26

A MULTIMODAL USER INTERFACE SYSTEM WITH FORCE FEEDBACK AND PHYSICAL MODELS

Christophe Ramstein

Industry Canada
Centre for Information Technology Innovation
1575 Chomedey Blvd Laval, Quebec H7V2X2
Tel: (514) 973-5700
E-mail: cramstein@citi.doc.ca

KEY WORDS: Multimodal interface, force feedback device, physical modelling, sound synthesis, force synthesis, user interface development systems.

ABSTRACT: An experimental hardware and software platform aimed at enriching a graphic user interface (GUI) through the use of auditory and haptic information is described. MUIS (Multimodal User Interface System) allows for the design and creation of multimodal feedback interfaces, and includes a force-feedback device with three degrees of freedom: the Pantograph. Interface objects (windows, icons and pop-up menus) are represented in a physical structure that is well adapted to the real-time synthesis of coherent and redundant sounds and forces.

1 INTRODUCTION

As Graphical User Interfaces (GUI) are becoming the standard in human-computer interaction, it has become essential to develop alternate modes of interaction on the one hand to replace visual feedback for blind users and on the other hand to reinforce it for seeing users. With multimodal interfaces, several sensorimotor channels come into play in the communication process (Coutaz & Caellen 1990), including graphics, speech, sounds and force-feedback.

This paper presents our analysis of the design of a multimodal feedback interface (with a focus on computer outputs), through which the objects in a GUI, such as windows, icons and pop-up menus, can be perceived via the three principal channels of communication: visual, auditory and haptic (tactile and kinesthetic). We shall describe the conceptual basis as well as the hardware and software architecture of our experimental platform: the Multimodal User Interface System (MUIS).

2 PREREQUISITES

2.1 Design Principles

Since the days of Aristotle, it has been known that the senses have both common and unique characteristics. Colour, for instance, is perceived by the eyes alone, just as pitch is perceived only by the ears, whereas direction can be determined simultaneously by both. Furthermore, the various human senses are not simply equivalent ways of perceiving significant categories of events: they differ in both their precision and speed. Sight, for example, is the quickest and most efficient sense for the perception of spatial events, while hearing is best adapted to the perception of temporal events (Welch & Warren 1987). Thus, we feel that, in the interest of taking GUI technology a step further, using not one but several modalities (visual, auditory and haptic) could enhance both user performance and ease of use. According to this hypothesis, MUIS provides force and sound feedback in addition to feedback on the graphic level.

K. Nordby et al. (eds.), *Human-Computer Interaction*

However, the arbitrary association of several modalities may cause degradation in data perception (Sherrington 1920), and possibly even a transformation in the data. Therefore, a coherent model of this combination must be defined. A physical model presents the qualities required to combine modalities consistently and thus enrich and facilitate data perception. In effect, it provides a means of representing objects on the basis of physical considerations; the perceptible manifestations of the objects (through visual, auditory and haptic means) are therefore linked by causal relationships.

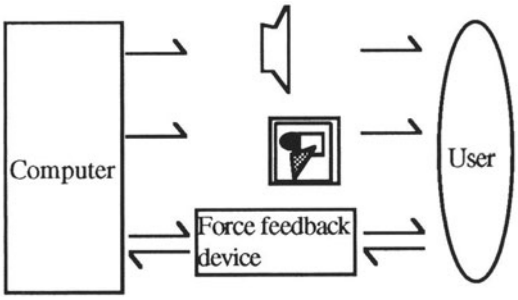

Figure 1: multimodal communication diagram.

2.2. An Haptic Device: the Pantograph

To provide the user with force feedback, a mechanical device is needed—one that not only acts as a sensor (i.e computer mouse) but that can also be used to stimulate tactile and kinesthetic senses to perceive interface objects. Since the existing force feedback devices (i.e Cadoz, Lizowski & Florens 1990; Minski, Ouh-Young, Steele & Brooks 1990) do not meet ergonomic criteria required for an iconic interface (in terms of workspace, maximum force, and precision), we will use the Pantograph, which was designed and created within the framework of this project.

