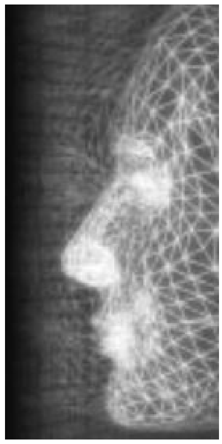
COMPUTER ANIMATION AND VIRTUAL WORLDS
Comp. Anim. Virtual Worlds 2006; **17**: 411–419
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/cav.144

Integrating physically based sound models in a multimodal rendering architecture



By Federico Avanzini and Paolo Crosato*

This paper presents a multimodal rendering architecture that integrates physically based sound models with haptic and visual rendering. The proposed sound modeling approach is compared to other existing techniques. An example of implementation of the architecture is presented, that realizes bimodal (auditory and haptic) rendering of contact stiffness. It is shown that the proposed rendering scheme allows tight synchronization of the two modalities, as well as a high degree of interactivity and responsiveness of the sound models to gestures and actions of a user. Finally, an experiment on the relative contributions of haptic and auditory information to bimodal judgments of contact stiffness is presented. Experimental results support the effectiveness of auditory feedback in modulating haptic perception of stiffness. Copyright © 2006 John Wiley & Sons, Ltd.

Received: 10 April 2006; Revised: 2 May 2006; Accepted: 10 May 2006
KEY WORDS: multimodal interaction; sound rendering; stiffness

Introduction

Most of virtual reality (VR) applications built to date make use of visual displays, haptic devices, and spatialized sound displays. Multisensory information is essential for designing immersive virtual worlds, as an individual's perceptual experience is influenced by interactions among sensory modalities. As an example, in real environments visual information can alter the haptic perception of object size, orientation, and shape.[1] Similarly, being able to hear sounds of objects in an environment, while touching and manipulating them, provides a sense of immersion in the environment not obtainable otherwise.[2] Properly designed and synchronized haptic and auditory displays are likely to provide much greater immersion in a virtual environment than a high-fidelity visual display alone. Moreover, by skewing the relationship between the haptic and visual and/or auditory displays, the range of object properties that can be effectively conveyed to the user can be significantly enhanced.

*Correspondence to: Federico Avanzini, Department of Information Engineering, University of Padova, Via Gradenigo 6/A, 35131 Padova, Italy. E-mail: avanzini@dei.unipd.it

Contract/grant sponsor: EU Sixth Framework Programme—IST Information Society Technologies.

Recent literature has shown that sound synthesis techniques based on physical models of sound generation mechanisms allow for high quality synthesis and interactivity, since the physical parameters of the sound models can be naturally controlled by user gestures and actions. Sounds generated by solid objects in contact are especially interesting since auditory feedback is known in this case to provide relevant information about the scene (e.g., object material, shape, size). Physically based sound models of impulsive and continuous contact[3,4] have been applied to the development of an audio-haptic interface for contact interactions.[5]

Many studies on bimodal (auditory and haptic) perception in contact interaction are focused on continuous contact (i.e., scraping or sliding). Lederman and coworkers have provided many results[6,7] on the relative contributions of tactile and auditory information to judgments of surface roughness. Guest *et al.*[8] have also focused on audio-tactile interactions in roughness perception. McGee *et al.*[9] studied bimodal perception of *virtual* roughness, that is, roughness of synthetic haptic and auditory textures. Impulsive contact interaction (i.e., impact) are apparently less investigated. DiFranco *et al.*[10] studied the effect of auditory feedback on haptic stiffness perception. In the setup, a contact detection event in the haptic rendering pipeline triggered headphone reproduction of a real impact sound. Sounds were recorded by tapping various tools against surfaces of various materials. One drawback in using recorded samples is that auditory feedback is static and is not controlled by user actions. Moreover, there is no control in the design of sounds. Studies in ecological acoustics[11,12] provide useful indications about which auditory cues are relevant to stiffness/hardness perception, and can be exploited in the design of synthetic sound feedback.

In this paper, we discuss the integration of physically based contact sound models into a multimodal rendering architecture. We first review related literature and compare our sound modeling approach to other proposed techniques. Next, we discuss bimodal (auditory and haptic) rendering of contact stiffness: our physically based impact sound model is presented, a mapping between model parameters and auditorily perceived stiffness is proposed, and the architecture for audio-haptic rendering is presented. Finally, we report upon an experiment on bimodal stiffness perception.

Sound Models for VR

According to Hahn *et al.*[2] three main problems have to be addressed in sound generation for VR applications. First, sound *modeling* techniques have to be used that allow an effective mapping between parameters generated from motion and sound control parameters. A second critical issue is *synchronization* between sound and other modalities. The last stage of sound *rendering* entails the process of generating sound signals within a given environment, that is, tracing the emitted sound energy within the environment and including perceptual processing in order to account for listener effects. The whole process of rendering sounds can be seen as a pipeline analogous to an image-rendering pipeline.

