P. Morasso. Spatial control of arm movements. *Experimental Brain Research*, 42: 223–227, 1981.