The Pantograph is a force-feedback device which, when connected to a computer, provides positioning and returns computer-calculated forces to the user (Ramstein & Hayward 1994). The Pantograph has three degrees of freedom (DOF), with a basic structure of 2 DOF and a third DOF in the form of a button. The device was designed with a large bandwidth and optimized in light of considerations such as work space, inertia, response and structural properties (see Hayward, Choksi, Lanvin & Ramstein 1994 for more information), while at the same time meeting the application's ergonomic requirements. The Pantograph's basic mechanical structure (2 DOF) enables the user to move a point in a 10-cm-by-16-cm space. Conversely, two powerful, accurate motors synthesize 10 newtons pick forces. The force-feedback button provides a 1-cm range of movement with 1 newton pick forces (see Matther 1994 for more information).

3 PHYSICAL MODEL OF INTERFACE-OBJECT FORCES

Forces that are synthesized in the course of the interaction should be modelled beforehand, translated into algorithms, then calculated in real time and reconstructed for the user by means of the Pantograph. In order to facilitate this design process, MUIS offers a natural high-level language for describing interface objects themselves (shapes and parameters) instead of their force features.

3.1 Structural Analysis

Each interface object (window frames, buttons, icons and pop-up menus) is defined as a three-dimensional polygon with a certain viscosity. The pointer, controlled by the Pantograph, is comparable to a mass that is subject to a gravity field which can be moved about the visible surface of the interface objects. The force returned to the Pantograph at any given moment is the resultant of the forces present, applied to the pointer: gravity, friction and object reaction.

3.2 Basic Shapes

We have defined two types of basic polygons. The first is a rectangular enclosure with sides characterized by width L and depth H (see Figure 2). The sides exert a force on the pointer which moves it into the object, thus giving the user the haptic sensation of a real enclosure. This type of object is used for buttons, icons, headers and menu items as well as for demarcating desktop workspace.

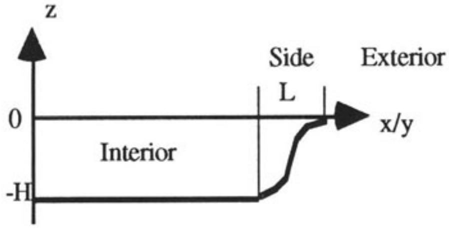

Figure 2. Position profile of a rectangular enclosure in planes 0xz and 0yz.

The second polygon represents a hollow rectangular frame (see Figure 3). The difference between this and the first is that the forces present are used to keep the pointer within the frame itself rather than within the rectangle it demarcates. It is characterized by the width of an interior edge (L_{int}), the width of an exterior edge (L_{ext}) and by depth H. This position profile is used for window frames. Though unable to see it, the user can discern the contour of the window well enough to identify and resize it.

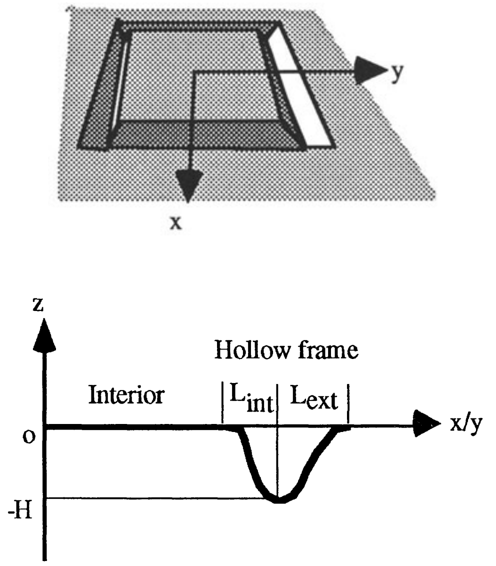

Figure 3. Position profile of a rectangular frame in planes Oxz and Oyz.

3.3 Definition of Force

The resultant force F applied to the pointer of mass m is the vector sum of gravity field G, reaction force T and friction force R. Hence, F=m.G+T+R. (see figure 4). M_n (x_n, y_n, z_n) is the position of the pointer at moment n and V_n is the pointer speed. P(x) and Q(y) are the position profiles which describe the polygon in planes Oxz and Oyz respectively. In plane Oxz, the forces are rendered by the following equations:

$$F_x(x_n) = (m \cdot G \cdot \sin(a) + R) \cdot \cos(a) \quad (1)$$

$$F_z(x_n) = (m \cdot G \cdot \sin(a) + R) \cdot \sin(a) \quad (2)$$

with a=arctan ($P'(x_n)$) and $P'(x_n)$ designates the derivative of x_n of profile position P

Similar equations are obtained for profile position Q in plane Oyz.