Until recently, the primary focus for sound generation in VR applications has been in spatial localization of sounds. On the contrary, research about models for sound sources and mappings between object motion/interaction and sound control is far less developed.

Physically Based Approaches

Sound synthesis techniques traditionally developed for computer music applications (e.g., additive, subtractive, frequency modulation[13]) provide abstract descriptions of sound signals. Although well suited for the representation of musical sounds, these techniques are in general not effective for the generation of non-musical interaction sounds.

The term physically based sound modeling[14] refers to a different set of sound synthesis algorithms, based on a physical description of the sound-generating mechanisms. Since these models generate sound from computational structures that respond to physical input parameters, they automatically incorporate complex responsive acoustic behaviors. Moreover, the physical control parameters do not require in principle manual tuning in order to achieve realistic output. Research in ecological acoustics[15,16] aids in determining what sound features are perceptually relevant, and can be used to guide the tuning process.

A second advantage of physically based approaches is interactivity and ease in associating motion to sound control. As an example, the parameters needed to characterize collision sounds, e.g., relative normal velocity, are computed in the VR physical simulation engine and can be directly mapped into control parameters of a sound model. The sound feedback consequently responds in a natural way to user gestures and actions.

Finally, physically based sound models can in principle allow the creation of dynamic virtual environments in which sound-rendering attributes are incorporated into data structures that provide multimodal encoding of object properties: shape, material, elasticity, texture, mass, and so on. In this way, a unified description of the physical properties of an object can be used to control the visual, haptic, and sound rendering, without requiring the design of separate properties for each thread. This problem has already been studied in the context of joint haptic-visual rendering, and recent haptic-graphic APIs[17,18] adopt a unified scene graph that takes care of both haptics and graphics rendering of objects from a single scene description, with obvious advantages in terms of synchronization and avoidance of data duplication. Physically based sound models may allow the development of a similar unified scene that includes description of audio attributes as well.

Contact Sound Modeling

Research on physically based sound models for VR is mostly focused on *contact* sounds between solids, that is, sounds generated when objects come into contact (collision, rubbing, etc.). Less has been done on liquid and aerodynamic sounds. A synthesis technique for liquid sounds was proposed by van den Doel,[19] while a method for creating aerodynamic sounds (e.g., sound generated by swinging swords, or by wind blowing) was presented by Dobashi *et al*.[20]

Various approaches have been proposed in the literature for contact sound modeling. Van den Doel and coworkers[3] proposed modal synthesis[21] as a framework for describing the acoustic properties of objects. Pre-computed contact force signals are used to drive the modal synthesizer, under the assumption that the sound-producing phenomena are linear, thus being representable as source-filter systems. The modal representation is naturally linked to many *ecological* dimensions of the corresponding sounds: modal frequencies depend on the *shape* and the geometry of the object, the *material* determines to a large extent the decay characteristics of the sound,[22] amplitudes of the frequency components depend contact location, and so on. DiFilippo and Pai[5] applied the techniques described above to audio-haptic rendering: in the proposed architecture contact forces are computed at the rate of the haptic rendering routine (e.g., 1 kHz), then the force signals are upsampled to audio rate (e.g., 44.1 kHz) and the resulting audio force is used to drive the modal sound model.

A different approach was proposed by O'Brien *et al.*[23] Finite-element simulations are employed for the generation of both animated video and audio, by analyzing surface motions of animated objects, isolating vibrational components that correspond to audible frequencies, and mapping surface motions to acoustic pressure waves in the surrounding medium. In this way, complex audio-visual scenes can be simulated, but heavy computational loads prevent real-time rendering and the use of the method in interactive applications.

We have proposed modal synthesis techniques,[4,24] in which the main difference with the previously discussed works lies in the approach to contact force modeling. Instead of adopting a feed-forward scheme in which resonators are excited with pre-computed driving forces, these models embed computation of non-linear contact forces in the sound rendering thread. Despite the complications that arise in the synthesis algorithms, this approach provides some advantages. Better quality is achieved due to accurate audio-rate computation of contact forces: this is especially true for impulsive contact, where contact times are in the order of few milliseconds. Interactivity and responsiveness of sound to user actions is also improved. This is especially true for continuous contact, such as stick-slip friction.[24] Finally, physical parameters of the contact force models provide control over other ecological dimensions of the sound events: as an example, in the next section we discuss a mapping between parameters of an impact force model and auditorily perceived stiffness.

