This is the authors' draft. The final version is pusblished in
Displays 28 (2007) 118–128

Design and implementation of a graphic-haptic display system

Farrokh Janabi-Sharifi, [a,*] Vincent Hayward, [b]
Qing Yuan Wang [c]

[a] *Department of Mechanical and Industrial Engineering, Ryerson University, Toronto, Ontario, Canada, M5B 2K3*

[b] *Center for Intelligent Machines, McGill University, Montréal, Québec, Canada, H3A 2A7*

[c] *Motorola Canada Ltd.*

Abstract

Despite current advances in multimedia environments, tracing the geometrical structures of graphical images using force feedback remains an important research issue. In this paper, the development and implementation of a Multi-Modal Display System (MMDs) for tracing 2D boundaries in graphic images are discussed. A method is proposed that provides a type of haptic feedback designed to assist a user to trace the contours of objects seen in images. This method is an example of a family of haptic synthesis methods whereby the force field explored by the user is dynamic in the sense that it depends both on movement as well as on the object being haptically represented. The proposed performance-based method provides users with a movement guidance through an active haptic sense rather than common impedance technique. The tracing effectiveness of the proposed method is verified experimentally.

Key words: haptic interface, boundary extraction, image display, user interface

1 Introduction

Real world decisions are usually made based on the combination of senses from environment. Among possible combinations, haptic and vision combination contributes to a large classes of our daily tasks. While vision is the quickest and most efficient sense for the perception of spatial events, touch provides tactile and kinesthetic information. Integrating visual and tactile information, not only provides visually absent characteristics of surfaces, but also, as psychophysical studies show, facilitates the manipulation of objects in terms of reducing cognition loads, errors, time and energy required to complete a task. In human-machine and interactive computer systems, visual and haptic information are expected to be transferred to the user through a coordinated graphic-haptic display system. A coordinated graphic-haptic display can play important roles in several application areas, such as teleoperation, virtual surgery simulation, entertainment, scientific visualization (Hollerbach, 2000), etc. In this work, an effective method for design and implementation of a haptic display for tracing contours in images of a graphic display is proposed.

Tracing contours in digital images constitutes an important class of our daily tasks. This is needed, for example, to quantitatively assess the size of structures in medical images (e.g., to monitor growth over time), or to outline their shape for registration. A good example is provided by NIH's Visible Human Project where thousands of slices were manually traced to segment tissues.

* E-mail: fsharifi@ryerson.ca. Tel: +1-416-979-5000x7097. Fax: +1-416-979-5265.

Despite assistance from machine vision techniques, most of the work was done manually (Crawfored-Hines et. al., 1998). Professional image processing tools such as Photoshop™ (for general purpose images) or Manifold™ (for domain specific images) offer automatic assistance to tracing contours, but always in conjunction with manual input. As can be readily appreciated by graphic interface users, manual tracing is straining and error prone. Therefore, there is an opportunity to take advantage of haptic feedback techniques to provide assistance to the task of tracing the contours of the image objects, to facilitate inspection and manipulation.

There have been an extensive research on haptic displays, mostly related to *physics-based* modeling and rendering of deformable objects. Examples include work of Zilles and Salisbury (1994) where a point-based rendering approach is taken to represent a god-object. Other approaches to physics-based modeling have been also reported (Srinnivasan and Basdogan, 1997). The main challenge is to achieve an optimal balance between the complexity of physical models and the realism of haptic and graphic displays in real-time (Duriez et al., 2006). There is little prior work compared to ours. Recent work is that of Fukui and Shimojo (1992) who employed an input device to investigate the human sensitivity to virtual shapes while following their contours when force feedback is available. In other words, almost no attention has been paid to *performance* of the user. They used an input device comprising a two-dimensional force sensor mounted on the head of an XY recorder. This devices was connected to a computer commanded the device to move in the direction of force applied to the user. Of particular interest here is the work of Kruze (1997) who studied methods to produce 2D "haptic images" from 2D graphics, and that of Shi and Pai (1997) who developed a system to demonstrate the automatic generation of haptic interaction with 3D objects derived from stereo 2D images.

These works approach the problem of information extraction from an image in various ways, but invariably, the haptic feedback is synthesized from a static force field which is entirely a function of the graphics under consideration. This is depicted in Figure 1. One problem with translating a picture into a static force field– for example based on the image gradient– is that the synthesized force feedback does depend on the task which is expressed by the movements of the user. That is what makes the previous work non-performance-based. In contrast, in this paper we introduce an approach whereby the force field is a function of both the image and the movements of the user. Our approach is to automatically extract features from an image and to use them to synthesize a haptic feedback only if the user's movement agrees with what is extracted automatically. This implies that the graphic-to-haptic translation step must be performed inside the interaction loop as represented by Figure 2. In other words, each movement causes the interaction loop to create a different force field. That is, the role of the machine is limited to task reinforcement. It does not force the user to a particular location and hence it is never inhibitory, i.e., it does not provide haptic guidance as described in (Faygin et al., 2002). Therefore, the resulting fields can be termed "dynamic" to distinguish them from "static" fields as in previous approaches.

The organization of the paper is as follows. In section 2, general methodology is given. Section 3 includes haptic display design with the focus on graphic-to-haptic translation methodology. Section 4 describes the experimental results, and finally, section 5, concludes the work.

2 Approach

In this paper, a multi-modal display system is developed. It consists of two subsystems: a graphic display system and a haptic display system. These two subsystems and information flow underlying interactions between a human user and the machine interface are shown in Fig 3.

Designing the proposed multi-modal interface between human and machine requires to analyze the tasks that the human is asked to perform and to predict the human performance. The framework presented in this paper is a proposal to assemble several components, including: (i) a graphic user interface for displaying the two dimensional images; (ii) localization of the interaction between a two dimensional cursor which is following the operator's movement and two dimensional images; and (iii) a haptic image display model with algorithms to calculate the reaction force, and application of the reaction force to the operator's hand in real-time. Since this paper is primarily concerned with two-dimension images, the intensity data can be directly displayed pixel by pixel on the screen. A haptic device will be used to build the communication channel between man and machine in terms of providing the position and carrying out the force command produced by the haptic rendering approach for tactual perception. Haptic devices, such as PenCAT/Pro™ (Immersion Canada Inc.), functions not only as an input device, also as an output device by means of force feedback. They complement an information flow between the human and the machine, leading to an enhanced subject's performance and reduced visual load.

While there are other interaction models (such as ray-based models of Srinnivasan and Basdogan (1997)), the point interaction paradigm greatly simplifies both algorithm and device development. Additionally, a point-based paradigm permits desirable bandwidth and force fidelity that enable a surprisingly rich range of interactions. Furthermore, such a paradigm reduces the problem of computing appropriate interaction forces to one of tracking the motion of a point among objects and generating the force components representing the interaction with these objects. The PenCAT/Pro™, used in our work, belongs to this kind of haptic interfaces. It provides the input position, and an output force vector defined by two components – the force magnitude and the force direction.

Our work is concerned with the creation of a graphic-haptic display system that interprets an image in haptic domain. Therefore, the focus of this paper will be on graphic-to-haptic translation.

Human haptics studies show that there is a strong link between the sensations experienced by a human hand, and the motions the hand performs to acquire that knowledge. Two coupled subsystems comprise the human haptic system: the motor and the sensory subsystems (Figure 3). Unlike the visual system, in haptic devices, not only the sensory system information (about the object) but also the motions used to gain that information are important. Most of haptic simulation design work is about the domain-independent considerations. Force models are often designed with reference to the physical characteristics of objects in order to replicate the haptic perception of the interaction. We call this force modelling approach as *physics-based force modelling*. In this paper, a new approach to force modelling, called *performance-based force modelling*, is proposed. We base our force modelling design on domain-dependent considerations, since the human performance has a great influence on the result of haptic perception. Specifically for tracing boundaries, a good performance is regarded as having fewer errors and quicker speed. If the user's hand is guided by a haptic perception, his or her attention can be devoted to the object features and the exploration time and error will be greatly reduced.