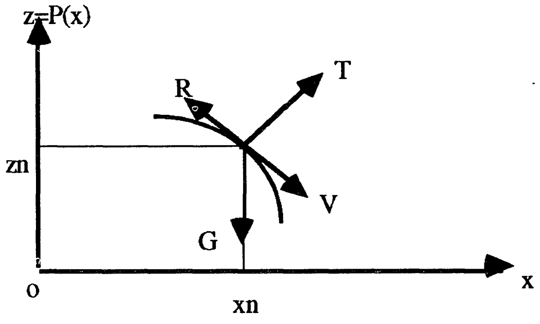

Figure 4. Forces applied to mass m at moment n in plane Oxz

Equations 1 and 2 show that the resultant force is a direct function of the P and Q derivative. Applied to a line segment (of constant derivation), the resulting force will be weak and of little relevance. To significantly enrich the fields of force, we have defined transients (L_{int} and L_{ext} in figures 2 and 3 as sinusoidal curves. Finally, to optimize the real-time calculation of these functions, we made two approximations. Force R, a function of the pointer speed ($R=-q \cdot V$), has no component in dimension 0z. Next, we approximated the force profile by segments.

3.4 Parametric Analysis

Each interface object is defined by length, depth and friction parameters. Many different values are possible for each parameter. The question to be asked is the following: What values will allow objects to be identified, without ambiguity and with the greatest ease and compatibility with the appearance of standard graphic interfaces? For example, values must be defined such that the synthesized force is strong enough to be perceived with the force feedback (the Pantograph) yet weak enough that it neither requires a major effort nor constitutes a source of fatigue and assessment error.

Sighted and non-sighted people were enlisted to participate in a pre-evaluation phase in order to determine the parametric values. This phase, described in (Ramstein & Martial & Dufresne 1994), enabled us to study the feasibility of pure-force feedback (without sound or graphics). It also showed us the importance of context (the user's mental representations) in identification tasks. In effect, haptic perception does not enhance the formation of an overall view of a spatial setting. This type of perception is sequential-perception. But a person exploring spatial information via the gestural channel overcomes this drawback by gradually constructing a mental representation of the setting which he or she then uses.

Table 1 shows the parameter values used for the various windows object types: Icon, Menu Item, Menu Header, Button (resize and close) and Move Button. The values of parameters L and H are defined with a fluctuating resolution of 640 x 480 x 100 in space 0xyz. Two mass values are specified: one in 0x and the other in 0y. Actually, perceptual haptic differences of the user's hand, according to the direction of movement, can be taken into account, and thus a more balanced perception of horizontal and vertical transients can be formed. To facilitate object identification, which becomes increasingly difficult with smaller objects, we have defined a friction force which is inversely proportional to the size of the object. Thus, the smaller the object, the greater the force working against the movement and the more easily perceptible the size and nature of the object.

Class	L	H	Mx	My	Friction
Menu Header	3	.3	2.0	1.7	0.05
Menu Item	3	.3	1.3	0.8	0.01
Icon	15	.15	1.6	2.0	0.5
Button	5	.5	1.2	1.3	1.0
Move Button	5	.5	0.9	1.2	0.7

Table 1. Parametric values of objects defined by a rectangular enclosure

Table 2 shows the parametric values retained for the window frame. A single mass value is defined for dimensions 0x and 0y. Two depth values are given: one (H_{ext}) for the exterior transient of the frame and a second (H_{int}) for the interior transient. The force synthesized in the frame is not symmetric, making it possible to identify the interior and exterior of the frame, and thus the side (top, bottom, left, right) where the cursor is located. Depths H and h depend, respectively, on the lengths L and l, to within one multiplication constant.

L_{ext}	H_{ext}	L_{int}	H_{int}	M	friction
6	0.6	3	0.3	1.5	0.3

Table 2. Parametric values of the window frame

4 PHYSICAL MODEL OF INTERFACE-OBJECT SOUNDS

From the point of view of perception, haptic stimulations have limitations (sequential-perception, not adapted for temporal events, etc.). This section describes the sound-synthesis technique used in MUIS for reinforcing those haptic stimulations seen previously.

4.1 Quick Overview

Thanks to the many sound-synthesis techniques available, iconic interface sounds are either prerecorded (sound samples) or synthesized (real-time synthesis algorithm) (Blattner, Sumikawa & Greenberg 1989; Gaver 1986; Gaver 1993). The advantages of real-time synthesis are low memory requirements and the ease with which sound parameters (volume, timbre, etc.) can be controlled. The drawback of this approach is the CPU cost, since sound samples must be calculated in real time at high sampling frequencies. With MUIS, we have experimented with a sound synthesis technique that incorporates a physical model. It has proven satisfactory for its complexity, realism and ease of parametric control (Dufresne, Martial & Ramstein 1994).