In this paper, our sound models[4,24] are integrated into a multimodal rendering architecture (see Figure 1), which extends typical haptic-visual architectures.[25] The sound rendering thread runs at audio rate (e.g., 44.1 kHz) in parallel with other threads. Computation of audio contact forces is triggered by collision detection from the haptic rendering thread. Computation of 3D sound can be cascaded to the sound synthesis block. In the next section,

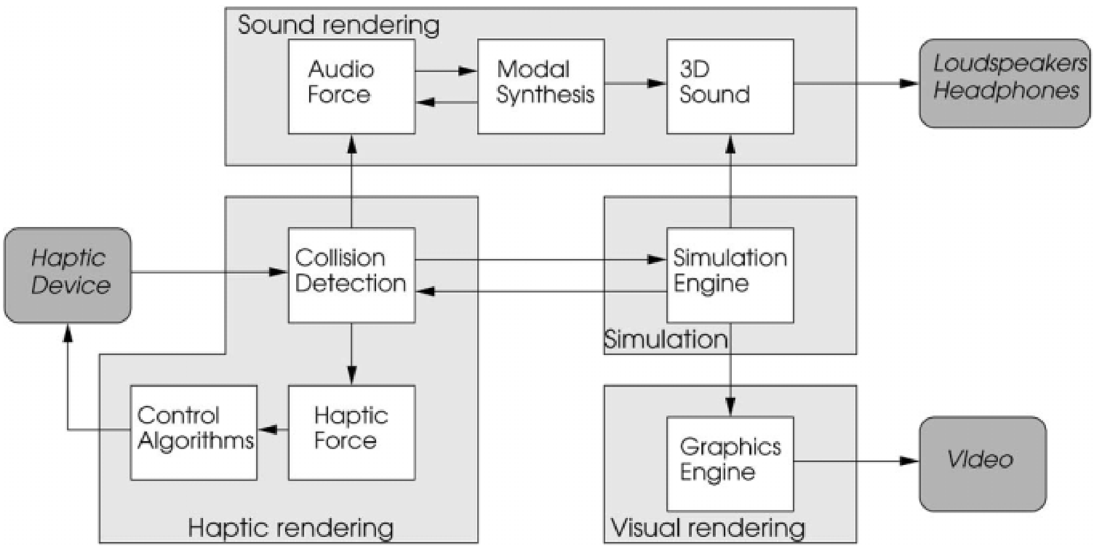


Figure 1. An architecture for multimodal rendering of contact interactions. Adapted from Figure 3 in Reference [25].

we exemplify this scheme for the case of impulsive contact.

Bimodal Rendering of Contact Stiffness

Rendering a virtual surface, that is, simulating the interaction forces that arise when touching a stiff object, is the prototypical haptic task. Properly designed visual[26] and/or auditory[10] feedback can be combined with haptics in order to improve perception of stiffness. Physical limitations (low sampling rates, poor spatial resolution of haptic devices) constrain the range for haptic stiffness rendering. As an example, the nominal maximum closed-loop control stiffness for the Phantom® Omni™ device used in this study is 500 N/m, which is far from typical values for stiff surfaces.[27] Ranges for haptic stiffnesses are usually estimated by requiring the system to be passive,[28] thus guaranteeing stability of the interaction, while higher stiffness values can cause the system to become unstable, that is, to oscillate uncontrollably. The addition of visual and auditory feedback can compensate for such limitations and enhance the range of perceived stiffness that can be effectively conveyed to the user.

A Physical Model for Impact Sounds

The theory of modal analysis[21] states that, given a differential system composed of N coupled equations representing a resonating object as a network of N masses connected with springs and dampers, a geometrical transformation can be found that turns the system into a set of decoupled equations. The transformed variables $\{q_n\}_{n=1}^{N}$ are generally referred to as *modal displacements*, and obey a second-order linear oscillator equation:

$$\ddot{q}_n(t) + g_n\dot{q}_n(t) + \omega_n^2 q_n(t) = \frac{1}{m_n} f_A(t) \qquad (1)$$

where q_n is the oscillator displacement and f_A represents any audio driving force, while ω_n is the oscillator center frequency. The parameter $1/m_n$ controls the 'inertial' properties of the oscillator (m_n has the dimension of a mass), and g_n is the oscillator damping coefficient and relates to the decay properties of the system. Modal displacements q_n are related to physical displacement through an $N \times L$ matrix $\mathbf{A}$, whose elements a_{nl} weigh the contribution of the nth mode at a location l. If f_A is an impulse, the response q_n of each mode is a damped sinusoid and the physical displacement at location l is given by

$$x_l(t) = \sum_{n=1}^{N} a_{nl} q_n(t) = \sum_{n=1}^{N} a_{nl} e^{-g_n t/2} \sin(\omega_n t) \qquad (2)$$

In general, however, the contact force has a more complex description, and depends on the state (displacement and velocity) of the colliding modal objects. In this section, we restrict our attention to impact force modeling, and use a model originally proposed by Hunt and Crossley[29] and previously applied to modal sound synthesis.[4] The non-linear impact force between two colliding objects is in this case

$$f_A(x(t), v(t)) = \begin{cases} k_A x(t)^\alpha + \lambda_A x(t)^\alpha \cdot v(t) & x \geq 0 \\ 0 & x < 0 \end{cases} \qquad (3)$$

where the *compression* x at the contact point is defined as the difference between the displacements of the two bodies, and $v(t) = \dot{x}(t)$ is the compression velocity. The condition $x \geq 0$ states that there is actual compression, while for $x < 0$ the two objects are not in contact.