With the proposed approach, the conversion of graphic data into haptic signals breaks down into two sub-problems (Figure 2). The *first* is the extraction of information from graphic images of perceptual importance in the haptic domain. Image processing is required to transform the raw intensity data into a suitable feature map (e.g., image boundary map), containing the locations of various features. The *second* is the conversion of this information into haptic signals of functional importance to the targeted task. This implies that the graphic-to-haptic translation step must be performed inside the interaction loop. More specifically, in our approach, the force synthesized at each point of the interaction is in the direction of the boundary found by the machine and its intensity is determined by the degree of agreement between the trace generated by the user and this boundary. Thus, it would be required to arrange also for a motion detection subsystem to calculate the motion of the cursor point with respect to the image.

3 Haptic Display System

The main challenge in designing a haptic display system is to produce a perceptually relevant haptic output which will convey something meaningful about the input (image).

When we manually explore the shape of an object with a probe, tactual information arises from correlation between a position trajectory and a force trajectory. It can be shown that in some cases, the force direction alone can provide shape information (Morgenbesser and Srinivasan, 1996; Robles-De-La-Torre and Hayward, 2000, 2001; Hayward and Li, 2003). The hypothesis used here is that the haptic perception of a boundary results, among other possibilities, from continuous changes in the force direction, in addition to the path traced by the probe. This also agrees with what is observed when people explore the shape of objects (Lederman and Klatzky, 1996).

Here, we focus on low-level vision-haptic translation. By analogy to low level visual image features (Marr, 1980), the following definitions are suggested that pertain to low level haptic features.

Definition 1 *A haptic edge is a local change in the force magnitude when the edge is crossed.*

Like in the image domain, in the haptic domain a complete boundary is formed by linking a set of edges to give rise to outlines or contours of the structures under consideration.

Definition 2 *A haptic boundary is such that when traced there is no change in force magnitude, but a continuous direction change. When moving away from the boundary, the force vanishes.*

These definitions are each associated with a different type of movement: "crossing" and "tracing", each corresponding to elementary image features: edges and boundaries.

3.1 Haptic Synthesis For Tracing

As outlined in Figure 2, the synthesis of dynamic force fields involves three processing steps. A suitable system specialized to contour tracing is represented in Figure 4 and consists of boundary extraction, motion detection, and graphic-haptic translation. These steps are described in the following subsections.

3.2 Boundary Extraction

The application required the boundary extraction process to be robust to noise, to be efficient, and to estimate accurately the orientation of the boundary. These requirements are adequately fulfilled by a three step process depicted in Figure 5. The location and orientation of each edge is detected and then grouped to form a boundary map. A suitable edge detection was selected to be robust to noise, to posses efficiency, and to yield high accuracy. The edge detection process involved first extracting the regions of the strongest edges which were the regions where the signal-to-noise ratio was high. Edge orientation was then estimated in these regions only and as accurately as possible. The others were discarded. Grouping was accomplished by a local weighted average of the remaining edges which further contributed to attenuate the noise.

A standard Sobel operator was first applied to the entire image to provide a gradient magnitude map. This is used to isolate out the regions of highest gradient. These high gradient regions are then reprocessed using convolution mask optimized for orientation accuracy (Lyvers and Mitchell, 1998). Let $I_{i,j}$

denote the intensity of each pixel, $\boldsymbol{S}^x$ and $\boldsymbol{S}^y$ be the horizontal and vertical masks, and $e^x_{i,j}$ and $e^y_{i,j}$, the horizontal and vertical responses, respectively. Then,

$$\boldsymbol{S}^x = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix}, \boldsymbol{S}^y = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix}, \tag{1}$$

$$e^x_{i,j} = \boldsymbol{S}^x \otimes I_{i,j}, \quad e^y_{i,j} = \boldsymbol{S}^y \otimes I_{i,j}. \tag{2}$$

The magnitude $|e_{i,j}|$ of the image gradient was computed as follows:

$$|\mathbf{e}_{i,j}| = \left[{e^x_{i,j}}^2 + {e^y_{i,j}}^2\right]^{1/2}. \tag{3}$$

In order to filter out the regions of low gradient, the image was segmented by maximizing the separatability of two classes of gradient levels using Otsu's method (Otsu, 1979). Given N pixels, the probability distribution of any gradient level (n_h) was $P_h = n_h/N$. Given L the maximum gradient value, the probability of occurrence of each gradient level k was $\omega_h(k) = \sum_{h=1}^{k} P_h$, $k = 1, \ldots, L$. The class mean level and the between-class variance were $\mu_T(k) = \sum_{i=1}^{L}(hP_h)$ and $\sigma_B^2(k) = [\mu_T\omega(k) - \mu(k)]^2/\omega(k)[1 - \omega(k)]$, respectively. The threshold K which maximized $\sigma_B(k)$ was then selected in order to produce a binary image.

Edge orientation estimated from the Sobel operator response is not accurate for angles within arctan(1/3) and $\pi/4$ in each quadrant (Lyvers and Mitchell, 1998). Within the high gradient regions, if the angle $\angle e_{i,j} = \tan^{-1}[e^y_{i,j}/e^x_{i,j}]$ was found to lie within arctan(1/3) and $\pi/4$, the edges were reprocessed using the convolution masks optimized for orientation accuracy in this range. Such masks were proposed by Climent et al. (1998):

$$C^x = \begin{bmatrix} -123 & 25 & 74 \\ -254 & 0 & 254 \\ -74 & -25 & 123 \end{bmatrix}, \qquad C^y = \begin{bmatrix} -74 & -254 & -123 \\ -25 & 0 & 25 \\ 123 & 254 & 74 \end{bmatrix}. \tag{4}$$

The difficult problem of grouping edge elements into a boundary can be simplified for haptic synthesis purposes. The output edge map had the vector quantity $\mathbf{e}_{i,j}$ associated with each image point. This map was sufficiently dense to be used directly as a boundary map, provided that an averaging process, further discussed in subsection 3.4 was applied.

3.3 Motion Detection

During interaction, the user moves the haptic device handle, causing the cursor to move along a trajectory $\boldsymbol{p}(t)$. For graphic-haptic translation, described in the next subsection, it is required to determine the motion direction. Given an infinitesimal movement bringing the cursor from point $\boldsymbol{p}$ to $(\boldsymbol{p} + d\boldsymbol{p})$, with $d\boldsymbol{p} = [dx, dy]^\top$, the movement direction, if it exists, was described by the unit vector $\delta = [dx/|d\boldsymbol{p}|, dy/|d\boldsymbol{p}|]^\top$. In practice, δ has to be estimated from discrete and noisy measurements of $\boldsymbol{p}$ noted $\bar{\boldsymbol{p}}_k$ obtained at discrete updates.

A robust technique to estimate δ from noisy discrete measurements was provided in (Hayward and Armstong, 2000). The technique is such that δ is also defined when $\boldsymbol{p}$ is stationary, i.e., when $|d\boldsymbol{p}| = 0$ or when it is very small. Compute the location of a point $\boldsymbol{w}_k$ from a measurement $\bar{\boldsymbol{p}}_k$ as:

$$\boldsymbol{w}_k = \begin{cases} \bar{\boldsymbol{p}}_k - \frac{\bar{\boldsymbol{p}}_k - \boldsymbol{w}_{k-1}}{|\bar{\boldsymbol{p}}_k - \boldsymbol{w}_{k-1}|}\epsilon, & \text{if } |\bar{\boldsymbol{p}}_k - \boldsymbol{w}_{k-1}| > \epsilon, \\ \boldsymbol{w}_{k-1}, & \text{otherwise.} \end{cases} \tag{5}$$

where ϵ is set according to the resolution of a particular device. A robust estimate of δ is found from scaling the quantity $\bar{\boldsymbol{p}}_k - \boldsymbol{w}_k$ to one.

$$\widehat{\delta_k}(\bar{\boldsymbol{p}}_k) = \frac{1}{\epsilon}(\bar{\boldsymbol{p}}_k - \boldsymbol{w}_k). \tag{6}$$

This neither involves division nor differences of measurements close to resolution, and hence is numerically robust.

