IPT & EGVE Workshop (2005)
R. Blach, E. Kjems (Editors)

Attribute Correlations between Haptic and Auditory Modalities

Bridget Baird Özgür Izmirli Bradford Heacock Donald Blevins
Ammerman Center for Arts & Technology, Connecticut College

Abstract
Correlations between attributes in the haptic domain and those in the auditory will assist in multimodal applications. This paper looks at the auditory attributes of pitch and reverberation and the haptic attributes of vibration, size and viscosity and seeks to correlate changes in one of the modalities with changes in the other. A total of 4 pairs of experiments were performed. The results indicate significant correlations in 3 of the pairs: pitch/vibration, pitch/viscosity, reverberation/size. There was no significant correlation found between reverberation and viscosity. These results have significance for designers of multimodal interfaces in virtual environments.

Categories and Subject Descriptors: H 1.2 [Information systems] : User/machine Systems, Human Factors

1. Introduction

Multimodal user interfaces aim at increasing the information bandwidth between the computer and the user. In the process of adding new modalities, one needs to take into consideration the degree of information represented by each modality as well as the resulting mixture of cross-modal interactions that may interfere with or reinforce each other. The bandwidth can be understood in more general terms to encompass and be influenced by concepts such as immersion, iconic retention, social loading [Mac00], preference and ergonomicity. Evidently, there are no formulas that model the effects of these concepts and each design needs to be tested in its own individual domain. Design guidelines for sensory interactions play an important role in interface design and can make designers aware of issues such as information overload within a specific modality, distribution of information across modalities or reinforcement of the message in more than one modality. Therefore, in order to further our knowledge of these design guidelines it is of interest to know which attributes of one modality correlate with some other attribute in another modality. Much research has been reported on addition, substitution or reinforcement of a second and even third modality in application specific interfaces. In these user interfaces one of the modalities is generally visual. However, little research has been done in exploring correlations among non-visual modalities. We are currently interested in exploring the relationship between haptic and auditory modalities and are conducting experiments to shed some light on users' understanding of correlations between pairs of attributes. Once a firm understanding has been established regarding the degree of correlation for each attribute pair in these two modalities, their joint relationship to the visual modality can then be explored. However, even without the visual link there are many applications that would benefit from the findings of such research, especially in cases when the visual sense is occupied elsewhere or even absent. The results could also be used in combined designs of tactons [BB04], hapticons [ME03] and earcons [BSG89] or in more general multisensory displays of abstract data [Nes04]. Nesbitt gives a hierarchical classification of abstract data displays in which the design mappings can be considered independent of modality [Nes04]. He distinguishes between spatial, temporal and direct metaphors and discusses their use in transferring mappings between modalities in the design of multi-sensory interfaces.

The relation between auditory and haptic modalities has been studied in various contexts: in multimodal graphs for blind users [YB03], in generating tightly coupled audio and haptic stimuli [DP02], in studying surface texture and touch produced sounds [LKMH02], in experimentally determining just noticeable differences due to asynchrony in haptic and auditory stimuli [ABAW03], in studying the effects of correlated cues in roughness perception [WP04] and in the design of multimodal object interaction [BI01]. In this paper, we report on work that aims to answer some of the questions on how specific attributes in these two modalities correlate. In our previous work [PBIB04] we have reported on correlations between a small set of attributes in the auditory and haptic modalities. In exploring the correlations among pairs of attributes, the most salient ones that have high differentiability need to be considered first. Therefore, in order to start with the most dominant attributes, in our previous study we used attributes that were well-understood, quantifiable and straightforward to generate. The correlations tested were auditory pitch versus haptic texture (surface roughness,) and auditory loudness versus haptic texture. These correlations were tested in both directions. The study showed that users established a strong correlation between loudness and roughness (louder correlating with rougher) and also between pitch and roughness, although not always in the same direction.

© The Eurographics Association 2005.

delivered by
EUROGRAPHICS
DIGITAL LIBRARY
www.eg.org
diglib.eg.org

The haptic attributes used in the current work were kinesthetic and vibrotactile. We selected the following three haptic attributes: viscosity of the haptic space in which the stylus moved, the size of an enclosure felt from the inside and vibration frequency. The auditory attributes were pitch (of harmonic tones) and reverberation. All attributes were nontemporal. Our hypothesis was that users would make correlations between pairs of auditory-haptic attributes.

