Multisensory Interaction: Real and Virtual

Dinesh K. Pai

Computer Science Department
Rutgers University, Piscataway, NJ
dpai@cs.rutgers.edu

Abstract. Human physical interaction is inherently multisensory, using vision, touch, hearing, and other senses. We discuss the constraints and opportunities for constructing computational models for multisensory interaction. Such models could be used to develop better robots and human interfaces, and to understand human interaction.

1 Introduction

Robots and human-computer interfaces today are in a state of sensory deprivation. This is in stark contrast with how humans interact with their every day environment, utilizing a very large number and variety of sensors even for apparently simple tasks. For instance, consider the act of putting a glass of water down on a table. Vision is used to detect a potential location for the glass, manipulation of the glass utilizes four kinds of skin mechanoreceptors [16,21], muscle spindle sensors and golgi tendon organs, and success of the task is verified by all the above sensors and the satisfying thud we hear. In this paper, we will examine how we can construct and utilize such multisensory models.

Before we begin, we briefly consider why multisensory processing is so common in human perception and so potentially useful for robots. The fundamental reason is that many real phenomena produce significant multisensory cues. This is particularly true for contact and impact between a robot and its environment, which produce forces that could be detected by a force sensor and vibrations that could be detected by accelerometers on the skin [11]. The vibrations also propagate through the air and could be detected as sound, and the forces cause the visually detected changes in motion. These measurable signals are highly correlated, and represent the same contact event. Therefore this correlation could be exploited in several ways, depending on the nature of the task. For instance, we could use it to make perception more robust; this has been the primary motivation in much of the robotics literature on sensor fusion (including several papers in this volume). The correlation could also be exploited for sensory substitution and augmentation; for instance contact pressures can be used to predict body posture [47].

Robots today have very few sensors in part because of the expense and complexity of using sensors. Some have also focused on sensorless robotics motivated by the desire to better understand and exploit important physical task constraints [6]. But the economics of sensors is improving rapidly with advances in MEMS and microelectronics. We believe multisensory systems will be far more common in the near future, and could have a significant impact on robotics and human computer interfaces. They could also help us to better understand human behavior.

2 Multisensory processing in humans

In this section we briefly review multisensory interaction in humans. Our goals are two-fold. First, even though we are far from understanding human perception, we can still get clues to how multisensory processing is organized in the human brain, which is an exceptionally sucessful perceptual system. Second, many important applications of robotics and virtual environments are for interaction with humans. It is therefore useful to know how humans process the output generated by our engineered systems.

Early work in human perception focused on the modularity of perceptual systems; however it is becoming increasingly clear that multisensory processing is more of the rule than the exception [37,46]. Large areas of the brain, known as “multimodal association areas,” are concerned with multisensory integration. Furthermore, the different sensory modalities interact even at the early stages of neural information processing in the brain stem [39,8].

Perhaps the most familiar multisensory phenomenon is the ventriloquism effect. The spatial location of a sound is perceived to be the location of a correlated visual stimulus. This phenomenon is so common that it is an integral part of the design of television sets and movie theaters.

Speech perception is also multisensory. A well known illustration of this is the McGurk Effect [24] where vision alters speech perception; for instance, the sound “ba” is perceived as “da” when viewed with the lip movements for “ga.” Notice that in this case, the percept is different from both the visual and auditory stimuli.

In haptics, perception of surface stiffness is influenced by visual [38] and auditory [5] cues. Perception of surface texture is also multisensory, utilizing haptic, visual, and auditory cues [21], though the evidence for integration is mixed in this area [10].

Early work on how multisensory cues are combined favored “dominance” or “capture” hypotheses, in particular, visual dominance over discrepant auditory or haptic cues [33]. However, it became clearer that this view is overly simplistic, and that responses depend on specific details of the stimuli and task. “Modality appropriateness” hypotheses were soon favored (e.g., [45,22]). This view has recently been placed on a more quantitative footing, with experiments indicating that multisensory cues are combined to optimally reduce the variance of an estimate [7].

Indeed, auditory cues can have suprising influences on visual perception; for instance, auditory beeps can create illusory visual flashes [36]. Determining whether two objects bounce off each other, or simply cross, is influenced by hearing a beep (or even seeing a flash) when the objects could be in contact [35]. Cross-modal influences are not symmetric; visual adaptation to 3D motion can influence auditory perception, but not *vice versa* [18], and haptic texture cues affect visual texture discrimination but not the other way around [9].