For clarity, the discrete time index k will be dropped in the rest of the paper.

3.4 Graphic-Haptic Translation

As a basic graphic-haptic translation principle, all the image features (in this case boundaries) will be transferred to their corresponding haptic counterparts. The boundaries traced interactively are based on the user's judgment while moving a cursor over the image. Our system provides reinforcement in the form of a force response to movement. The approach was to provide such force if the movement agreed with tracing the estimated boundary, that is, in a direction tangent to it. The response was the strongest when the cursor moved exactly on the boundary and vanished elsewhere. At the limit, crossing a boundary at a right angle should not create any response. The force magnitude and the force direction had to be synthesized independently.

Recall that during preprocessing, boundaries were not explicitly computed but represented by a density of edge vectors magnitudes. What was required was an operator that not only responded to proximity to a boundary but also to its strength resulting from the mutual alignments of individual edges.

Neurological evidence shows that although the visual acuity is predominant in the foveal region, peripheral vision is known to play a major role for responding to patterns or textured stimuli which cover a large region of the visual field (Kertesz and Hampton, 1981). By analogy to the characteristic of the spatial visual acuity distribution in the retina (Kertesz and Hampton, 1981), this can be expressed by first defining a weighting function $n(r)$ in a neighborhood of radius m centered at the cursor location $\boldsymbol{p}$ and by computing a weighted average of the relative contributions of edge magnitudes as a function of their distance r to cursor location $\boldsymbol{p}$. A simple piece-wise linear function as in Figure 6 was selected.

A boundary map was computed in preprocessing as a weighted average of all the edges associated to the neighborhoods centered at $\boldsymbol{p}_{i,j}$ and containing the pixels $l = 0, \cdots, q$:

$$\boldsymbol{b}_{i,j} = \sum_{l=0}^{q} n(r_l)\boldsymbol{e}_l \tag{7}$$

In order to detect misalignments between the direction of motion and the boundary, the final response force model had to resemble the inner product between the cursor motion direction δ and the average edge direction $\boldsymbol{b}$. When the cursor position $\boldsymbol{p}$ was in pixel i, j, the response force would be calculated from:

$$\boldsymbol{f}_{i,j} = \begin{cases} |\boldsymbol{f}_{i,j}| = F_m |\delta(\boldsymbol{p}) \cdot \boldsymbol{b}_{i,j}| \\ \\ \angle \boldsymbol{f}_{i,j} = \operatorname{sgn}(\delta(\boldsymbol{p}) \cdot \boldsymbol{b}_{i,j}) \angle \boldsymbol{b}_{i,j} \end{cases} \tag{8}$$

where F_m was a force intensity scaling factor.

In other words, if there was a strong edge, and if the user's trace was on top of the boundary, the force had a strong magnitude tangentially to the boundary, and the user's hand was pushed along the trace whether it was made clockwise or counterclockwise. However, crossing a boundary or stopping on it produced no force, thus never inhibiting the user's intentional movements.

This force model depended on two parameters: the force intensity F_m and neighborhood size m. The following considerations were made to adjust them in the absence of systematic guidelines. The force intensity should allow a user to experience a boundary and in accordance to its visual appearance. It had to be strong enough to be felt, yet sufficiently weak so as to be annoying. The boundary region size needed to be large enough to yield a smooth force change while being small enough to localize a boundary accurately.

4 Pilot Study

A pilot study was conducted to investigate the effects of the described dynamic force field on tracing performance. A few experiments were arranged. The purposes of these experiments were to show that most people tended to perform and learn better with the proposed display system than without it.

The bi-modal graphic-haptic display system consisted of a PC (PII) with a 19” monitor to graphically display the images, a PenCat/Pro™ (Immersion Canada Inc.) as the the haptic device connected to PC, and software for processing both image and haptic operations. The TouchWare platform (of PenCat/Pro™) was used to implement both visual and haptic operations. The haptic device had a pen-like handle moving in a 14 $\times$ 10 cm work-area (Figure 7). The pen interface had two buttons for standard mouse operations under Windows environments and a contact sensor for detecting the contact with user’s hand. It returned forces up to 5 N in a plane. The haptic device provided 3600 dpi position resolution. Because of its direct-drive design, it operated silently and opposed virtually no friction when unpowered.

The importance of real-time force and motion calculations led us to separate the interfaces into two sections: synchronous and asynchronous. The synchronous section processes each step in PenCAT/Pro™ dynamic force calculation. All calculations were executed within a predetermined time frame (real-time constraints). The asynchronous section defined the graphic interface spatial structure and force attributes of the image boundary, such as the location, the force intensity component and the force direction. When the application was opened, the elements of the synchronous section were dynamically constructed based on their attributes defined in the asynchronous section image boundary. The two sections communicated at each startup. The synchronous section continuously transmitted the position of the device cursor and encountered boundaries, calculated the cursor movement and returned force data. In turn, the asynchronous section prompted the synchronous section for structural changes. For example, changing the force parameter defined in the asynchronous section also resulted in the output force change.

The dynamic force calculation was performed at a rate of 400 Hz in a hard real-time loop. A 19" Viewsonic™ computer monitor with resolution of 1600 ×1024 was used as visual display. A graphical user interface (GUI) (Figure 7b) enabled the parameters to be set for optimal comfort for each individual subject and to view gray scale images. The software processed two-dimensional images as previously described and exported a boundary image. The boundary image, however, was not shown on the GUI. The workspace of PenCat/Pro™ was in proportion to the GUI area. An environment similar to MS Windowsraisebox^R was generated through TouchWare. At the right bottom corner of the window, the cursor movement trajectory was recorded and was used for tracing boundary performance analysis. A pop-up menu bar appeared on the top of the desktop. Each menu had a header and menu items. For example, the force parameter could be adjusted dynamically by either the experimenter or the user, depending on the needs of a particular experiment.

4.2 Subjects

Twenty unpaid volunteers, all happened to be male, from students population in an age range of 21-30 participated in the experiments. All happened to be right handed. All the participants were healthy with no hand or arm motion disorder. All had some experience with a computer mouse, but none with a haptic device. Participants were randomly divided into two even groups.

4.3 Procedure

The subjects of group 1 were first exposed to only graphic display without any haptic feedback and subsequently to bi-modal graphic-haptic display. The order of the two displays were revered for participants of group 2. Each group was first asked to perform practice trials to familiarize themselves with the interface and to adjust the parameters to their satisfaction. Lines and semi-circles at four orientations of 0°, 45°, 90°, and 135° were introduced.

After learning phase, subjects were tested on three different examples designed to present manual tracing difficulties of different kinds (Figure 8). Image a (Figure 8a) was a geometrical rectangular shape with straight edges. Image b (Figure 8b) was the capital letter H that contained many corners. Image c (Figure 8c) was a highly noisy gray-scale image resulting from an ultrasound scan.

Subjects were asked to perform the tracing task based on the viewed image and the movements of the cursor over it. They were seated in front of the haptic and graphic displays with the elbow, forearm and wrist supported on the desk as shown in Figure 7a. During the experiments, the subjects were given no feedback as to their performance although they could see the trace in a separate panel that was not overlaid over the image as shown in Figure 7b.

During trials, they were presented with the three images in a random sequence. The task was to move the cursor to a point of the visible outline and to trace it in a single closed trajectory. The trace was recorded. The time needed to produce it was measured by starting a timer at the onset of cursor movement after a cue was given by the experimenter, and stopping it when the trace was completed. Subjects were asked to accomplish the task as quickly and accurately as possible.

All the trials were performed in four conditions: with and without haptic feedback for comparison, and tracing clockwise and counter-clockwise for bias compensation. All trials were performed three times in the same conditions.

Figure 9a shows the boundary map which is hidden from the subject but translated into haptic cues during performance. Figure 9b depicts a traced made without haptic cues while Figure 9c illustrates a similar trial but with haptic cues. Although this has not been quantified, the haptically assisted trace appears to be much smoother. In fact it will be seen that it was traced faster than the unassisted trace.