Eight experiments in four pairs were conducted in order to test our hypothesis. The first and second experiments tested auditory pitch versus haptic vibration frequency. In experiment 1 pitch was given and the subject's task was to match vibration frequency. The order was reversed in experiment 2: given a level of vibration, pitch needed to be matched. This constituted the first pair. The second pair (experiments 3 and 4) used pitch and viscosity, the third pair used reverberation and size of the haptic space and finally the fourth pair used reverberation and viscosity (see Table 1 for a listing of the experiments). Each dependent (controlled) attribute had five levels and the subjects were forced to choose from five levels in the other modality. Pitches ranged from low to high; reverberation levels ranged from no reverberation to high reverberation; vibration levels ranged from a very low buzz (low frequency) to extreme buzzing (high frequency); viscosity levels ranged from none (air) to very dense. For the haptic characteristic of size (which could also be a visual characteristic) we chose levels which ranged from a very small enclosure (as felt by the user) to a large one. Because the auditory characteristics included pitch, we inquired, in a questionnaire, about the participant’s musical background. We also thought that some familiarity with the principles of physics might play a role in making the correlations and so we also included questions on that subject. For the correlations, we predicted that higher pitch would correspond to vibration of higher frequency, higher pitch would correlate with less viscosity, higher reverberation would correlate with larger size and lower reverberation would correlate with higher viscosity.

Although we tried to keep the attributes as basic and direct as possible it is worth noting that the pitch-size and reverberation-size experiments come close to nomic mappings as described by Gaver [Gav86]. A nomic mapping in our context means that the sound will have a dependence on the physical properties of the object. The common example of a nomic mapping is between the size of an object and the perceived pitch of a sound emitted from an object. In our case, the mapping was between the size of the enclosure and its reverberation. We predicted a positive correlation between reverberation and size. i.e. more reverberation would be matched with larger enclosure size.

In a related study, Stevens, Brennen and Parker [SBP04] reported on a recognition task using sound transformations of common auditory icons. They suggested that pitch has a nomic relationship to object size, similar to the relationship of reverberation to distance and that of volume ramping to motion direction and they hypothesized that multi-icon recognition accuracy would decrease as the parameters increased. In another study, Nicol, Brewster and Gray [NBG04] proposed a system to generate sound for audio interfaces using a timbre space approach. Müller-Tomfelde [Mül04] proposed sound feedback in a haptic environment to improve the learning process of motor skills. Sound was used to provide cues that helped position the tool tip. More specifically, beats (created by the difference of two frequencies) were used to signal proximity to a landmark. Another distance cue was based on reverberation. Chu [Chu03] found that haptics can improve the time needed to locate audio onsets in audio navigation applications. His experiments used the spectral content of a tone to render haptic feedback. In yet another study, Jeong and Gluck [JG03] incorporated haptic and auditory displays into a visual geographic information systems application and found that haptic displays performed better than auditory displays or combined displays but that nevertheless the users preferred combined displays.

2. Experimental Setup

In order to simulate the haptic properties of vibration, viscosity and size, we used a PHANToM force-feedback device from SensAble Technologies in conjunction with the GHOST software development kit. Audio tones were transmitted via closed headphones to minimize outside interference. The screen in front of the participant contained information and directions and displayed a small sphere that denoted the tip of the PHANToM stylus. This sphere moved when the participant moved the device and was present mainly for orientation. When there was a necessity to denote size, there was an invisible virtual wall that users could feel with the haptic device. Graphics were rendered in World Toolkit (Sense8). All audio and haptic attributes were rendered at five different levels.

Reverberation was generated using a synthetic model called the Schroeder reverberator [Zöl99]. The Schroeder reverberator uses comb filters and all-pass filters with a variety of delay lengths to simulate reverberation. The reverberation time was the only controlled parameter. The reverberation time is the time it takes for the impulse response to drop to 60dB below its original amplitude. The output was monophonic and delivered binaurally. We intentionally refrained from introducing 3D or stereo spatial auditory cues in order to keep the sound quality of reverberation separate from any sense of space that might be conveyed by these cues. The reverberation time was varied from 0 to 3 seconds in 5 equal steps. The sound we used was a collection of three concatenated nonsense syllables. The recordings were taken from the CUNY nonsense syllable collection [CUNY]. The syllables were extracted from their carrier phrases and concatenated, leaving brief periods of silence between the syllables. The syllables used were ('Ba', 'Eet' and 'Chi'.) The reasons for choosing nonsense syllables were both to use natural sounds and also to ensure the least iconic or symbolic association with those sounds.

