Evaluation of Visuo-haptic Feedback in a 3D Touch Panel Interface

Abstract
In this paper we evaluate the relation between visual and haptic feedback in a 3D touch panel interface and show the optimal latency for natural interaction. We developed a system that consists of an autostereoscopic display and a high-speed stereo camera. With this system, virtual objects are stereoscopically-presented, and the objects respond to the finger movement that is obtained using the stereo camera. We conducted an experiment to evaluate visual and haptic synchronization and the result showed that visual and haptic feedback was the most synchronized with latencies around 150 ms, while finger and button movements were more synchronized with smaller latencies. We also conducted a comparison experiment to explore which synchronization is more important and the result showed that the visual synchronization of finger and button movements is more important than visual and haptic synchronization.

Author Keywords
Stereoscopic display; Passive haptic feedback; Latency

ACM Classification Keywords
H.5.2 [Information interfaces and presentation]: User Interfaces – Input devices and strategies.
Introduction
Recently, touch panels are widely used in our living environment. Touch panels are not only easy to use, but also they can be used for general purposes since screen layout and UI components can be changed depending on the situation. However, we cannot feel much like we are touching objects directly since presentation and operation are restricted on a 2D surface. On the other hand, direct manipulation of virtual 3D objects that are stereoscopically or optically displayed in the air has been studied [1, 2, 3]. However, only visual and audio feedback is provided and users do not feel much haptic sense. Vibration devices can be used to provide tactile feedback [4], but users have to wear the devices. Non-contact tactile display using an ultrasound transmitter can also be used to provide haptic feedback [5], but it requires a large space.

In another way, the surface of the touch panel can be used to provide users with passive haptic feedback. In [6], users touch a 2D surface to select a 3D stereoscopic object and the relation between the 3D positions of rendered objects and the on-surface touch points was analyzed. In [7], users also touch a 2D surface to select a 3D stereoscopic object and depth perception of the users was evaluated. These studies show that 3D objects can be selected with passive haptic feedback by touching a 2D surface of the touch panel. However, the touched object does not move according to the finger movement and the users do not feel much like they are manipulating the object. In addition, when the depth of the rendered object is not around the display, the users feel unnatural since the depth of the finger and that of the object are different. A robotically actuated surface can solve this problem [8], but it requires mechanical hardware.

In this study, we use a system that consists of an autostereoscopic display and a high-speed stereo camera. With this system, virtual objects are stereoscopically-presented, and the objects respond to the finger movement that is obtained using the stereo camera. In this system, virtual buttons are floating a little in front of the display. First, a user’s finger touches the surface of a button and the button moves backward according to the finger position. After that, the finger touches the surface of the display and gets passive haptic feedback from the surface. The timing of haptic feedback is inconsistent with that of visual feedback, which may cause a feeling of strangeness.

We can synchronize haptic feedback and visual feedback by adding a delay to the button movement. However, the delay also causes slow response, which impairs the sense of reality in object manipulation.

In this paper we evaluate the relation between visual and haptic feedback and show the optimal latency for natural interaction.

System Overview
For the experiment, we developed a system shown in Figure 1. The system consists of an autostereoscopic display and a high-speed stereo camera. The display was an 8.4” parallax barrier autostereoscopic display from VMJ Inc., which shows relatively good 3D image quality. The minimum 3D viewing distance of the display is 0.65m, which is short enough for users to manipulate rendered 3D objects by hand. The stereo camera consists of two monochrome IEEE 1394 high-speed cameras Grasshopper from Point Grey Research Inc. with a lens having a focal length of 5 mm. We used...
the cameras at the frame rate of 120 fps (frame per second) with the image size of 640 × 480 pixels.

Since high-speed cameras have shorter latency than standard cameras, we can reduce the latency of the entire system and we can use a wide range of latencies (from short to long) for the experiment.

We created a 3D graphical user interface of a touch panel, in which users can push three buttons which are displayed stereoscopically.

**Fingertip Recognition**

The positional relationship of the cameras, the display and the finger, and the world coordinate system are shown in Figure 2.

First of all, the system binarizes the images from the cameras to obtain the finger region. We put a black board on the right side of the display, so that the system can extract the finger region easily and stably. Next, the system scans the binarized image line by line.

When a line that has five or more white pixels is first found, the system regards it as a fingertip line.