The force model (3) includes both an elastic component $k_A x^\alpha$ and a dissipative term $\lambda_A x^\alpha v$. The latter accounts for viscoelastic losses during collision. The parameter k_A in Equation (3) is the force *stiffness* and is in general a function of the mechanical properties of the two bodies, while λ_A is the force *damping weight*. Additionally a variable exponent α is introduced, whose value depends on the surface geometry of the contact (e.g., $\alpha = 3/2$ for the particular case of two spheres in contact).

Auditory Stiffness

Freed[11] has investigated the ability of listening subjects to estimate the hardness of hammers made of various materials, from the sound that they generated when striking metallic pans of variable sizes. His experiments showed that the useful information for hardness rating is contained in the attack transients, namely in the first 300 milliseconds, of the sounds. Loudness and descriptors related to the spectral centroid (average value and temporal variability in the first 300 milliseconds) were used as predictors in a multiple regression analysis, and were found to account for 75% of the variance of the hardness ratings. In a related study, Giordano[12] argues that the duration τ of the contact between the two

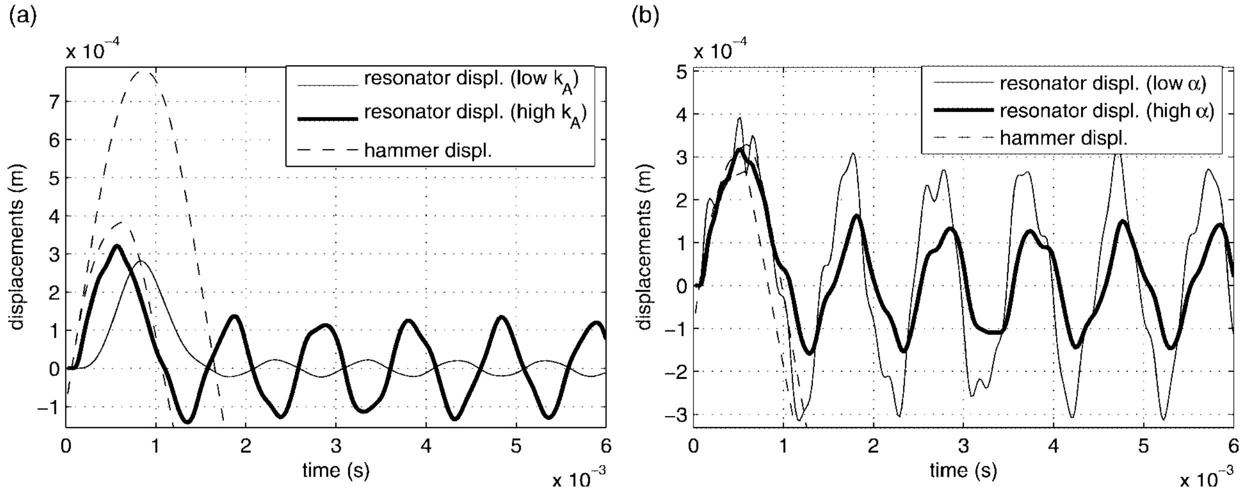


Figure 2. Examples of transient attacks obtained from the impact sound model: (a) short versus long initial bumps, obtained by varying the force stiffness k_A; (b) single versus multiple contacts obtained by varying the exponent α.

objects during the stroke has an influence on hardness perception, and that τ variations are likely to explain at least in part data reported by Freed.[11] Specifically, an increase in τ determines a decrease in the loudness of the radiated signal, and in the amount of energy at high frequencies (and thus in the spectral centroid), since vibrational modes with a period higher than τ are minimally excited.

Based on similar considerations, we have investigated[4] the dependence of contact time τ and the attack spectral centroid on the parameters of the impact force model (3). The following equation was derived for τ:

$$\tau = \left(\frac{m}{k_A}\right)^{\frac{1}{\alpha+1}} \cdot \left(\frac{\mu^2}{\alpha+1}\right)^{\frac{\alpha}{\alpha+1}} \cdot \int_{v_{\text{out}}}^{v_{\text{in}}} \frac{dv}{(1+\mu v)\left[-\mu(v-v_{\text{in}}) + \log\left|\frac{1+\mu v}{1+\mu v_{\text{in}}}\right|\right]^{\frac{\alpha}{\alpha+1}}} \quad (4)$$