At the end of each session, subjects filled a survey that contained multiple choice questions. In order to minimize the bias, subjects were not told about the existence of another session (e.g., to use graphic-haptic display for the group that first started with only image display). The survey contained subjective questions with fuzzy values related to degree of satisfaction with each display functionality. The *first* question was: "How easy was the interface to use? (i) difficult; (ii) not easy; (iii) easy; and (iv) very easy." The *second* question asked: "How was your sense of control and coordination? (i) poor; (ii) not bad; (iii) good; and (iv) excellent." The *third* question targeted mental fatigue and asked: "How was the required level of your concentration? (i) high; (ii) average; (iii) light; (iv) almost none." The *fourth* question targeted physical fatigue and asked: "How demanding was the task physically? (i) very; (ii) average; (iii) mild; (iv) almost not." The *fifth* question was: "How was your overall satisfaction with system functionality? (i) disappointed; (ii) satisfied;

(iii) very satisfied." The *final* question was given at the end of second session and had a negatively formatted question inquiring: "Would you prefer the other display? (i) Yes; (ii) No."

4.4 Results

One objective of the experiment was to quantify the effect of haptic cues on the performance in terms of speed and accuracy. Speed is captured by the mean execution time $\langle t \rangle$ required to complete a tracing task for a given image. Tracing accuracy was assessed by computing the ratio of the area between the manual trace and the nominal image boundary to the total area of the nominal boundary, i.e., $e = S'/S$ as illustrated in Figure 10. Note that error measured in this fashion yields a very small number as the surface created by the cursor deviation is typically small compared to the total surface.

The results from two groups were combined to remove the effect of order, i.e., bias for each display. Time and error mean ($\bar{t}$, $\bar{e}$) and standard deviation (σ_t, σ_e) in application of each method to individual images were calculated. The results were shown in Figures 11 and 12).

Execution time mean and standard deviation ($\bar{t}$, σ_t) were (26.12, 1.94), (35.32, 1.83), and (7.25, 1.60) sec for images a, b, and c (Figure 8), respectively, when only graphic display was used. When the proposed bimodal graphic-haptic display system was utilized, execution time mean and standard deviation were (20.2, 1.49), (27.9, 1.47), and (4.85, 1.01) sec for images a, b, and c, respectively.

Similarly, percentage error mean and standard deviation ($\bar{e}$, σ_e) were measured to be (0.647, 0.092), (0.347, 0.038), and (0.299, 0.036) for images a, b, and c (Figure 8), respectively, when only graphic display was used. When bimodal graphic-haptic display was utilized, error percentage mean and standard deviation were (0.508, 0.061), (0.294, 0.023), and (0.224, 0.022) for images a, b, and c, respectively.

The results showed a difference in means of each method when applied to the same image, with consistently graphic-haptic display indicating better performance. To investigate the reliability of the differences in means of two methods when they are applied to the same image, t-test was used. A high t-value was obtained for each of the six cases (three images and two measures of $\langle t \rangle$ and e) and hence, null hypothesis was rejected for all cases. Therefore, it was concluded that the differences in the means were reliable. For example, for investigating the difference in $\langle t \rangle$ means of both methods when applied to image a (Figure 9a), a t-value of 10.8 was obtained at the common significance level of $\alpha = 0.05$. Additionally, one might investigate the theoretical difference in means by calculating the 95% confidential interval CI. The confidence interval CI on $\langle t \rangle$ means was obtained to be [4.49 7.0585], [6.3070 8.4930], and [1.5298 3.29] for images a, b, and c, respectively. Similarly, 95% confidence interval on e means were calculated to be [0.0876 0.1904], [0.0592 0.1003], and [0.0556 0.0949] for images a, b, and c, respectively. Interestingly, when comparing two methods for the same image, the ratio of standard deviations remained below 2, allowing us to use standard t-test.

Participants also responded to the survey questions about their experience with the application. The goal was to evaluate the level of satisfaction and preferences of the users regarding the displays. For brevity, only the results related to Image c (Figure 9c) were included in this paper. Analyzing the first question (Figure 13) revealed mixed response of the users regarding the ease of the display use. About 90% of the users using bimodal display found it "easy" to use (versus almost 80% of the users utilizing graphic display who found graphic display "easy" for use). On the extreme sides, however, the users seemed to have demonstrate mixed feeling about the displays. On one hand, 15% of the graphic display users favor graphic display as "very easy to use" (versus 7.5% of bimodal display users) while on the other hand, 5% of graphic display users seem to disfavor it as "not easy to use" (versus 2.5% of bimodal display users). Therefore the results in this regard are non-conclusive. The reason for mixed reception of the users regarding the use of bimodal display was probably due to the unfamiliarity of the users in using bimodal display. In the meanwhile the subjects were quite familiar with the use of single mode graphic display and they were expected habitually to favor graphic display for the ease of use. However, when a specific question regarding the sense of control and coordination was raised (Figure 13), the users clearly favored bimodal display. It was also interesting to assess the reception of the users regarding the level of concentration required by each display. Figure 14 depicts answers regarding this assessment. The results suggested that bimodal display would reduce level of concentration required for tracing. Physical demand would be also a factor in deciding about the interface type. In particular, bimodal display imposed some (although minor) levels of force to the users hands. Figure 14 plots the survey results regarding this issue. As shown in Figure 14, interestingly, the users perceived almost the same level of physical demand in using both displays. A general question regarding the users satisfaction with the displays functionality (Figure 15) revealed that users showed better level of acceptance of the received functionality in using bimodal display. In order to establish the preferences of users in using both displays, negative question was raised. Subjects who used the image display last, were asked whether they would have preferred to receive the additional force feedback. Likewise, the users who utilized graphic-haptic display last, were asked whether they would have preferred not to receive additional force mode. Figure 15 summarizes the answers to this question. In both groups, a clear majority of the users favored bimodal display.

4.5 Discussion

Manual tasks usually involve a speed-accuracy tradeoff, a longer time is associated to a better precision. The results in Figures 11 and 12 indicate that, despite the fact that error is not significantly different between the two tracing conditions, the execution time is markedly reduced with the addition of haptic cues. Also, it is noted that the standard deviation is consistently lower in using proposed graphic-haptic display than solely image display. As already mentioned, the haptic assistance appears to reduce hesitation and hence the many small movements seen during purely visual tracing tasks. This results in a smoother and faster trace.

The values of the parameters corresponding to the various images were adjusted by the users in the pre-experiment phase. The resulting parameter selections reveal differences between subjects which is an important consideration in haptic interface design. However it is too early to be in a position to identify the relative importance of the many factors which may contribute to the performance.

The survey results showed that, from the users perception, the bimodal display would enhance the required level of control and coordination, and reduce the concentration demand, without imposing a noticeable physical burden on the user. Furthermore, the results revealed a superior level of satisfaction with the bimodal display functionality.

5 Conclusions and Future Work

Traditional methods for inspecting images rely almost entirely on our powerful sense of vision to convey information. In order to compensate for shortcomings of mono-modal visual inspection tasks, a multi-modal graphic-haptic display system was designed. The proposed system consisted of a graphic display and a haptic display.

As the core of our haptic display system, a performance-based approach was proposed for graphic-haptic translation. Two basic image features, edge and boundary, were identified as features that should contribute to the synthesis of haptic cues designed to assist the performance of contour tracing tasks. This led to a force model that generates a dynamic force field, that is, a field that depends on the user's movements as distinct from a static field used in previous approaches to haptic cue synthesis. A slightly reduced standard of deviation in all experiments when using bimodal display supports the aforementioned suggestion.

Despite the limited scope of the study, experiments indicate that the proposed force model on average measurably increases the tracing performance of users in terms of the speed/accuracy tradeoff. Also, the proposed integration would help the users to decrease their visual concentration and cognitive workload.