All these results indicate that the ecological validity of multisensory cues may play an important role in perception. Therefore, it is important to retain these natural relationships when developing systems for multisensory perception or display.

3 Models for multisensory interaction

What sort of computational models of the physical environment are needed for multisensory interaction? Most models in use today in robotics and graphics descend from models developed for scientific and engineering analysis. We propose that multisensory interaction has very different requirements, and therefore needs very different types of computational models.

We explore some of these differences below. In broad terms, the relevant computational models could be classified, based on their purpose, into models for *analysis* (e.g., object recognition) and models for *synthesis* (e.g., virtual environment simulation). For concreteness, in this paper we will focus on models for synthesis but many of the same considerations apply to models for analysis. The distinction between the two is further reduced in "analysis-by-synthesis" approaches. These have long been popular as models of human perception; for instance, according to the motor theory of speech, our perception is mediated by our internal models of how speech is produced [41]. There is recent evidence that perception of motion is also mediated by the motor areas of the brain [32,19,40].

We will also focus on contact models. Contact is both the central problem for robot and human interaction with the physical world and provides excellent examples of multisensory stimuli since we can perceive contact using haptics, vision, and audition.

3.1 Accuracy vs. Responsiveness

The first crucial difference between models for science and models for interaction is the relative importance of accuracy vs. responsiveness. Computations about the physical world are always approximate, since we have to work with finite dimensional approximations to reality, with uncertain parameters. In general, one could improve accuracy by constructing more detailed models and making more precise measurements, but this comes at the cost of *latency*, i.e., the elapsed time before an answer is obtained. For multisensory models we must also ensure *synchronization* of time between different sensory modalities. We group all such temporal considerations, such as latency and synchronization, in to a single category that we call *responsiveness*. How should we make this unavoidable tradeoff between accuracy and responsiveness?

We propose the following.

- For science, accuracy is the hard constraint that must be satisfied, while responsiveness is a soft constraint that we try to optimize based on available resources.
- For interaction, responsiveness is the hard constraint that must be satisfied, while accuracy is a soft constraint that we try to optimize based on available resources.

For instance, in science and engineering analysis, models of elastic deformation due to contact are concerned primarily with accurate predictions [17]. Therefore, one needs to model subtle material non-linearities, use dense meshes so that one does not miss important stress concentrations, use symplectic integrators to avoid excess numerical dissipation, etc. It does not matter much if a millisecond event takes days of simulation, but of course we look for fast algorithms such as multigrid methods and we buy fast supercomputers to minimize this latency if possible.

In contrast, for haptic interaction the primary requirement is to compute contact forces within a fixed amount of time; this available time may be determined by the haptic control loop (typically about 1 ms), or by the constraints of haptic perception (Pacinian corpuscles are most sensitive to vibrations of about 250 Hz, where they can detect motions of about 1 μm, and may be able to detect larger amplitude motions at up to 1 KHz [16][1]). Therefore, haptic interaction models need to exploit precomputation and multirate methods so that contact forces can be computed with a fixed, low latency [13–15,3,1]. It does not matter much if one uses local approximations such as point contact, but we look for better models that account for object geometry and material properties using Green's functions [14,20], and we seek to include non-linearities where possible without violating the responsiveness constraint (e.g., [29]).

The need to satisfy the responsiveness constraint fundamentally changes the types of models we should use. For instance, in science and engineering simulations, it rarely makes sense to explicitly precompute and store Green's functions, since this requires $O(n^2)$ storage for an $O(n)$ surface mesh, and the Green's functions change when boundary conditions change. But for interaction it has the tremendous advantage that in most cases the response can be computed extremely quickly by a simple memory look up and combination of a few Green's functions. This efficiency is retained even if the boundary conditions change slightly [13,14], and the storage costs could be reduced by compression using surface-adapted wavelet multiresolution techniques [12].

The choice of basic numerical methods used is also strongly influenced by the need for responsiveness. In some cases direct methods are attractive because one can explicitly bound the amount of computation required and ensure that the responsiveness constraints are satisfied. A more sophisticated approach is to use anytime methods (e.g., [4]) which can produce an approximate solution at any time. It is possible to reorganize an entire simulation in an anytime fashion, including the time integration of differential equations, so that solution accuracy improves monotonically with available computational resources [26].