4.2 Physical Model of Vibrant Structure

Each interface object produces a characteristic sound when encountered by the pointer. What we propose here is an original approach based on the instrumental paradigm situation suggested by Cadoz (Cadoz & al 1990). From the auditory point of view, we regard the interface as a complex musical instrument composed of a stimulative structure (the pointer) and vibrating structures (interface objects). Each object category (icons, window frames, etc.) is represented by a vibrating structure. Unlike traditional synthesis techniques, the physical model approach offers an intuitive-representation framework as well as highly realistic features, while staying within reasonable CPU-cost parameters.

A vibrating structure can be seen as the combination of a set of masses interconnected by visco-elastic elements. To simplify the definition, all masses and all visco-elastic elements have the same characteristics: a mass M, an elasticity value K and a friction value Z for the connective elements. The following model was used: the model of a string (such as that of a musical instrument) comprising the juxtaposition of N masses interconnected by visco-elastic links (see figure 5).

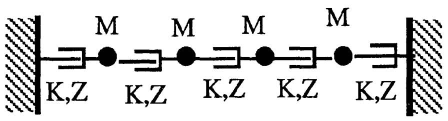

Figure 5. Physical model of the string used for sound synthesis

When the pointer interacts physically with a vibrating structure, the structure oscillates and produces a sound. This is exactly how sounds are produced on a piano: a hammer strikes a string that corresponds to a given pitch, thus making the string vibrate. The speed at which the string is struck determines volume of the sound produced. However, in our computerized context, the stimulation mode will be considered in its most basic form. When the pointer enters an interface object, the mass of the corresponding string is shifted from its initial position to a predefined position. The sound produced is therefore not dependent on the dynamics of the pointer (most notably, speed).

4.3 Physical Model Calculation Algorithm

In order to represent and calculate the preceding string model in real time, we used the modelling language defined by Cadoz & al. A brief discussion of the components follows. The physical system is made up of a set of masses interconnected by visco-elastic elements (springs). These elements are depicted as recursively defined sequences.

Expression of position

The position of each mass can be expressed as a function of its preceding positions and of the external force applied to it. According to the basic mechanical equation, the sum of vectorial forces applied to a given mass m equals the product of the acceleration a of this mass and the mass itself.

$$F = M \cdot a \qquad (3)$$

$P_n(x_n,y_n)$ is the position of one string-mass at moment n in a 2-D space (0xy). The speed of the mass is approximated by differentiating two successive positions. To simplify following equations, the sampling rate is equal to one.

$$V_n = (P_n - P_{n-1})$$

The acceleration of the mass is approximated by the differentiation of two successive speeds:

$$A_n = V_n - V_{n-1} = (P_n - 2 \cdot P_{n-1} + P_{n-2}) \quad (4)$$

At moment n, equation (3) combined with (4) becomes

$$F_n = M \cdot A_n = M \cdot (P_n - 2.P_{n-1} + P_{n-2}) \qquad (5)$$

Position Pn at moment n of each mass is expressed as a function of the two preceding positions and the sum of the applied forces. Thus, starting with (5):

$$P_n = w \cdot F_n + 2 \cdot P_{n-1} - P_{n-2} \qquad (6)$$

where $w = 1/M$

Expression of visco-elastic force

The force exerted by a spring on two masses P1 and P2 is calculated using its length. The expression of a visco-elastic force is approximated by the following equation:

$$F_n = -K \cdot (P1_n - P2_n - L0) - \zeta \cdot V_n \qquad (7)$$

where $V_n = (P1_n - P2_n - P1_{n-1} + P2_{n-1})$
K is the tautness constant of the spring
ζ is the friction constant of the spring

The force applied to mass P1 is F_n and the force applied to mass P2 is -Fn.

Computational algorithm

At each phase of sampling n, the computer calculates:

1. The force exerted by each spring as a function of its length and masses positions, using equation 7.
2. The new position of each mass, using equation 6.

When balanced, the string does not vibrate, and thus produces no sound. Whenever one of the masses is separated from its initial position and released, the system oscillates and, like its real homomorph, the resulting vibration produces an audible sound.