The proposed system can be improved in several directions. There are a few existing sources of difficulty including sampling errors, delays due to computation or data transmission and haptic interface limitations. A revisiting engineering effort would improve transparency of forces reflected. Improved performance could be also achieved by incorporating several different operator fields and by combining them in a coherent way to access fine details missing from current implementation. Moreover, smoothing filters are necessary to be applied to the image before it is sampled to prevent aliasing of high frequency components in the sampled signal. Our study also indicated difficulties in haptic spatial processing of geometric properties and identification of objects via a single-point haptic interface. Further effort is required to make haptic devices become more effective by incorporating additional tactile feedback to match the sensory and motor capabilities of the human haptic system, thus promoting further the tactual fidelity. Finally, using the psychological data in the design process is well known difficulty. Our haptic boundary model, of course, remains restricted by diverse psychological factors such as gender, personality, emotional experience with devices, and the difficulty of casting this diverse knowledge in a unified framework. With some effort, it is possible to gradually replace present formulations for the model of human information processing with more sophisticated versions that capture in greater subtleties larger numbers of psychological results.

Acknowledgment

This work was supported by the Natural Sciences and Engineering Research Council of Canada through Research Grants # 203060-98 and # 203060-02.

References

Basdogan, C., Ho, C., and Srinivasan, M.A., 1997. A ray-based haptic rendering technique for displaying shape and texture of 3-d objects in virtual environments. In: Proc. Conf. ASME Winter Annual Meeting, Dallas, TX, Dsc-vol. 66, pp. 77–84.

Climent, J., Grau, A., Aranda, J., and Martinez, A.B., 1998. A high precision operator to determine edge orientation. In: Proc. UKACC 1998 Int. Conf. Control, no. 455, Swansea, UK, pp. 95–99.

Crawford-Hines, S., Anderson, C., and McCracken, T., 1998. Machine-learned assist for boundary controur tracing. In: Proc. Second Visible Human Project Conf., R.A. Banvard, F. Pincirolo, P. Cercevi (Eds.), Bethesda, MD, National Library of Medicine, available online at http://www.nlm.nih.gov/research/visible/vhpconf98/MAIN.HTM.

Duriez, C., Dubois, F., Kheddar, A., and Andriot, C., 2006. "Realistic haptic rendering of interacting deformable objects in virtual environments. IEEE Trans. Visualization and Computer Graphics 12(1), 36–47.

Feygin, D., Keehner, M., and Tendick, F., 2002. Haptic guidance: Experimental evaluation of a haptic training method for a perceptual motor skill. In: Proc. Int. Symp. Haptic Interfaces for Virtual Environments and Teleoperator Systems, Orlando, FL, pp. 40–47.

Fukui, Y., and Shimojo, M., 1992. Edge tracing of virtual shape using input device with force feedback. Systems and Computers in Japan 23(5), 1052–1059.

Hayward V., and Armstrong, B., 2000. A new computational model of friction applied to haptic rendering. In: Experimental Robotics VI, Lecture Notes in Control and Information Sciences, vol. 250, Corke, P.I., and Trevelyan,

J. (Eds.), Springer-Verlag,New York, pp. 403–412.

Hayward, V., and Yi, D., 2003. Change of height: An approach to the haptic display of shape and texture without surface normal. In: Experimental Robotics VIII, Springer Tracts in Advanced Robotics, Siciliano, B. and Dario, P., (Eds.), Springer-Verlag, New York, pp. 570–579.

Hollerbach, J.M., 2000. Some current issues in haptics research. In: Proc. IEEE Int. Conf. Robot. Automat., San Francisco, CA, pp. 757–762.

Kretesz, A.E., and Hampton, D.R., 1981. Fusional response to extrafoveal stimulation. Investigative Opthalmology and Visual Science, 21 600–605.

Kurze, M., 1997. Rendering drawings for interactuve haptic perception, In: Proc. ACM CHI Conf.: CHI'97, Atlanta, GA, ACM Press, pp. 423–430.

Lederman, S.J., and Klatzky, R.L., 1996. Action for perception: Manual exploratory movements for haptically processing objects and their features. In Wing, A., Haggard, P., Flanagan, R. (Eds.), Hand and Brain: Neurophysiology and Psychology of Hand, San Diego: Academic, pp. 431–446.

Lyvers, E.P., and Mitchell, O.R., 1998. Precision edge contrast and orientation estimation. Trans. Pattern Analysis and Machine Intelligence 10(6) 927–937.

Marr, D., 1980. Vision. W.H. Freeman and Company.

Morgenbesser, H.B., and Srinivasan, M.A., 1996. Force shading for haptic shape perception. In: Proc. Fifth Annual Symp. on Haptic Interfaces for Virtual Environment and Teleoperator Systems, ASME Dyn. Syst. and Cont. Div., DSC-vol. 58, pp. 407–412.

Otsu, N., 1979. A threshold selection method from gray level histograms. IEEE Trans. Systems, Man, and Cybernetics 1(1) 62–66.

Robles-De-La-Torre, G. and Hayward, V., 2000. Virtual surfaces and haptic shape perception. In: Proc. Haptic Interfaces for Virtual Environment and Teleoperator Systems Symposium, ASME Dyn. Syst. and Cont. Div., DSC-vol. DSC-69-2, pp. 1081–1087.

Robles-De-La-Torre, G. and Hayward, V., 2001. Force can overcome object geometry in the perception of shape through active touch. Nature 412 445–448.

Shi, Y., and Pai, D.K., 1997. Haptic display of visual images. In: Proc. Virtual Reality Annual International Symposium: VRAIS'97, Albuquerque, NM, pp. 188–191.

Srinnivasan, M.A., and Basdogan, C., 1997. Haptics in virtual environments: taxonomy, research results, and challenges. Computers and Graphics, 12(4) 393–404.

Zilles, C.B., and Salisbury, J.K., 1994. A constraint-based god-object method for haptic display. In: Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems, Pittsburg, PA, pp. 146–151.

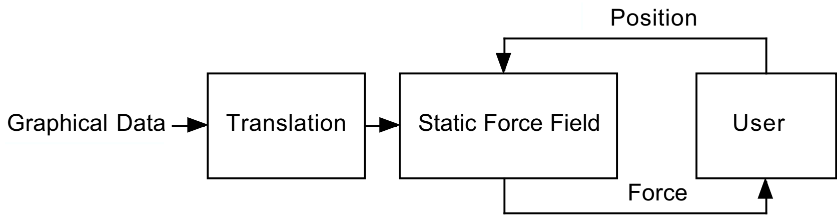


Fig. 1. Block diagram of the standard haptic interaction.

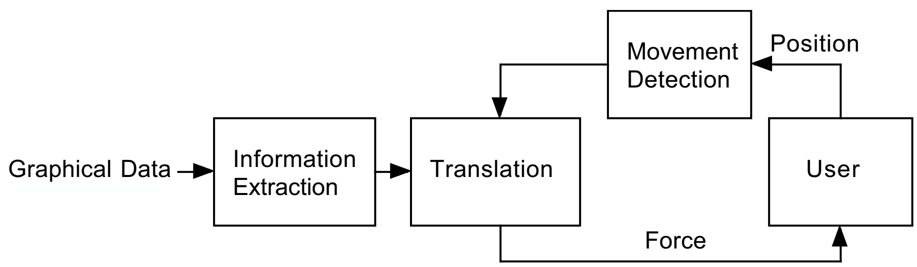


Fig. 2. Block diagram of dynamic haptic interaction.