3.2 Signals vs. Features

How should we organize our computations for multisensory interaction? In the end, a simulation must produce unimodal *signals*, such as sequences of images to display on

[1] Note that these are constraints in the sense that once they are satisfied, there is little or no additional benefit to, say, computing the haptic forces at 10 KHz.

screen, forces to apply with a haptic interface, and sounds to play over loudspeakers. But one can not simply compute these signals separately because their underlying causes are tightly coupled. This seems obvious but, unfortunately, it is one of the main reasons why contact interactions in video games today appear unrealistic — motion is simulated by one method, while the haptic response is computed separately (using a separate haptic API), and the accompanying sound may be "simulated" by playing a previously recorded audio file. Essential correlations between the signals are lost by this decoupling of different sensory modalities.

One may then be tempted to build a monolithic physical simulation which produces all the required signals; for instance, a contact simulation could be constructed using the Finite Element Method, which can produce visible large scale deformations, contact forces, and the contact sounds emitted by vibrating surface patches (e.g., [25]). This has the advantage of automatically producing the required correlations, but it is not well suited for multisensory interaction. The problem is that different modalities place different requirements on a simulation. From a perceptual point of view, visual motion of the surfaces must be computed at about 60 Hz, but haptic forces must be computed at about 1 KHz, and sounds must be computed at about 48 KHz. On the other hand, visual motion must be computed at a relatively high spatial resolution while the emitted sound can be lumped more coarsely, due to differences in spatial acuity between human vision and hearing. Therefore a monolithic simulation will have to use very fine meshes and very small time steps, making it very difficult to satisfy the responsiveness constraints.

We propose instead to organize a multisensory interactive simulation based on *spatiotemporal features*, and to simulate the occurrence of these features separately from their multisensory "rendering" (i.e., conversion to signals). Features have played an important role in robotic perception for a very long time, but there has been little corresponding work in simulation[2] (but see, for example, [2]).

For example, in our approach, an impact is a spatiotemporal contact feature, frictional sliding is another. An interactive simulation proceeds by computing the active features and their properties. For instance, the simulation could decide that an impact has occured at a particular time and at a particular location on object, based on approximate collision detection with a bounding volume hierarchy. Once this impact has been decided and parameters such as the total impulse delivered are computed, the impact feature can be rendered as different sensory signals separately. It could be rendered haptically at 1 KHz using a stiff wall model (e.g., [34]), and the impact sound could be efficiently rendered using a modal vibration model at 48 KHz (e.g., [43]).

Such feature-based simulation has some important advantages. First, it appears to match human perception, which is believed to process stimuli by extracting features that are relatively stable. Such features have been extensively studied in visual perception [23], but they have not received much attention in multisensory perception [39]; more computational modeling is needed in this regard. Even though

[2] In a sense, discrete event simulation methods could be considered as a kind of feature-based simulation.

it is not yet clear what multisensory features are used in human perception, it seems likely that the spatiotemporal features will correspond to qualitative modes of physical behavior (impact vs. continuous contact, sliding vs. rolling, etc.). By simulating features explicitly, we increase the likelihood that perceptible correlations between different modalities are preserved. Second, by separating the rendering of different sensory signals, it is possible to take advantage of modality-specific algorithms and hardware to satisfy the responsiveness constraint. For instance, visual deformation can be well represented by a few deformation modes with complex mode shapes — this can be efficiently rendered at high spatial resolution using vertex processors on modern graphics hardware [15]. On the other hand, sound rendering requires high temporal resolution that can be achieved either using efficient algorithms [44], or by using wavetable synthesis on DSPs available in modern sound cards.

Two earlier examples of systems for multisensory interaction based on some of these principles are our FoleyAutomatic system [42] which combined audio-visual interaction with rigid-body contact models, and our ArtDefo system [13,14] which combined haptic-visual interaction with deformable models.

3.3 Physically based vs. Reality-based

Finally, where should these models come from? In science and engineering analysis, one starts with a formally specified theory to be validated or a design to be evaluated. The physical model could be very sophisticated, but the model and its parameters are compact and given by the user as a part of the specification. For instance, a contact simulation may use a sophisticated nonlinear viscoelastic model, but it may require only a few material and shape parameters to be specified. The main priority is to simulate the subtle consequences of the physical model.

Unfortunately, such an approach is not well suited for a robot or a human trying to manipulate a novel object. Real objects, particularly biological and naturally occuring ones, often have complex internal structure and internal variations in material parameters. So, a physically based FEM model of elastic deformation would require a large number of internal material parameters to be estimated from observations of the object, by solving a delicate inverse problem. Therefore our robot has a different priority: how quickly and easily can it construct a sufficient model (including its parameters) from measurements, including the costs of acquiring the required measurements.