4.4 Physical Parameters of the Vibrating Structure

The sound phenomena produced by the preceding physical model depend entirely on the physical parameters of the string, notably:

1. the number N of masses in the model
2. the value M of each mass
3. the value K of link elasticity
4. the value ζ of link viscosity
5. the distance L0 when the masses are balanced

The definition of these parameters must include the following criteria:

1. Relative sound harmony (overall aesthetics).
2. Relative distinction: each object must be easily distinguishable from the others.
3. Rapid association of sounds and corresponding interface objects must be possible.

The parameters are defined so as to render the sounds of each object characteristic, pleasant and suggestive of the object. To highlight the state of an object (selected or not), the friction parameter is divided by a constant value. This change reduces stress on the string (and thus prolonging its life), while conserving its timbre.

5 CONCLUSION

To date, MUIS has been used exclusively for carrying out ergonomic evaluations of the impact of bimodal feedback (forces and/or sounds) on the interaction process of blind and "visually occupied" users, particularly in relation to basic graphic interface tasks: identifying, selecting, moving and resizing. These evaluations have shown the feasibility and usability of our bimodal interface for non-sighted persons, as well as the greater ease of use and improved performance resulting from bimodal feedback in comparison with monomodal situations (force only, sound only)(Dufresne, Martial & Ramstein 1994).

However, it must be understood that the observed results are not optimal and should be improved with further work on the system. Other auditory features and sound synthesis techniques may be used and evaluated in terms of their ability to reinforce haptic perception. While force feedback increases the precision of pointing tasks and brings the feeling of reality, nonspeech sounds are best adapted for representing temporal events and play a great part in the rapid completion of the tasks. It is precisely with this complementarity and redundancy that we hope to be able to offer users the ease of use and performance features that their multiple perceptual motor skills give them in physical reality, and which cannot be found with existing GUIs, which are limited simply to the use of a hand-eye interaction loop.

ACKNOWLEDGMENTS

I am grateful to Vincent Hayward for helping me to put an idea in concrete form. Likewise, I want to thank Odile Martial and Aude Dufresne.

REFERENCES

Blattner, M., Sumikawa, D & Greenberg, R. (1989) "Earcons and icons: their structure and common design principle". Human-Computer-Interaction, 4(1), pp.11-44, 1989.

Cadoz, C., Lizowski L. & Florens, J. L. (1990). “Modular Feedback Keyboard.” Computer Music Journal, 14, N2, p.47-5. MIT Press, Cambridge Mass (1990).

Coutaz, J. & Caellen J. (1990). “PRC communication homme-machine : opération de recherche concertée interface homme-machine multimodale,” June 1990.

Dufresne, A., Martial, O. & Ramstein C. (1994) “Multimodal User Interface System for Blind and "Visually Occupied" Users: ergonomic evaluation of the haptic and auditive dimensions”, In submission.

Gaver, W. (1986). "Auditory icons: using sounds in computer interfaces". Human-Computer Interaction, 2, pp.167-177.

Gaver, W. (1993) “Synthesizing Auditory Icons.” Interchi ’93, pp. 228-235, April 24-29, 1993.

Hayward, V., Choksi J., Lanvin G. & Ramstein C. (1994). “Design and Multi-Objective Optimization of a Linkage for a Haptic Device.” Proc. of the 4th Workshop on Advances in Robot Kinematics, ARK ’94. July 1994. Ljubljana, Slovenia. Published by Kluver Academic.

Matther, M. (1994) “Force feedback button for the PANTOGRAPH.” CITI research report, March 1994.

Minski, M., Ouh-Young, M., Steele, O., & Brooks, F. P. J. (1990) “Feeling and Seeing: Issue in Force Display.” Computer Graphics 24 (2), April 1990, pp. 235-243.

Ramstein, C. & Hayward, V. (1994). “The Pantograph: A Large Workspace Haptic Device for Multimodal Interaction.” In proceedings of the CHI ’94 Conference, Interactive Experience, Boston, 1994.

Ramstein, C., Martial, O. & Dufresne A. (1994). “Évaluation ergonomique d’une interface à rétroaction bimodale conçue avec le MUIS (Multimodal User Interface System).” CITI research report, July 1994.

Sherrington, C. S. (1920) “Integrative Action of the Nervous System.” New Haven: Yale University Press, 1920.

Welch, R. B & D. H. Warren (1987) “Intersensory Interactions.” In Handbook of Perception and Human Performance, Volume I, Section IV, 1987.