where $v_{\text{in}}, v_{\text{out}}$ are the normal velocities before/after collision, respectively, and $\mu = \lambda_A/k_A$ is a mathematically convenient term. Equation (4) states in particular a power-law dependence of τ on the force stiffness: $\tau(k_A) \sim k_A^{-1/\alpha+1}$. A study[30] on synthetic impact sounds obtained from model (3) provided quantitative results that confirm the correlation between spectral centroid of the attack transients and τ. An example of the effect of k_A on the sound attack is provided in Figure 2(a). The dissipative component of the contact force also has a slight effect on the centroid: as λ_A is lowered, the amount of energy transferred to the higher partials is increased, and the centroid increases accordingly, even though τ remains approximately constant. Similarly, the centroid increases as α decreases, even though the contact time varies slowly. This effect is illustrated in Figure 2(b): as α is lowered, the magnitude of the impact force increases significantly and eventually multiple bounces of the vibrating surface on the striking object are produced, with a consequent increase of the centroid.

In summary, the results reported in this section show that manipulation of the impact force parameter k_A affects in a predictable way the contact time and the sound attack transient, and has, therefore, a major influence on the perception of impact hardness.

Realization

The contact sound synthesis models are currently implemented as plugins for the open-source real-time synthesis environment pd (Pure Data†). The numerical realization of the sound models is based on the analysis presented in previous works.[4,24]

The software experimental setup consists of two processes which communicate by means of a shared memory area (see Figure 3). The first process is responsible for graphic and haptic rendering and is realized with the Openhaptics™ Toolkit developed by Sensable.‡ Graphics is built in OpenGL while the haptic rendering thread

†http://crca.ucsd.edu/~msp/.
‡http://www.sensable.com.

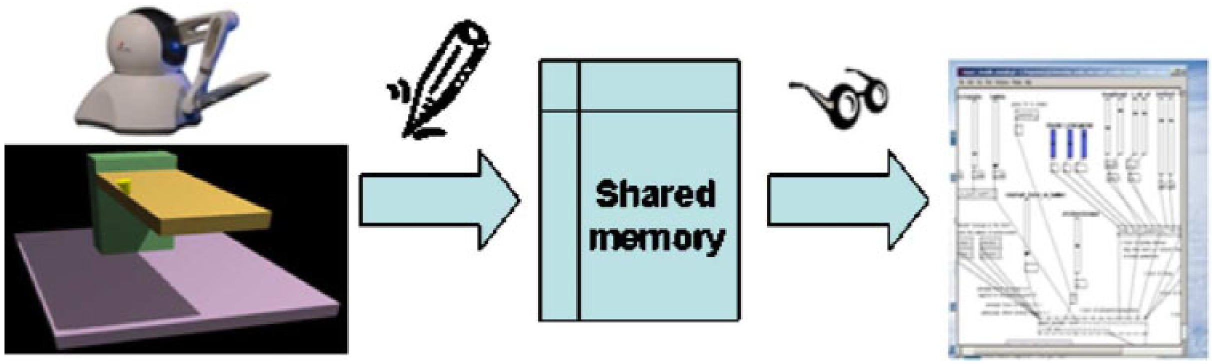


Figure 3. The software architecture of the experimental setup.

uses a state machine architecture which resembles very closely the model upon which OpenGL is built. An event-catching engine driven by a function callback model is adopted to monitor contact events. When such an event occurs, haptic data necessary for sound synthesis are written into the shared memory area. The second process is a 'patch' run by `pd`, which reads data from the shared memory area and renders contact sounds according to the current physical parameters. This architecture implements the general scheme depicted in Figure 1.

Low communication latency is critical in order to ensure unitary perception rather than perception of two distinct auditory and haptic events. The main sources of delay in the process are graphic rendering, sound rendering, and the write/read access to the shared memory area. As for the latter, the code was heavily optimized for maximum speed: many simulations of cyclic write/read accesses showed that the delay introduced is in the order of microseconds, thus being irrelevant if compared to microseconds order delays introduced by sound and graphic synthesis. The resulting latency is well below typical experimental estimates for temporal windows of auditory-tactile integration (see for example, References [31,32]). During the experimental test reported in the next section, no subjects perceived any kind of noticeable intermodal latency.

An Experiment on Bimodal Stiffness Perception

The architecture described above has been tested on an experiment for the assessment of the effectiveness of auditory feedback in modulating haptic perception of stiffness.

Setup

The graphic display provided to subjects is shown in Figure 4(a). The haptic stiffness was given a constant value of 400 N/m. According to literature,[27] this can be considered an average value, with 'soft' values being below 300 N/m and 'hard' values starting above 600–700 N/m. A picture of the experimental setup is provided in Figure 4(b).

Auditory stiffness levels were obtained by varying the parameter k_A (see Equation (3)), while all the remaining parameters of the physical sound model were held constant. The fundamental frequency and modal frequency distribution of the struck object were chosen on the basis of the equations for the ideal bar, where density and Young's modulus were given intermediate values between wood and glass. The modal decay times were also chosen to match intermediate values between wood and glass, in order not to provide a definite auditory perception of material. Impact force parameters other than k_A were held constant.