This suggests that models for multisensory interaction should be more directly based on the input-output measurements that can be easily and quickly acquired. We have previously called these *reality-based models* [28], to distinguish them from standard physically based models. As the costs of sensors and memory drop, such reality-based models will become increasingly viable.

For instance, rather than estimate the internal material distribution, we could construct a local linear approximation of the input-output deformation behavior of an object using Green's functions [27,20]. Thus, in addition to the responsiveness advantage discussed in Sec. 3.1, a key advantage is the relative ease with which reality-based models can be contructed. We need relatively few measurements, taken only at the surface (i.e., contact tractions measured with a force sensor and global surface displacements acquired using stereo vision).

3.4 Acquiring and rendering multisensory models

New human interfaces and robotic systems need to be designed to address the specific requirements of multisensory interaction. We have previously described how multisensory models could be acquired using robotics [28]. The UBC Active Measurement facility (ACME) was developed specifically for this purpose, and could acquire multisensory models of small objects, including shape, surface roughness and friction, sound response, and deformation response [30,27].

However, it is still a challenge to develop multisensory virtual environments that feel as natural as the real world. To explore the issues we are developing an environment called the HAVEN (Haptic, Auditory, and Visual Environment), which we briefly preview here. It is a physical space designed for the multisensory human interaction and measurement. (see Figure 1). It consists of a specially constructed and instrumented chamber in which users can interact with other users, physical objects, and computer simulations. The chamber is designed to provide acoustic and optical isolation from ambient noise. As users interact in the HAVEN, their motion, applied forces, and sounds are simultaneously measured using a multitude of sensors located on the walls, on the ceiling, on the table top, and also attached to manipulated objects and probes, and even attached to the user's hand. Examples of sensors include a 1 KHz Vicon motion capture system, a microphone array, a pressure sensor pad with ten thousand capacitative sensors, and a hand-held WHaT (Wireless Haptic Texture sensor) [31]. This enormous amount of information is digitized and processed in real time, on a cluster of computers connected using a gigabit network. To provide rapid multisensory feedback to users, the HAVEN contains two projection displays, a multi-speaker array for auditory display. Haptic devices will be incorporated in the near future.

4 Summary

Human information processing and interaction is multisensory in significant ways. Therefore, human interfaces should take this into account, and robotic systems could benefit from using similar techniques. We have argued that computational models for multisensory interaction have very different requirements than traditional models used in science and engineering, and we have suggested how new models could be built.

Acknowledgements

I would like to thank my students and colleagues for contributing to these ideas in many subtle ways, in particular U. Ascher, M. Cline, K. van den Doel, T. Edmunds,

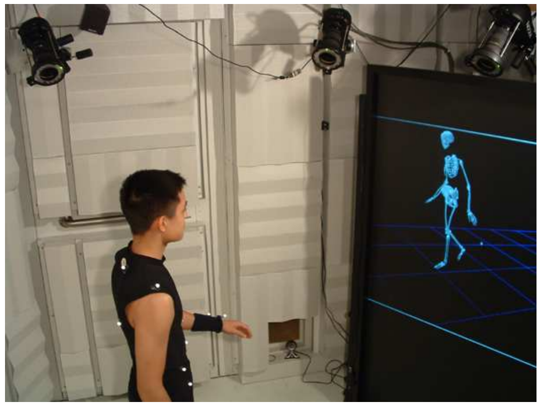

Fig. 1. Current status of the HAVEN (Haptic, Auditory, and Visual Environment. The figure shows one corner of the HAVEN, with the motion capture cameras on top, a rear projection display, and acoustically treated walls in the background. The subject in the foreground is wearing retroreflective markers that can be tracked in real time.

D. L. James, P. G. Kry, J. Lang, J. E. Lloyd, S. Sueda, and K. Yin. This work was supported in part by NSF grants EIA-0215887 and IIS-0308157, and the IRIS Network of Centres of Excellence.