Harmonic sounds were generated for the pitch attribute. Each of the five sounds had five equal-amplitude harmonics. The fundamental frequencies ranged from 100 Hz to 653 Hz, as in our previous work, but this time the range was divided into four intervals. As a result, the musical interval distance corresponded to a value slightly larger than 8 semitones. All tones were normalized to have the same relative loudness at estimated listening levels during the experiment by applying a linear approximation to values read from equal loudness curves. Loudness compensation is necessary in order to decouple and eliminate the effects of a perceived loudness change that would otherwise be experienced with changing pitch. The normalization was verified according to the ISO 532B loudness standard [ISO75] and experimentally by asking a small group of subjects. Each sound was 1.5 seconds long with 1/30th second linear fade-in and fade-out time envelopes.

Five different levels of vibration frequency were generated in the range of 20 to 60 Hz. in increments of 10 Hz. The highest vibration the PHANToM device could tolerate with stability was determined experimentally. The baseline vibration was chosen so that the subject would experience the vibration but not the movement due to the individual cycles at low frequencies.

The five levels of viscosity were obtained by creating a force opposing the direction of movement. The value of the force was obtained by multiplying the current velocity by a factor, whose range started at 0 and then changed linearly.

Different size levels were achieved by creating a box around the tip of the haptic device, centered in the haptic space. The user could touch the (smooth) surfaces of this box but not leave the box. The sizes of the boxes ranged from 1cm x 2.7cm x 1cm to 5cm x 13.5cm x 5cm in equal increments.

3. Experiments

For these experiments we used 11 subjects, all volunteers. There were 3 females and 8 males. Their ages ranged from 18 to 21. Their musical backgrounds ranged from slight familiarity to extensive knowledge. Most of them had taken some sort of physics course. Likewise their familiarity with computers and/or video games ranged from very little use to extensive. We did not find any apparent indication that there was any influence of these backgrounds in how well the subjects correlated the attributes in question.

Each participant performed four pairs of experiments. Each experiment pair began by having the participant explore the audio and the haptic attributes: the 't' key incremented the attribute (when a limit was reached the direction was reversed). Sounds could be repeated with the 'r' key. In addition to exhibiting the range of each attribute, this exploration gave the subjects a chance to become familiar with the haptic device. Once the user was finished exploring the attributes, 10 trials for each half of the experiment pair were given. In each trial, subjects were given an audio (respectively haptic) attribute and then asked to cycle a haptic (respectively audio) attribute until the best match was found. There was no time limit on how long the subject could take to choose a match. For example, a subject would be given a particular pitch and then be asked to change the vibration frequency level. The subject would then keep changing the vibration frequency until satisfied that the best choice had been made. Thus each subject performed a total of 80 trials. Changes in the audio were activated with the Q/E keys: Q moved in one direction and E in the other. Haptic changes were similarly activated by pressing the A/D keys. For each subject the direction of these keys remained consistent for the same property (e.g. pitch) but the initial mapping was randomly generated for each property and for each subject. For example, for some subjects 'Q' meant a higher pitch and 'E' a lower pitch and for others it meant the opposite. Side by side keys were used to eliminate any built-in bias toward 'up' meaning 'higher' or 'more'. Each subject performed 10 trials of one experiment and then 10 trials of the opposite match before moving on to the next experiment pair. There were four pairs of experiments in all (pitch to vibration and vibration to pitch; pitch to viscosity and viscosity to pitch; reverberation to size and size to reverberation; and reverberation to viscosity and viscosity to reverberation). The order of the pairs was randomly varied from subject to subject. Within each experiment, the level of the attribute that was initially given (one of five possibilities) was also randomly generated, while ensuring that all five levels were used twice.

At the conclusion of the four pairs of experiments the subject was asked to fill out a questionnaire. This questionnaire included questions about the musical background of the participant, familiarity with physics and basic feedback about the experiments.