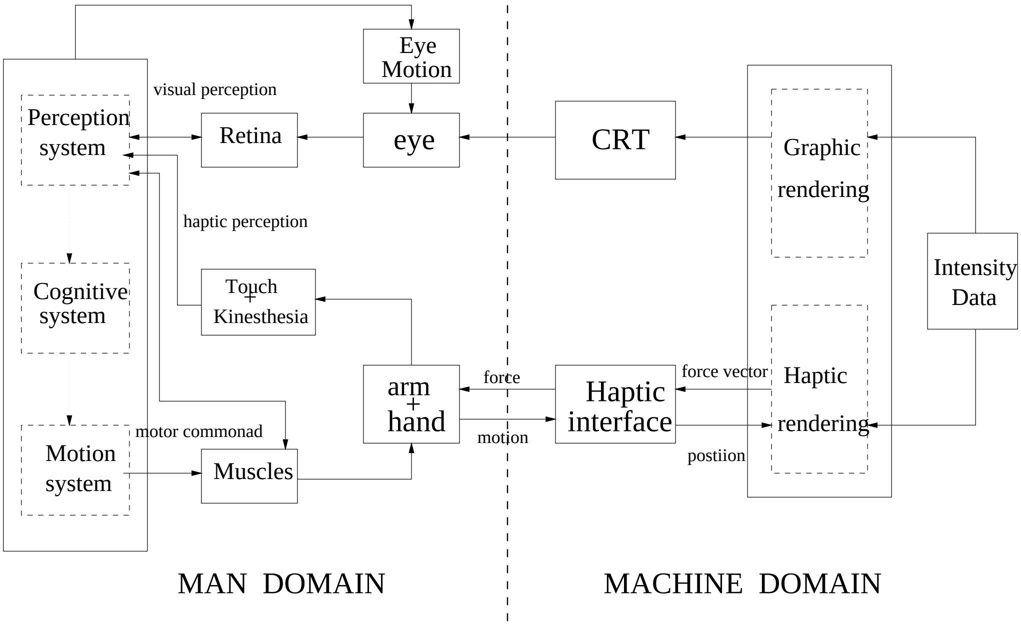


Fig. 3. Multi-modality interaction between humans and machines.

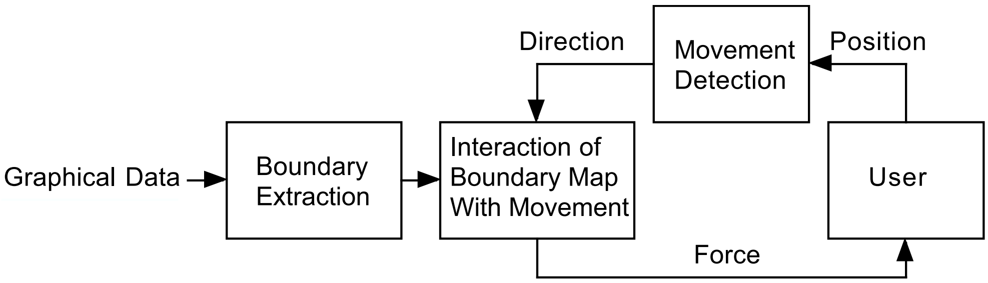


Fig. 4. Block diagram of dynamic haptic interaction for contour tracing.

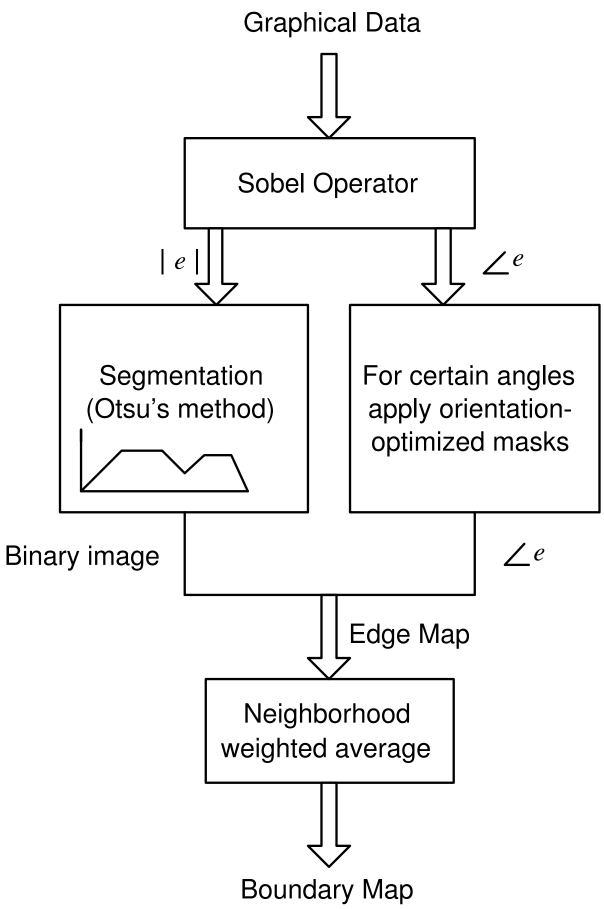


Fig. 5. Block diagram of the boundary extraction process.

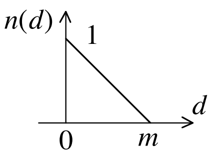


Fig. 6. Neighborhood weighting function.

a b

Fig. 7. (**a**) Subject performing task. (**b**) Graphical user interface. Top left panel: image to be traced. Top righ panel: parameter settings. Bottom panel: result of a tracing task.

a b c

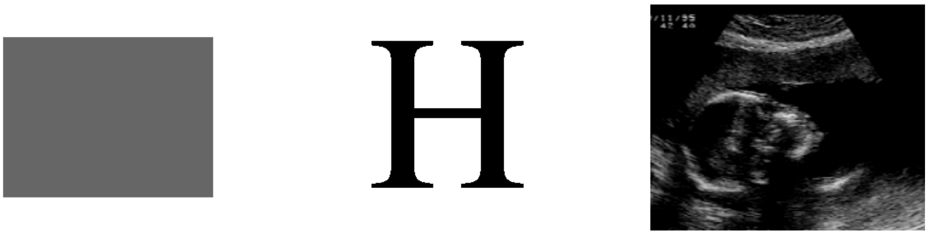


Fig. 8. (**a**) Rectangle. (**b**) Letter H. (**c**) Ultrasound scan.

a b c

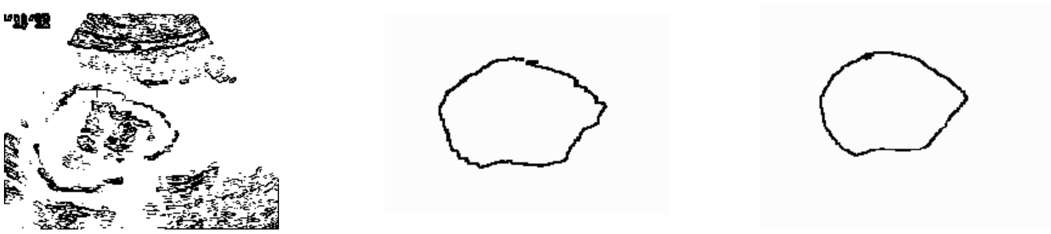


Fig. 9. (**a**) Boundary image. (**b**) Trace example without haptic cues. (**c**) With haptic cues.

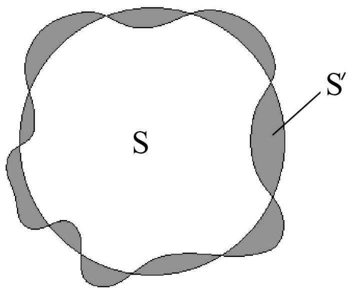


Fig. 10. Percentile error estimation $e = S'/S$.

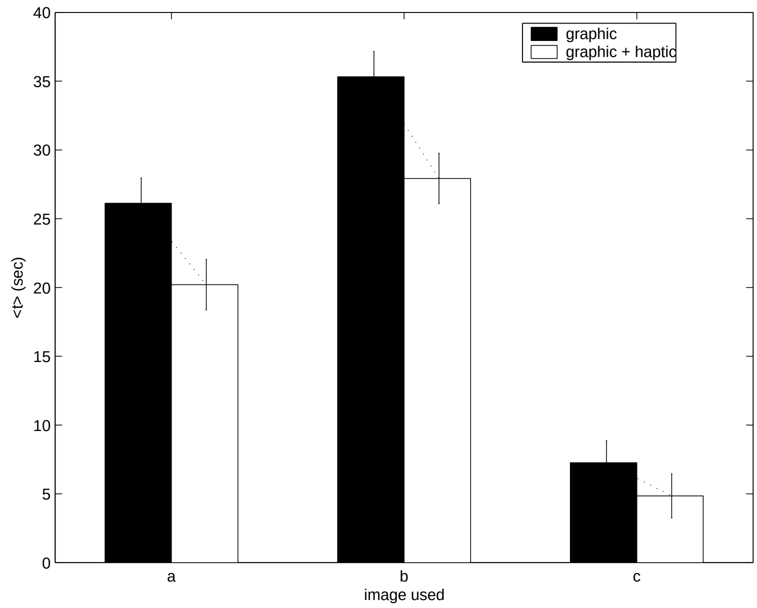


Fig. 11. Mean of execution time with standard deviation for each image.