Given this setup, the interval of variability for the stiffness k_A was determined empirically as the largest interval outside of which further stiffness variations do not produce noticeable effects in the physical model behavior. Finally, a series of seven exponentially spaced values for the auditory stiffness was sampled within this interval.

Subjects were instructed to judge the stiffness of the impact between a 'hammer,' represented by the stylus of the haptic device, and the bar in the display. Every object in the scene was rendered haptically, but only touching the upper bar produced a sound. The graphic display did not change between conditions, and was intentionally composed of stylized objects, in order to limit as much as possible the amount of visual information delivered to subjects. Participants

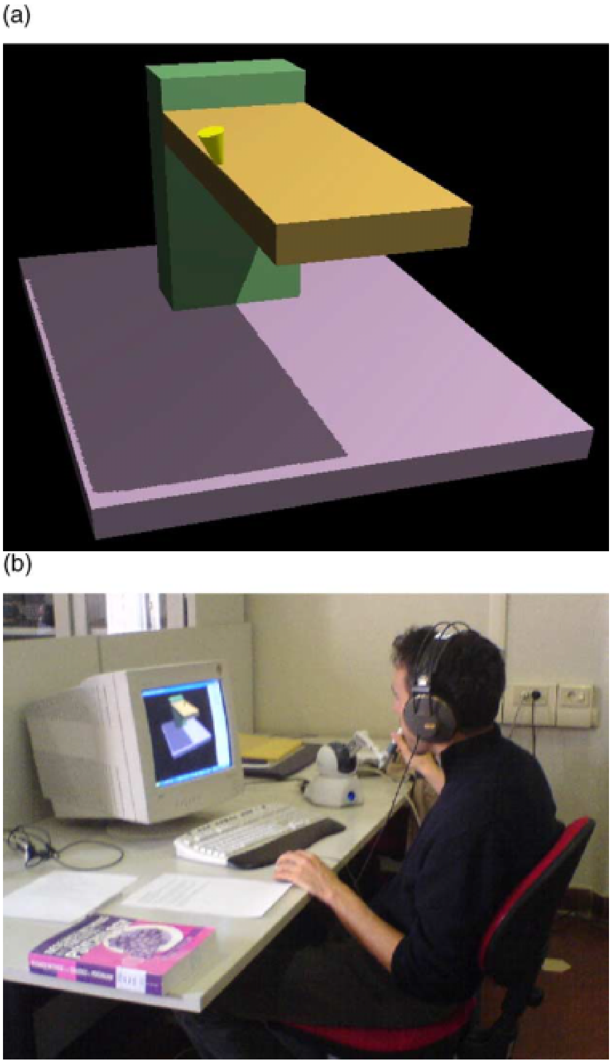


Figure 4. Experimental procedure: (a) interactive graphic display presented to the subjects (the small cone represents the tip of the Phantom® Omni™ stylus); (b) experimental setup.

were allowed to interact with each condition as long as desired.

Perceived stiffness was determined through an absolute magnitude-estimation procedure (similarly to the approach reported by Lederman *et al.*[7]): participants were instructed to assign the non-zero, positive number that best described the magnitude of the perceived stiffness of the stimulus, along a scale ranging from 1 to 8. Verbal labels were associated to each point of the scale, ranging from 'extremely soft' (1) to 'extremely stiff' (8).

Participants did not receive any training before the experiment, however auditory feedback conditions were

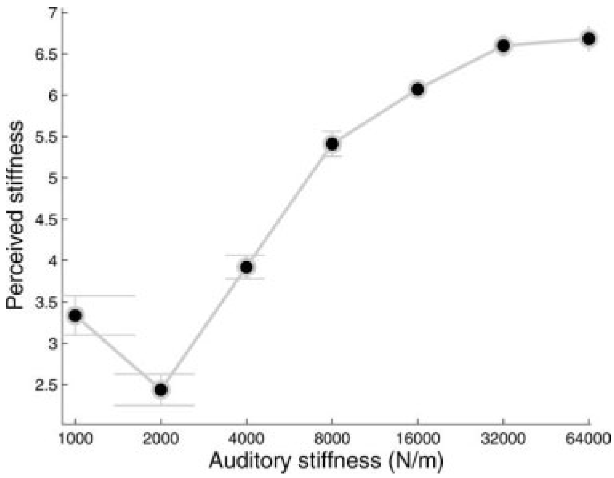


Figure 5. Linear plot of perceived stiffness judgments. Each point represents the average perceived stiffness for each subjects, and for each value of k_A.

presented with the following internal organization (not known to the subjects): first the seven stiffness levels were presented once each, then they were presented again three times each, and the 21 (level × repetition) combinations were randomized. In this way, the first seven conditions provided subjects with a minimal hidden training phase.