4. Results

In all of the data shown below, the pitches are ranked from 1 (lowest) to 5 (highest), the reverberation levels range from 1 (none) to 5 (most), the vibration levels range from 1 (lowest frequency) to 5 (highest frequency), viscosity levels range from 1 (none) to 5 (highest), and the size levels range from 1 (smallest) to 5 (largest). Data from the eight experiments were analyzed with a nonparametric correlation test using the Goodman-Kruskal gamma statistic and are summarized below. The data were also analyzed using Kendall's τ_c statistic, which produced similar results.

The data were first treated as an aggregate collection. For three pairs of experiments there was a strong and significant correlation ($p<0.0005$), as shown in Table 1 below. As an aggregate, subjects correlated higher pitch with higher vibration frequency and lower pitch with lower vibration frequency; they correlated higher pitch with lower viscosity and lower pitch with higher viscosity; and they correlated smaller size with less reverberation and larger size with more reverberation. Subjects did not make a significant correlation between reverberation and viscosity.

Experiment	**Gamma**	**Significance**
Given pitch, match vibration	0.619	<0.0005
Given vibration, match pitch	0.464	<0.0005
Given pitch, match viscosity	-0.617	<0.0005
Given viscosity, match pitch	-0.546	<0.0005
Given reverberation, match size	0.488	<0.0005
Given size, match reverberation	0.554	<0.0005
Given reverberation, match viscosity	-0.026	0.82
Given viscosity, match reverberation	0.175	0.123

Table 1. Aggregate Correlations

The same correlation test was applied to the 11 subjects individually for each of the 8 experiments. These results corroborated the aggregate results. Subjects correlated higher pitch with higher vibration frequency, lower pitch with higher viscosity, and larger size with higher reverberation. Figures 1-3 give histograms showing the gamma values for the first three pairs of experiments. The individual results for the last pair of experiments (reverberation/viscosity) also show no consistent correlation. Figure 4 shows the gamma values for those two experiments.

