Running head: SOUND AND TOUCH

Sound and Touch

Johan F. Hoorn

vrije Universiteit, Amsterdam, the Netherlands

http://www.businessworldindia.com/archive/010122/stephen%20hawking.jpg, Feb. 6, 2003

Correspondence Address:
Johan F. Hoorn
Faculty of Sciences
Department of Computer Science
Section Information Management & Software Engineering
Sub section Human Computer Interaction, Multimedia & Culture
Tel.: +31-(0)20-444 7614
Fax: +31-(0)20-444 7728

jfhoorn@cs.vu.nl

Abstract

This Chapter tries to explain that sound depends on colliding matter and can be described in terms of frequency and amplitude. It explains how antisound can be used to create silence, what echolocation is, and that the human perception of sound is logarithmic. Medical and military applications and home entertainment are discussed as an illustration. Then we turn to music, discuss what music scales and chords are, when music is considered 'beautiful,' and which areas of the brain are involved in perceiving music. MIDI systems and CD-ROMs are discussed for their relation to digitized music. Speech as the ultimate tool of sound communication is discussed for applications ranging from car navigation to lie detection. You will exercise with creating auditory icons and earcons. The second part of this Chapter deals with the concept of touch and how sound fits in to that. You will get introduced to an intelligent touchscreen that changes handwriting into digital text, a mouse that senses more than clicks, haptic devices in virtual reality applications, domotica, and you will exercise with creating hapticons.

Sound

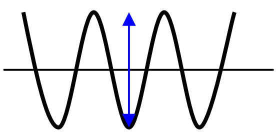
In the previous Chapters, we have focused on information that travels by light (reading, visuals, and motion). Light can travel through a vacuum, whereas sound cannot. Sound needs matter (gas, liquids, or solids) to be heard. Unlike light, sound depends on particles of matter that bang against each other and thus transmit energy. When you pluck the strings of an electric guitar or when you play a wav file, the string and the speaker box membranes push against the air molecules. A pattern of dense and thin layers of molecules is transmitted through the air corresponding to higher and lower air pressures, respectively. That is why you literally can 'touch' the sound or feel the basses in your bowels at a concert. Another difference with light is that light travels at 300,000 km/sec., which is about a million times faster than sound (± 331 m/sec. at 0°C; 343 m/sec. at 20°C; 354 m/sec. at 40°C). Radio waves are as fast as light and although they can transport sound very well, they themselves are quick current oscillations in electric circuits. Scientists have found all kinds of applications that make use of the particularities of light, sound, and radio to create human information representations.

Question 1: Name a few scanning applications in which sound is translated into visuals. Explain why sound is used for measurement and light is used for representation.

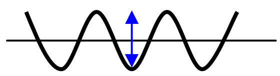
Question 2: Name a communication application in which light is translated into sound. Explain why light is used for transport and sound is used for representation.

Question 3: Why don't astronauts use sound to speak to each other in space? What do they use instead? Why don't they use this solution for visual contact?

Because of the higher molecule density, sound is faster and looses less energy when transported through solid material than through gas. In your room, sound travels through the air with 343 m/sec. In your aquarium, the speed of sound is 1,500 m/sec., and when you pound on the metal pipes of the central heating, your neighbours will be annoyed at the speed of 6,000 m/sec. The formula for (sound) wave speed (m/s) = frequency (Hz) * wavelength (m). In principle, speed is the same for every sound and depends only on the matter that is transmitting it (density and heat). It is the shape of the waves that makes different sounds. The psychological effect of loudness is the result of *amplitude* (expressed in dB). The amplitude of a wave is the height of the peak or trough relative to the baseline. High amplitude creates a perceptually 'hard' sound, low amplitude a 'soft' tone.



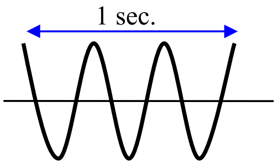
High amplitude
= 'hard' effect



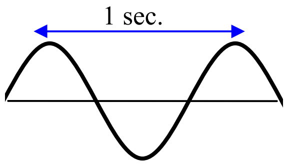
Low amplitude
= 'soft' effect

If the amplitude of a sound wave gets higher, the changes in air pressure become bigger and the sound is perceived as stronger. Thus, the perception of sound volume depends on the amount of energy a sound wave is carrying.