Results

Eighteen subjects (between 19 and 30 years old) participated in the experiment. All participants reported themselves as being right-handed, and as having both normal hearing and normal tactual/motoric capabilities in their hands. All of them were naive as to the purposes and hypotheses of the test, and all of them volunteered.[§]

One-way ANOVA on the mean magnitude estimates showed that the effect of the auditory stiffness level was statistically significant ($F = 122.87$, $p < 0.001$). Results are summarized in Figure 5. On average, subjects identified the increase in stiffness with good accuracy, especially in the range 2000–32 000 N/m, although near the extremal values the judgments are clearly less accurate.

In a post-experimental interview, subjects were asked questions about the test. The answers reveal that every subject's judgment was influenced, at least partially,

§A video of an experimental session is available at http://www.dei.unipd.it/~avanzini/papers.html.

by sound, but remarkably about 40% of the subjects perceived the haptic feedback changing together with audio and based their rating also on haptic feedback (although the haptic stiffness had the same value in all conditions). These results suggest that properly designed and synchronized contact sounds can elicit an auditory-haptic illusion and modulate the haptic perception of stiffness. Although the range of object stiffness that can be haptically displayed is limited by the force-bandwidth of the haptic device, the range perceived by the subject can be effectively increased through auditory feedback.

As already noted, the synthetic stimuli used in this study differ only in the values of k_A, while the modal parameters associated to the struck object are constant. Moreover, as described above, the modal parameters of the struck object were chosen to lie between values typical for wood and glass, in order not to provide a clear perception of material. As a result, the auditory cues associated to variations in stiffness are very subtle.

Conclusion

We have proposed a multimodal rendering architecture that integrates physically based sound models with haptic and visual rendering. An example of implementation of the architecture has been presented, which realizes bimodal rendering of contact stiffness. The proposed rendering scheme allows tight synchronization of the two modalities, as well as a high degree of interactivity and responsiveness of the sound models to gestures and actions of a user.

The findings from the experiment reported in the last section support the effectiveness of auditory feedback in modulating haptic perception of stiffness. Interestingly, a relevant portion (about 40%) of the subjects remarked in their answers to the post-experimental interview that they perceived variations in the haptic stiffness, although in every experimental condition the haptic stiffness had the same value. These results suggest that auditory cues can be successfully used to augment and modulate the haptic display of stiffness.

ACKNOWLEDGMENTS

This research was supported by the EU Sixth Framework Programme—IST Information Society Technologies (Network of Excellence 'Enactive Interfaces' IST-1-002114, http://www.enactivenetwork.org).

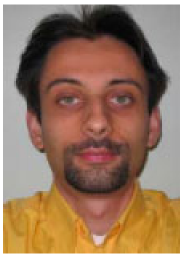
References

1. Welch RB, Warren DH. Intersensory interactions. In *Handbook of Perception and Human Performance: Sensory processes and perception*, Vol. 1. Boff KR, Kaufman L, Thomas JP (eds). John Wiley & Sons: New York, 1986; 1–36.
2. Hahn JK, Fouad H, Gritz L, Lee JW. Integrating sounds in virtual environments. *Presence: Teleoperators and Virtual Environment* 1998; **7**(1): 67–77.
3. van den Doel K, Pai DK. The sounds of physical shapes. *Presence: Teleoperators and Virtual Environment* 1998; **7**(4): 382–395.
4. Avanzini F, Rath M, Rocchesso D, Ottaviani L. Low-level sound models: resonators, interactions, surface textures. In *The Sounding Object*, Rocchesso D, Fontana F (eds). Mondo Estremo: Firenze, 2003; 137–172.
5. DiFilippo D, Pai DK. The AHI: An audio and haptic interface for contact interactions. In *Proceedings of ACM Symposium on User Interface Software and Technology (UIST'00)*, San Diego, CA, November 2000.
6. Lederman SJ. Auditory texture perception. *Perception* 1979; **8**(1): 93–103.
7. Lederman SJ, Klatzky RL, Morgan T, Hamilton C. Integrating multimodal information about surface texture via a probe: Relative contribution of haptic and touch-produced sound sources. In *Proceedings of IEEE Symposium Haptic Interfaces for Virtual Environment and Teleoperator Systems (HAPTICS 2002)*, Orlando, FL, 2002, pp. 97–104.
8. Guest S, Catmur C, Lloyd D, Spence C. Audiotactile interactions in roughness perception. *Experimental Brain Research* 2002; **146**(2): 161–171.
9. McGee MR, Gray P, Brewster S. Mixed feelings: Multimodal perception of virtual roughness. In *Proceedings of International Conference on EuroHaptics*, Edinburgh, July 2002, pp. 47–52.
10. DiFranco DE, Beauregard GL, Srinivasan MA. The effect of auditory cues on the haptic perception of stiffness in virtual environments. In *Proceedings of the ASME Dynamic Systems and Control Division*, (DSC-Vol.61), 1997.
11. Freed DJ. Auditory correlates of perceived mallet hardness for a set of recorded percussive events. *Journal of the Acoustical Society of America* 1990; **87**(1): 311–322.
12. Giordano B. *Sound source perception in impact sounds*. PhD thesis, Department of General Psychology, University of Padova, Italy, 2006.
13. Zölzer U (ed.). *DAFX—Digital Audio Effects*. John Wiley & Sons: Chichester Sussex, UK, 2002.
14. Cook PR. *Real sound synthesis for interactive applications*. A. K. Peters, Ltd: Natick, MA, 2002.
15. Gaver WW. What in the world do we hear? An ecological approach to auditory event perception. *Ecological Psychology* 1993; **5**(1): 1–29.
16. Gaver WW. How do we hear in the world? Explorations of ecological acoustics. *Ecological Psychology* 1993; **5**(4): 285–313.
17. Novint Tech. The interchange of haptic information. In *Proceedings of Seventh Phantom Users Group Workshop (PUG02)*, Santa Fe, October 2002.
18. Sensegraphics. Website. http://www.sensegraphics.se.