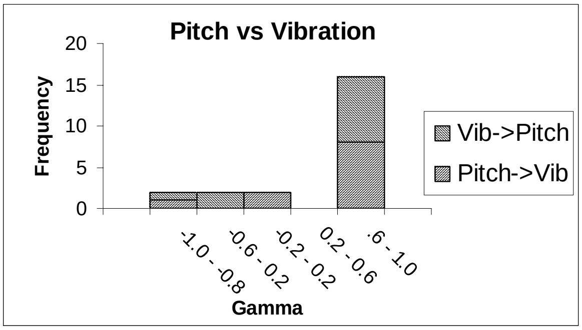

Figure 1. Pitch/Vibration Gamma Histogram

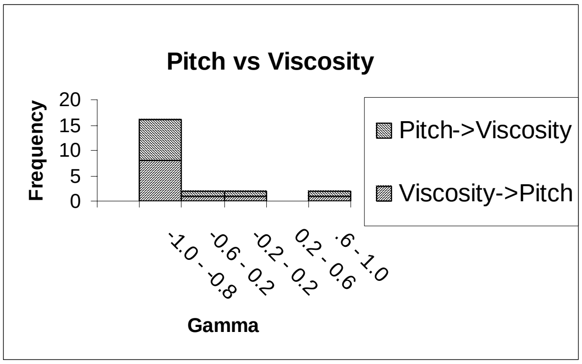

Figure 2. Pitch/Viscosity Gamma Histogram

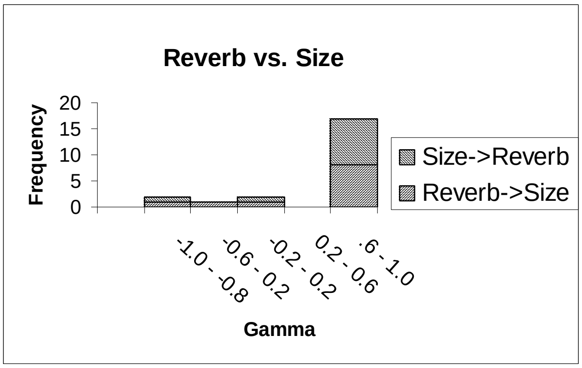

Figure 3. Reverberation/Size Gamma Histogram

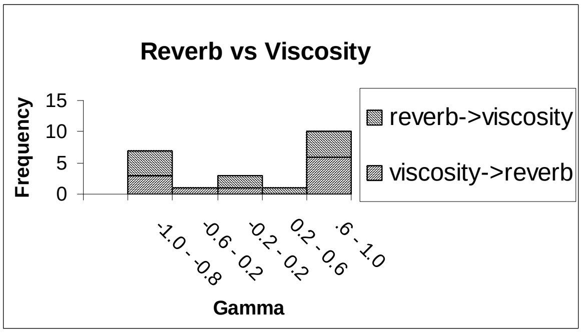

Figure 4. Reverberation/Viscosity Gamma Histogram

Using a confidence level of $p \leq .05$ we obtained the following results. For experiment 1 (given pitch, match vibration), 9 of the 11 subjects showed a significant correlation (|gamma|>.63) with 8 of the 9 correlating higher pitch with higher vibration. In experiment 2 (given vibration, match pitch), 7 of the 11 had significant correlation (|gamma|>.68) with 6 of the 7 correlating higher vibration with higher pitch. In experiment 3, in which subjects were given pitch and matched viscosity, 9 of the 11 showed significant correlation (|gamma|>.77) with 8 of the 9 matching higher pitch with lower viscosity. In experiment 4, 9 of the 11 had significant correlation (|gamma|>.61) with 8 of the 9 matching lower viscosity with higher pitch. In experiment 5, 8 of the 11 subjects had significant correlation (|gamma|>.87), with 7 of the 8 matching more reverberation to larger size. In experiment 6, 10 of the 11 subjects had significant correlation (|gamma|>.62), with 9 of the 10 matching larger size to more reverberation. In experiment 7, 8 of the 11 subjects had significant correlation (|gamma|>.82), but they were evenly divided about the direction of the correlation. And in experiment 8, 7 of the 11 subjects had significant correlation (|gamma|>.67), with 4 matching in one direction and 3 in the other. Summaries for these experiments are shown below in Table 2.

Experiment	**Subjects with p≤.05**	**Subjects with p>.05**
Given pitch, match vibration	9	2
Given vibration, match pitch	7	4
Given pitch, match viscosity	9	2
Given viscosity, match pitch	9	2
Given reverberation, match size	8	3
Given size, match reverberation	10	1
Given reverberation, match viscosity	8	3
Given viscosity, match reverberation	7	4

Table 2. Summaries of p values

5. Conclusions

In this paper, we report on correlation experiments using the auditory attributes of pitch and reverberation with the haptic attributes of viscosity, vibration frequency and size. Analysis of the aggregate and individual data have shown that subjects found significant correlations between pitch/vibration, pitch/viscosity and reverberation/size pairs. Specifically, we found that higher pitch correlated with higher vibration frequency, lower pitch with higher viscosity, and larger size with higher reverberation. For the viscosity/reverberation pair, although individual data showed significant correlations, the directions were divided and consequently the aggregate data did not show a strong correlation.

The work outlined in this paper is a second step in our effort to develop guidelines for cross-modal correlations

© The Eurographics Association 2005.

between auditory and haptic modalities. We believe that such information would be valuable for designers of multimodal interfaces and would have direct implications. Knowing the direction and the strength of the correlations may be useful for reinforcement of a visualization using these two modalities. Similarly, the results could be used to distribute information among attributes in auditory and haptic modalities in which closely correlated entities to be visualized would be mapped to correlated attributes. The results of these studies have also revealed that some attributes are uncorrelated. These results also bear important information for designers of multimodal interfaces by showing that there exist fewer restrictions in the ways these two attributes can be combined. In a larger context, the knowledge regarding the correlations between fundamental attributes in auditory and haptic modalities may be combined with the visual modality to test correlations for tri-modal stimuli. The results of this work will lead to a better understanding of the interactions between attributes in different modalities and increase our awareness of how these attributes are interrelated.

On a broader scale we anticipate using this knowledge in the field of way-finding in which three dimensional immersive virtual worlds will be created. The user will then experience multimodal cues that will aid in way-finding. A similar application is in the field of geographical information system interfaces where multilayered data make it an attractive area for multimodal interfaces.