By calculating the centroid of the white pixels in the fingertip line, the system obtains the coordinates of the fingertip in the image. After that, the system obtains the Y-coordinate of the fingertip in the three-dimensional space, which is used to calculate the distance of the finger from the display $D_{\text{finger}}$, on the following equation.

\[
Y_{\text{finger}}[\text{mm}] = \frac{y_L[\text{pixel}] + y_R[\text{pixel}]}{2(x_L[\text{pixel}] - x_R[\text{pixel}])} \cdot b[\text{mm}] \tag{1}
\]

$(x_L, y_L), (x_R, y_R)$ are the positions of the fingertip in the left and right camera images. $b$ is the baseline of the stereo camera. The Z-coordinate of the fingertip in the three-dimensional space, which is used to detect which
button is pushed, is calculated on the following equation.

\[ Z[\text{mm}] = \frac{f[\text{pixel}]}{x_l[\text{pixel}] - x_r[\text{pixel}]} \cdot b[\text{mm}] \]  \hspace{1cm} (2)

where \( f \) is the focal length of the cameras.

**Latency**

Before conducting the experiment with different latencies, we measured the original latency of the system including camera input, image processing, CG rendering and display output. We created a program in which a button displayed on the screen moves vertically and horizontally according to the finger and took a video of both the finger and button movements using a high-speed camera at the frame rate of 300 fps. By counting the lag between the movements, we can obtain the latency of the system \( l_s \). We measured the latency 10 times and the mean was just 100 ms and the standard deviation was 5.16 ms.

For the experiment, we added an artificial delay to the system. In order to add a delay without reducing the output frame rate, we stored the measured finger position data to a ring buffer. Let \( t \) be the current time and \( d_{\text{add}} \) be the added delay. The data at \( t - d_{\text{add}} \) is read out from the buffer and is used to move the button. The total latency becomes \( l = l_s + d_{\text{add}} \).

**Visual Feedback**

When \( D_{\text{finger}} \) is less than the distance of the button surface (17mm), the selected button moves back and forth according to \( D_{\text{finger}} \). The button surface coincides with the display surface when the finger touches the display surface. Figure 3 shows the visual output of the system.

**Evaluation of Visuo-haptic Feedback**

In order to find the optimal latency for the 3D touch panel interface, we evaluated the visual and haptic senses with varying the latency of the system.

**Experiment 1: Evaluation of visual and haptic synchronization**

We selected 10 subjects (6 males and 4 females) aged 21-23 years old. We changed the latency to 100, 150, 200 and 250 ms (added delays were 0, 50, 100 and 150 ms respectively) in the randomized order. We asked the subjects to touch buttons 10 times for each latency. Then we asked the subjects the following questions.
Figure 4A shows the result of the order of visual and haptic senses. When the latency was 100, 150 or 200 ms, more than half subjects answered that they felt visual and haptic sensation simultaneously. In contrast, when the latency was 250 ms, most subjects felt that the button starts to move after touching the screen. From the average scores, the best latency in terms of synchronization of visual and haptic feedback was around 150 ms.

- Did you feel the haptic sensation first or did the button move backward first?
- Did you feel that the button moves along with the finger movement while the finger was pushing the button?
- Did you feel that the button moves along with the finger movement while the finger was leaving from the button?

Figure 4B and Figure 4C show the graphs of visual synchronization between finger and button movements while the finger was pushing the button and while the finger was leaving the button respectively.

From the graphs, we can see that the smaller the latency was, the more synchronized the subjects felt in both pushing and leaving cases.

Experiment 2: Evaluation of the sense of reality
From the experiment above, visual and haptic feedback was the most synchronized with the latencies around 150 ms, while finger and button movements were more synchronized with smaller latencies. We therefore conducted another experiment to explore which latency gives the greatest sense of reality to users. We chose 100, 150 and 200 ms latencies to compare. We selected 18 subjects (12 males and 6 females) aged
21-36 years old and asked the subjects to use a pair of systems with different latencies and asked them the following question.

- Which did you feel like pushing a real button more?

This was repeated for all the ordered pairs of latencies (6 pairs) in the randomized order. The data was analyzed by the Thurstone method and the result is shown in Figure 5. From the result, the best latency in terms of the sense of reality was 100 ms.

![Scores](image)

**Figure 5.** Evaluation of the sense of reality using the Thurstone method.

**Conclusion**

In this paper, we developed a system that consists of an autostereoscopic display and a high-speed stereo camera. The system allows users to push 3D virtual buttons that are floating in front of the display. When a user’s finger touches the surface of a button, the button moves backward according to the finger position and then the finger touches the surface of the display and gets haptic feedback from the surface. The timing of haptic feedback is inconsistent with that of visual feedback, but we can synchronize haptic feedback and visual feedback by adding a delay to the button movement. We conducted an experiment to evaluate visual and haptic synchronization and the result showed that visual and haptic feedback was the most synchronized with the latencies around 150 ms, while finger and button movements were more synchronized with smaller latencies. We also conducted a comparison experiment to explore which synchronization is more important and the subjects felt the greatest sense of reality with the smallest latency, which showed that the visual synchronization of finger and button movements is more important than visual and haptic synchronization.

**References**