References

1. O. R. Astley and V. Hayward. Multirate haptic simulation achieved by coupling finite element meshes through norton equivalents. In *ICRA*, pages 989–994, 1998.
2. Stephen Chenney and D. A. Forsyth. Sampling plausible solutions to multi-body constraint problems. In *Proceedings of ACM SIGGRAPH 2000*, Computer Graphics Proceedings, Annual Conference Series, pages 219–228, July 2000.
3. S. Cotin, H. Delingette, J-M Clement, V. Tassetti, J. Marescaux, and N. Ayache. Geometric and physical representations for a simulator of hepatic surgery. In *Proceedings of Medicine Meets Virtual Reality IV*, pages 139–151. IOS Press, January 1996.
4. T. Dean and M. Boddy. An analysis of time dependent planning. In *AAAI-88*, 1988.
5. D. E. DiFranco, G. L. Beauregard, and M. A. Srinivasan. The effect of auditory cues on the haptic perception of stiffness in virtual environments. In *Sixth Annual Symposium on Haptic Interfaces for Virtual Environments and Teleoperator Systems*, pages DSC–Vol. 61, pp. 17–22. ASME, 1997.
6. M. A. Erdmann and M. T. Mason. An exploration of sensorless manipulation. *IEEE Journal of Robotics and Automation*, 4(4):369–379, 1988.
7. M.O. Ernst and M.S. Banks. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870):429–433, Jan 2002.
8. J.M. Groh, A.S. Trause, A.M. Underhill, K.R. Clark, and S. Inati. Eye position influences auditory responses in primate inferior colliculus. *Neuron*, 29(2):509–518, Feb 2001.
9. S. Guest and C. Spence. Tactile dominance in speeded discrimination of textures. *Exp Brain Res*, 150(2):201–207, May 2003.

10. S. Guest and C. Spence. What role does multisensory integration play in the visuotactile perception of texture? *Int J Psychophysiol*, 50(1-2):63–80, Oct 2003.
11. R. D. Howe and M. R. Cutkosky. Sensing skin acceleration for slip and texture perception. In *Proc. IEEE International Conference on Robotics and Automation*, pages 145–150, 1989.
12. Doug James and Dinesh K. Pai. Multiresolution green's function methods for interactive simulation of large-scale elastostatic objects. *ACM Transactions on Graphics*, 22(1):47–82, January 2003.
13. Doug L. James and Dinesh K. Pai. ArtDefo, accurate real time deformable objects. In *Computer Graphics (ACM SIGGRAPH 99 Conference Proceedings)*, pages 65–72, August 1999.
14. Doug L. James and Dinesh K. Pai. A unified treatment of elastostatic contact simulation for real time haptics. *Haptics-e, the Electronic Journal of Haptics Research*, 2(1), 2001. `http://www.haptics-e.org`.
15. Doug L. James and Dinesh K. Pai. DyRT: Dynamic response textures for real time deformation simulation with graphics hardware. *ACM Transactions on Graphics*, 21(3):582–585, 2002. Also *ACM SIGGRAPH 2002 Conference Proceedings*, July 2002.
16. E. R. Kandel, J. H. Schwartz, and T. M. Jessell, editors. *Principles of Neural Science*. McGraw-Hill/Appleton & Lange; 4th edition, 2000.
17. N. Kikuchi and J. Oden. *Contact problems in elasticity: a study of variational inequalities and finite element methods*. SIAM, 1988.
18. N. Kitagawa and S. Ichihara. Hearing visual motion in depth. *Nature*, 416(6877):172–174, Mar 2002.
19. E. Kohler, C. Keysers, M.A. Umilta, L. Fogassi, V. Gallese, and G. Rizzolatti. Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, 297(5582):846–848, Aug 2002.
20. Jochen Lang, Dinesh K. Pai, and Robert J. Woodham. Acquistion of elastic models for interactive simulation. *International Journal of Robotics Research*, 2002. (to appear).
21. S.J. Lederman and R.L. Klatzky. Multisensory texture perception. In G. Calvert, C. Spence, and B. Stein. Handbook of Multisensory Processes. MIT Press: Cambridge., 2004. (to appear).
22. S.J. Lederman, G. Thorne, and B. Jones. Perception of texture by vision and touch: multidimensionality and intersensory integration. *J Exp Psychol Hum Percept Perform*, 12(2):169–180, May 1986.
23. David Marr. *Vision*. W.H. Freeman and Company, San Francisco, 1982.
24. H. McGurk and J. MacDonald. Hearing lips and seeing voices. *Nature*, 264(5588):746–748, Dec 1976.
25. J. F. O'Brien, P. R. Cook, and G. Essl. Synthesizing sounds from physically based motion. In *Proceedings of ACM SIGGRAPH 2001*, Computer Graphics Proceedings, Annual Conference Series, pages 529–536, August 2001.
26. D. K. Pai and U. M. Ascher. Anytime integrators for interactive simulation. (to appear).
27. D. K. Pai, K. van den Doel, D. L. James, J. Lang, J. E. Lloyd, J. L. Richmond, and S. H. Yau. Scanning physical interaction behavior of 3d objects. In *Computer Graphics (ACM SIGGRAPH 2001 Conference Proceedings)*, pages 87–96, August 2001.
28. Dinesh K. Pai. Robotics in reality-based modeling. In *Proceedings of the International Symposium on Robotics Research, Snowbird, UT*, October 1999. To appear as *Robotics Research: the Ninth International Symposium*, Springer-Verlag.
29. Dinesh K. Pai. Strands: Interactive simulation of thin solids using cosserat models. *Computer Graphics Forum, The International Journal of the Eurographics Association*, 21(3):347–352, 2002.