19. van den Doel K. Physically based models for liquid sounds. *ACM Transactions on Applied Perception* 2005; **2**(4): 534–546.
20. Dobashi Y, Yamamoto T, Nishita T. Real-time rendering of aerodynamic sound using sound textures based on computational fluid dynamics. In *Proceedings of ACM SIGGRAPH 2003*, San Diego, CA, July 2003, pp. 732–740.
21. Adrien J-M. The missing link: Modal synthesis. In *Representations of Musical Signals*, De Poli G, Piccialli A, Roads C (eds). MIT Press: Cambridge, MA, 1991; 269–297.
22. Klatzky RL, Pai DK, Krotkov EP. Perception of material from contact sounds. *Presence: Teleoperators and Virtual Environment* 2000; **9**(4): 399–410.
23. O'Brien JF, Shen C, Gatchalian CM. Synthesizing sounds from rigid-body simulations. In *Proceedings of ACM SIGGRAPH 2002*, San Antonio, TX, July 2002, pp. 175–181.
24. Avanzini F, Serafin S, Rocchesso D. Interactive simulation of rigid body interaction with friction-induced sound generation. *IEEE Transactions of Speech Audio Process* 2005; **13**(6): 1073–1081.
25. Salisbury K, Conti F, Barbagli F. Haptic Rendering: introductory concepts. *IEEE Computer Graphics and Applications* 2004; **24**(2): 24–32.
26. Wu W-C, Basdogan C, Srinivasan MA. Visual, haptic, and bimodal perception of size and stiffness in virtual environments. In *Proceedings of the ASME Dynamic Systems and Control Division*, (DSC-Vol.67), 1999.
27. Kuchenbecker KJ, Fiene J, Niemeyer G. Improving contact realism through event-based haptic feedback. *IEEE Transactions on Visualization and Computer Graphics* 2006; **13**(2): 219–230.
28. Colgate JE, Brown JM. Factors affecting the Z-width of a haptic display. In *Proceedings of IEEE International Conference on Robotics & Automation*, San Diego, May 1994, pp. 3205–3210.
29. Hunt KH, Crossley FRE. Coefficient of restitution interpreted as damping in vibroimpact. *ASME Journal of Applied Mechanics* 1975; **42**: 440–445.
30. Avanzini F, Rocchesso D. Physical modeling of impacts: theory and experiments on contact time and spectral centroid. In *Proceedings of International Conference Sound and Music Computing (SMC06)*, Paris, October 2004, pp. 287–293.
31. Bresciani J-P, Ernst MO, Drewing K, Bouyer G, Maury V, Kheddar A. Feeling what you hear: auditory signals can modulate tactile tap perception. *Experimental Brain Research* 2005; **162**(2): 172–180.
32. Hötting K, Röder B. Hearing cheats touch, but less in congenitally blind than in sighted individuals. *Psychological Science* 2004; **15**(1): 60–64.

Authors' biographies:

Federico Avanzini is assistant professor of Computer Science at the Department of Information Engineering, University of Padova. During 2001, he worked as a Visiting Researcher at the Laboratory of Acoustics and Audio Signal Processing, Helsinki University of Technology. In 2002 he received a Ph.D. degree in Computer Science at the University of Padova, with a project on sound and voice synthesis by physical modeling. His research interests currently include physically based sound synthesis and processing techniques, audio rendering in multimodal human-computer interfaces, voice production, and articulatory modeling.



Paolo Crosato received the Master's degree in Computer Engineering from the University of Padova, Italy, in 2006 with a thesis on an audio-haptic platform for contact interactions. He is currently working at the Center for Computational Sonology, Department of Information Engineering, University of Padova, with a post-graduate fellowship.