The perception of high and low sounds is the result of *frequency* or number of oscillations per second (expressed in Hz). Each hair cell in the inner ear registers its own frequency. Actually, sound frequency is the number of wave peaks or troughs that is counted during one second. The wave peaks of a high-frequency sound lie closer together than those of a low-frequency sound wave and they reach the eardrum shorter after each other. A wave peak is an area where the matter is more dense (pushed together more strongly) than in the trough of a wave where the molecules lie at greater distance from each other. Wavelength is the distance between two subsequent peaks or troughs. A wave with a low frequency has a long wavelength. A wave with a high frequency has a short wavelength. High-frequency and therefore, short sound waves are perceived as having a high pitch. Low-frequency and therefore, long sound waves are perceived as having a low pitch. All materials have a natural vibration or resonance frequency. If the frequency of a sound is similar to the resonance frequency of a substance, the substance absorbs energy from the sound wave and starts trembling along.

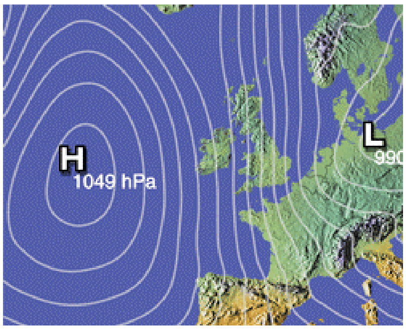


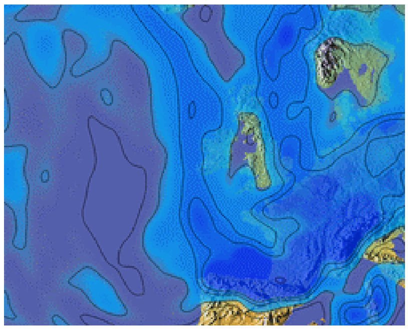
High frequency
= ‘high’ pitch



Low frequency
= ‘low’ pitch

Exercise 1: Try to be creative and represent a visual weather chart in terms of ‘areas of sound,’ so that the acoustic chart becomes understandable to visually handicapped people. In your Acoustic Isobar Chart, make sure that high and low air pressure areas correspond to qualities that are inherent in sound waves as well. In your Acoustic Rainfall Chart, how would you represent different concentrations of rain? And how a thundercloud?





Weather Isobar Chart (left) and Weather Rainfall Chart (right) (http://www.met.ie/forecasts/atlantic.asp Jan. 29, 2003)

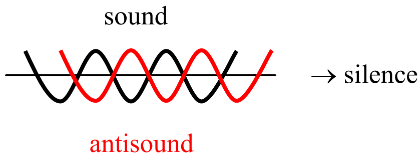
Question 4: Why can you hear that a sound comes from the left or right, given that all amplitudes are equal? What problem gives that for stereo or surround sound applications with headphones?

Question 5: Put sound to your motion visualization. Imagine you programmed a car simulation in which the user should aptly respond to a fire engine coming from the left and driving to the right with sirens on. How would you represent the Doppler effect? The pitch of the siren depends on whether the fire engine approaches or leaves. When approaching, the sound waves come quickly after each other. When leaving, the sound has long wavelengths. Describe the features of the sounds that you should present to give the user a lifelike sound experience of this traffic situation and motivate why.

Exercise 2: Personnel on the bridge of oil tankers see their environment only through electronic devices. Yet, the conditions of the sea are very important to know where to start slowing down (e.g., at the coast of Spain to end in Rotterdam harbor). Design an audio-visual representation of a calm sea with small waves that follow each other quickly and of a sea with slowly rolling high waves.

The difference between soft tones (a computer bleep) and extremely hard sounds (e.g., an exploding space shuttle) is logarithmic. In the decibel scale, adding 10 dB relates to multiplying the amplitude by 10. After 120 dB, the human hearing mechanism gets damaged at least for certain frequencies. You should keep in mind, then, that the sound of headsets (e.g., in free hands car kits or VR helmets) is blasted directly into the ears from metal devices, which means that sounds are perceived as hard and fast. Ergonomists use a criterion of acceptable peak amplitudes of 110 dB and an accepted average 90 dB over the complete workday. To reduce noise in the work environment, the industry (e.g., Signum Data) tries to ban the noisy fans and ventilators from large GHz processors by using so called heat-pipes and new power systems that do not need cooling. Other noisemakers such as the harddisk and CD/DVD drives are still waiting to be silenced.

Another way of silencing machines is the use of antisound. Because sounds are composed of frequencies and amplitudes you can add them up. If you add the peak of one wave to the trough of another (with equal amplitudes and frequencies) the net effect is zero, that is, no sound. A computer can actively generate such interference patterns, which of course, is a matter of exact timing. *Nature* online (Feb. 6, 2002) reports that antisound can diminish the noise inside a car by as much as 6 dB. Antisound is also used in medical scanning devices (fMRI).