30. Dinesh K. Pai, Jochen Lang, John E. Lloyd, and Robert J. Woodham. Acme, a telerobotic active measurement facility. In *Proceedings of the Sixth Intl. Symp. on Experimental Robotics*, 1999.
31. Dinesh K. Pai and Peter Rizun. The what: a wireless haptic texture sensor. In *Proceedings of the Eleventh Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, 2003. (to appear).
32. G. Rizzolatti and M.A. Arbib. Language within our grasp. *Trends Neurosci*, 21(5):188–194, May 1998.
33. I. Rock and j. Victor. Vision and touch: An experimentally created conflict between the two senses. *Science*, 143:594–596, Feb 1964.
34. S.E. Salcudean and T.D. Vlaar. On the emulation of stiff walls and static friction with a magnetically levitated input/output device. Technical report, Dept. of Electrical Engineering, University of British Columbia, Vancouver, BC, Canada, March 1994.
35. R. Sekuler, A.B. Sekuler, and R. Lau. Sound alters visual motion perception. *Nature*, 385(6614):308, Jan 1997.
36. L. Shams, Y. Kamitani, and S. Shimojo. Illusions. what you see is what you hear. *Nature*, 408(6814):788, Dec 2000.
37. S. Shimojo and L. Shams. Sensory modalities are not separate modalities: plasticity and interactions. *Curr Opin Neurobiol*, 11(4):505–509, Aug 2001.
38. M. A. Srinivasan, G. L. Beauregard, and D. Brock. The impact of visual information on the haptic perception of stiffness in virtual environments. In *Fifth Annual Symposium on Haptic Interfaces for Virtual Environments and Teleoperator Systems*, pages DSC–Vol. 58, pp. 17–22. ASME, 1996.
39. B. E. Stein and M. A. Meredith. *The Merging of the Senses*. MIT Press, 1993.
40. J.A. Stevens, P. Fonlupt, M. Shiffrar, and J. Decety. New aspects of motion perception: selective neural encoding of apparent human movements. *Neuroreport*, 11(1):109–115, Jan 2000.
41. K.N. Stevens and M. Halle. Remarks on analysis by synthesis and distinctive features. In W. Wathen-Dunn, editor, *Proc. of the AFCRL Symposium on Models for the Perception of Speech and Visual Form*, 1964.
42. K. van den Doel, P.G. Kry, and D. K. Pai. FOLEYAUTOMATIC: Physically-based sound effects for interactive simulation and animation. In *Computer Graphics (ACM SIGGRAPH 2001 Conference Proceedings)*, pages 537–544, August 2001.
43. K. van den Doel and D. K. Pai. Synthesis of shape dependent sounds with physical modeling. In *International Conference on Auditory Display (ICAD 96)*, Xerox PARC, Palo Alto, November 1996.
44. Kees van den Doel and Dinesh K. Pai. The sounds of physical shapes. *Presence*, 7(4):382–395, 1998.
45. R.B. Welch and D.H. Warren. Immediate perceptual response to intersensory discrepancy. *Psychol Bull*, 88(3):638–667, Nov 1980.
46. U. Werner-Reiss, K.A. Kelly, A.S. Trause, A.M. Underhill, and J.M. Groh. Eye position affects activity in primary auditory cortex of primates. *Curr Biol*, 13(7):554–562, Apr 2003.
47. KangKang Yin and Dinesh K. Pai. Footsee: an interactive animation system. In *Proceedings of the ACM SIGGRAPH Symposium on Computer Animation*, pages 153–159. ACM, July 2003.