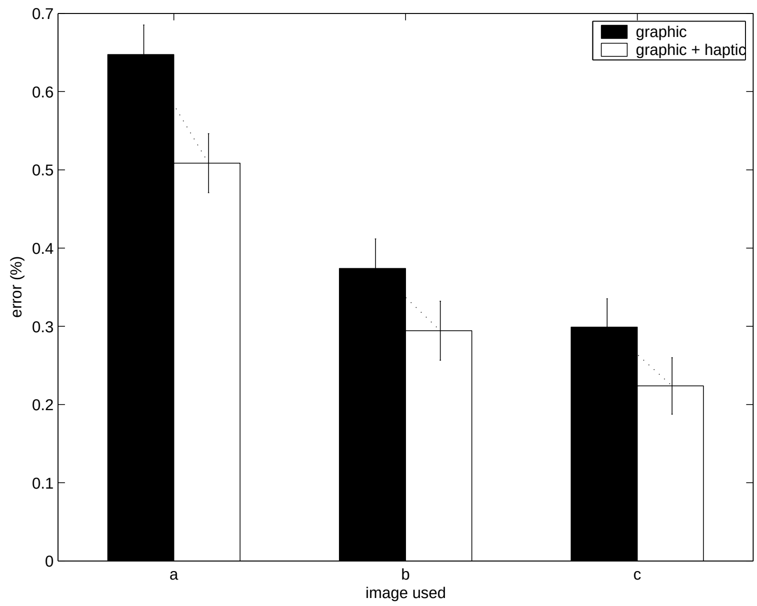
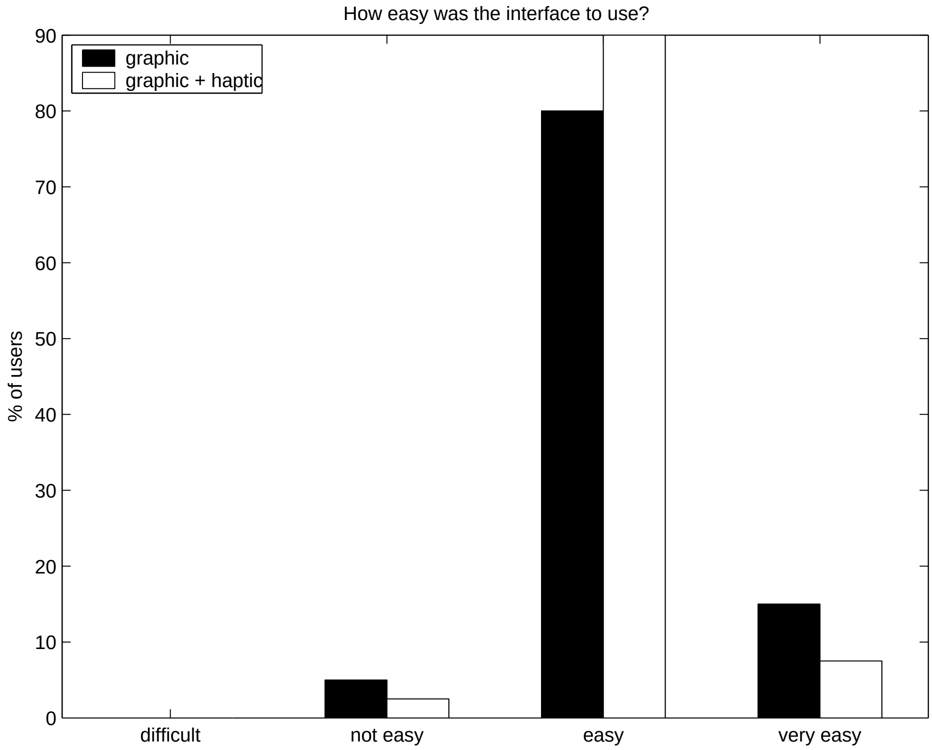
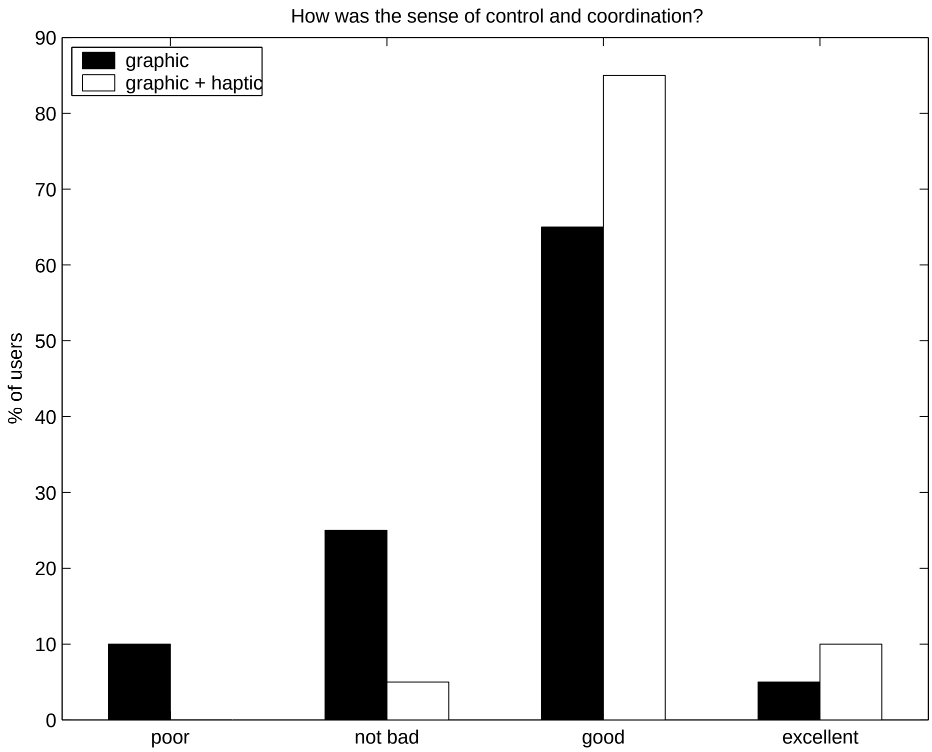


Fig. 12. Mean of percentage error with standard deviation for each image.

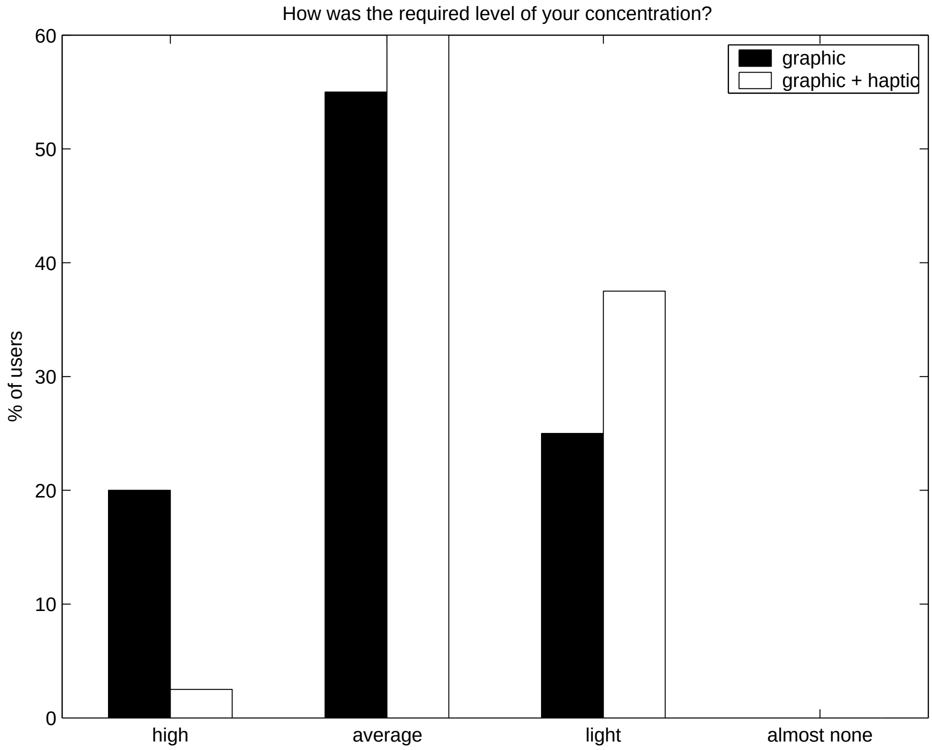


(Q1)