References

[ABAW03] Adelstein, B.D., Begault, D.R., Anderson, M.R. & Wenzel, E.M. Sensitivity to haptic-audio asynchrony. In *Proceedings of the 5th International Conference on Multimodal Interfaces*, 73-76, 2003.

[BI01] Baird, B. & Izmirli, O. A tool for designing interactions with multi-modal objects. In *Proceedings of the Eighth Biennial Symposium for Arts and Technology at Connecticut College*, 1-5, 2001.

[BSG89] Blattner, M., Sumikawa, D. & Greenberg, R. Earcons and icons: their structure and common design principle. *Human Computer Interaction 4(1)* 11-44, 1989.

[BB04] Brewster, S.A. and Brown, L.M. Tactons: Structured tactile messages for non-visual information display. In *Proceedings of Australasian User Interface Conference 2004*, Dunedin, New Zealand, Australian Computer Society, 15-23, 2004.

[Chu03] Chu, L.L. Haptic interactions for audio navigation. *Ph.D. Dissertation*, Stanford University, 2003.

[CUNY] CUNY Nonsense syllable collection. url: http://web.gc.cuny.edu/Speechandhearing/labs/dsal/cuny_developed_test_materials.htm

[DP02] DiFilippo, D & Pai, D. K. Contact interaction with integrated audio and haptics. In *Proceedings of the International Conference on Auditory Display*, 2002.

[Gav86] Gaver, W. W. Auditory icons: Using sound in computer interfaces. *Human- Computer Interaction, 2*, 167-177, 1986.

[ISO75] ISO "Acoustics. Method for calculating loudness level. International Standard ISO 532", *International Organization for Standardization*, 1975.

[JG03] Jeong, W. and Gluck, M. Multimodal geographic information systems: adding haptic and auditory display. *Journal of the American Society for Information Science and Technology Vol. 54*, No. 3, 229 –242, 2003.

[LKMH02] Lederman, S. J., Klatzky, R., Morgan, C. & Hamilton, C. Integrating multimodal information about surface texture via a probe: relative contributions of haptic and touch-produced sounds. In *Proceedings of the 10th International Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, 97-104, 2002.

[Mac00] MacLean, K. E. Designing with haptic feedback. In *Proceedings of IEEE Robotics and Automation (ICRA'2000)*, San Francisco, CA, April 22-28, 2000.

[ME03] MacLean, K., Enriquez, M. Perceptual design of haptic icons. In *Proceedings of Eurohaptics*, Dublin, Ireland, 2003.

[Mül04] Müller-Tomfelde, C. Interaction sound feedback in a haptic virtual environment to improve motor skill acquisition. In *Proceedings of ICAD 04-Tenth Meeting of the International Conference on Auditory Display*, Sydney, Australia, July 6-9, 2004.

[Nes04] Nesbitt K. MS-Taxonomy: A conceptual framework for designing multi-sensory displays. In *Proceedings of the Eighth International Conference on Information Visualisation (IV'04)*, 665-670, London, England, July 14-16, 2004.

[NBG04] Nicol, C., Brewster, S.A. and Gray, P.D. A system for manipulating auditory interfaces using timbre spaces. In *Proceedings of CADUI 2004, Madeira, Portugal, ACM*, 366-379, 2004.

[PBIB04] Peeva, D., Baird, B., Izmirli, O. and Blevins D. Haptic and sound correlations: pitch, loudness and texture. In *Proceedings of the Eighth International Conference on Information Visualisation (IV'04)*, 659-664, London, England, July 14-16, 2004.

[SBP04] Stevens, C., Brennen, D. and Parker, S. Simultaneous manipulation of parameters of auditory icons to convey direction, size, and distance: Effects on recognition and interpretation. In *Proceedings of ICAD 04-Tenth Meeting of the International Conference on Auditory Display*, Sydney, Australia, July 6-9, 2004.

[WP04] Weisenberger, J.M. & Poling, G.L. Multisensory roughness perception of virtual surfaces: effects of correlated cues. In *Proceedings of the 12th International Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, 161-168, 2004.

[YB03] Yu, W. & Brewster, S.A. Evaluation of multimodal graphs for blind people. *Journal of Universal Access in the Information Society 2(2)*, 105-124, 2003.

[Zöl99] Zölzer, U. *Digital Audio Signal Processing.* John Wiley and Sons, 1999.