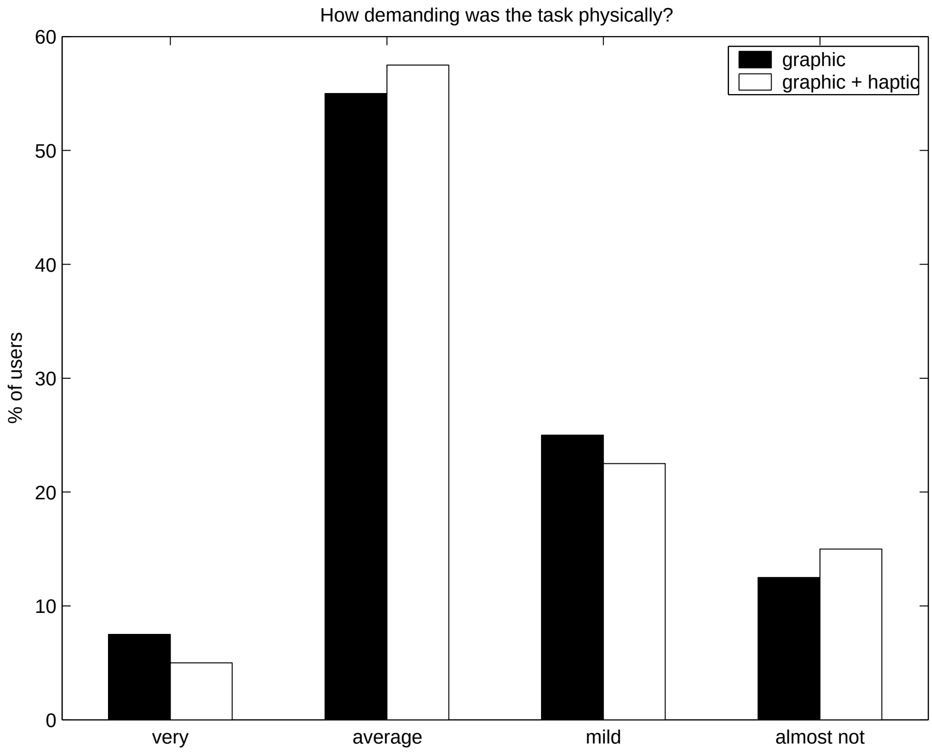


(Q2)

Fig. 13. User perception of the method applied to image 8c. (set 1)

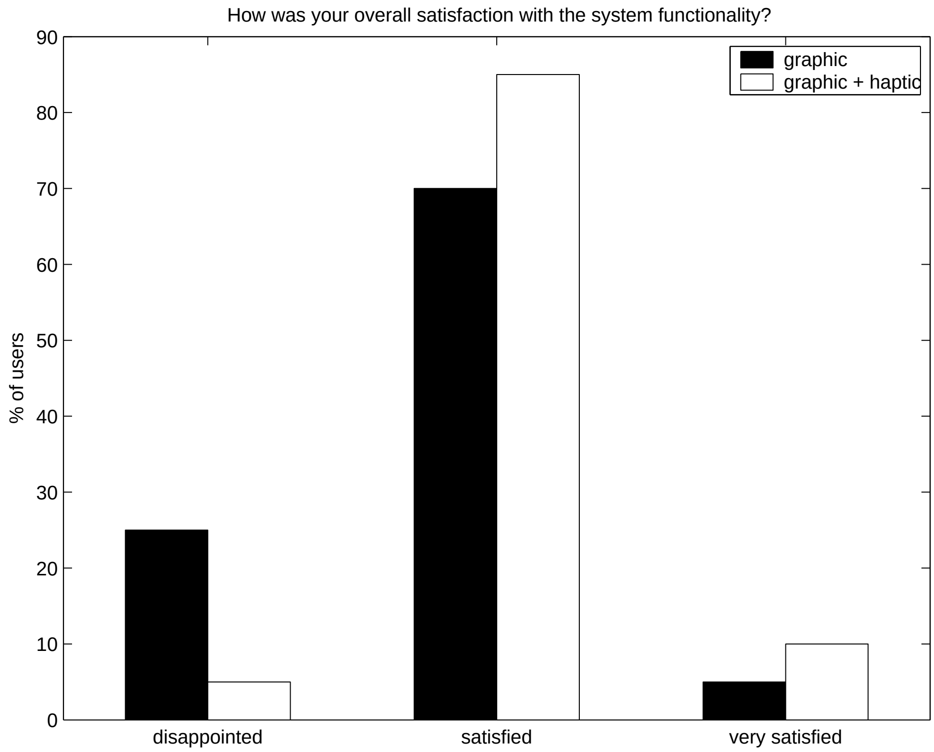


(Q3)

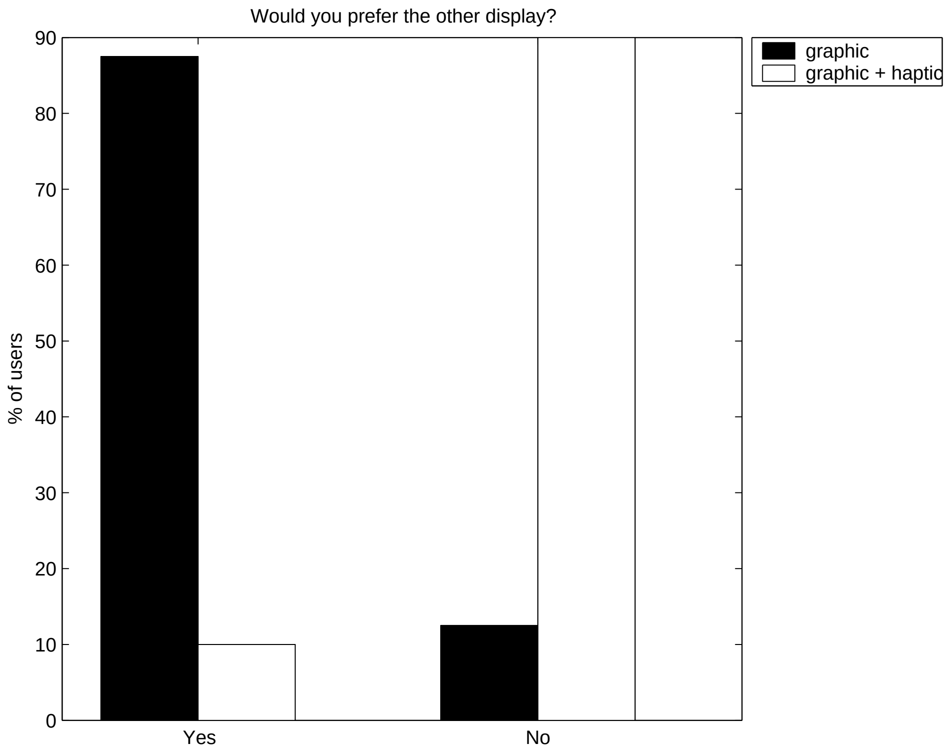


(Q4)

Fig. 14. User perception of the method applied to image 8c. (set 2)



(Q5)



(Q6)

Fig. 15. User perception of the method applied to image 8c. (set 3)