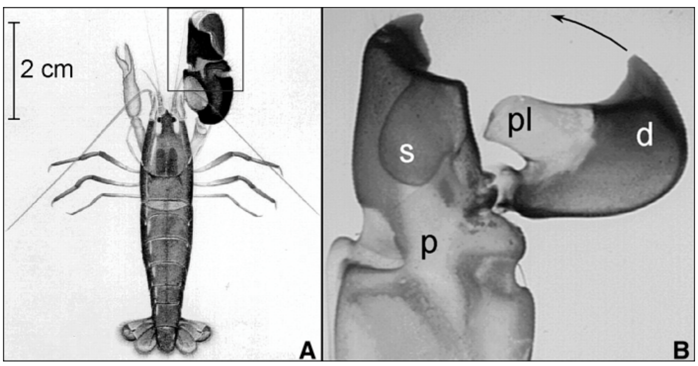
Muting can also be reached by absorbtion. Soft and irregular surfaces absorb more sound energy than hard and smooth surfaces, which return the sound as an echo. Although sound reflections like an echo go into multiple directions, their pitch remains the same. That is why producers mix in some echo for pop stars with feeble voices. The effect on the listener is that the voice is more voluminous.

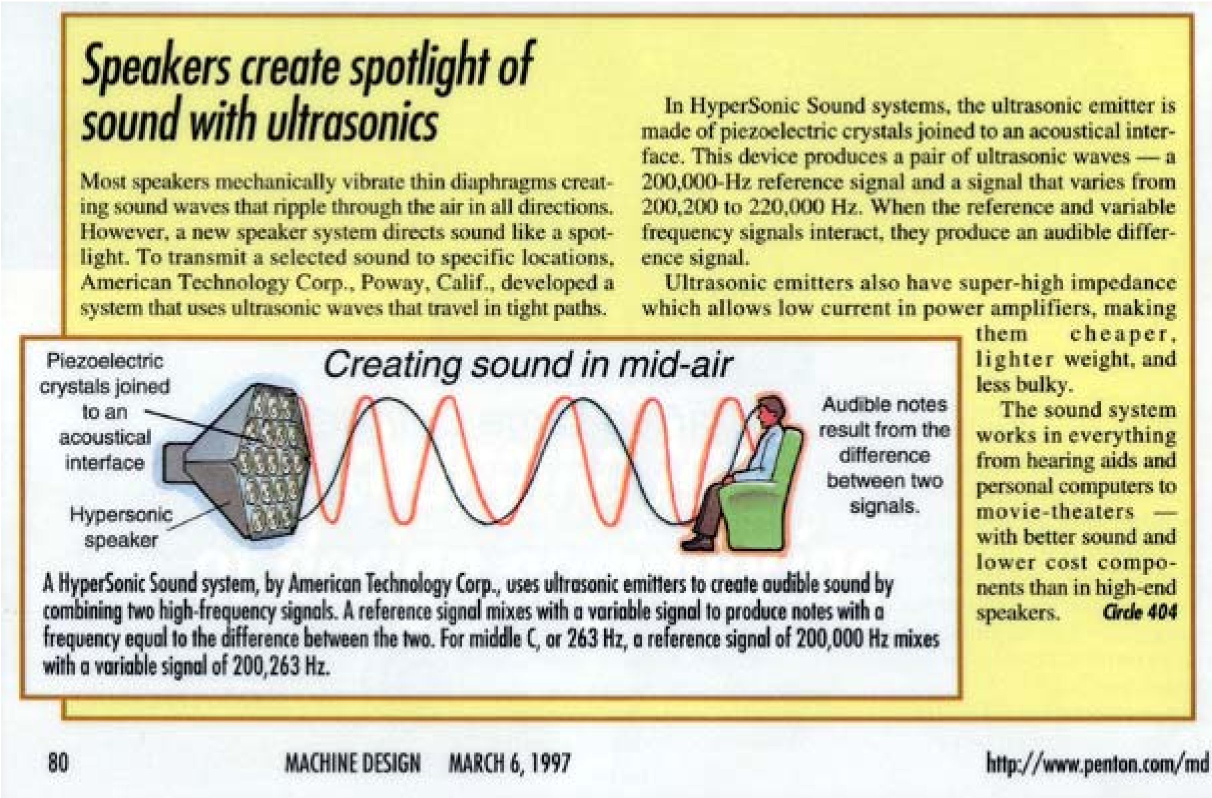
Echolocation or sonar works by using the formula: Sound speed = 2(distance to object) / (t_1 - t_2), where t_1 stands for the time of producing the sound and t_2 for the time it returns. The distance to the object is doubled because the sound travels back and forth. Echolocation or sonar is often done with ultrasounds; sound waves with a frequency above 20,000 Hz, which cannot be heard by humans. In hospitals, computer scans are often made by ultrasounds because they are not harmfull to, for instance, the foetus or the brain (as opposed to X-rays). The stronger the echo of ultrasounds, the harder the material that reflected it. From the reflection pattern built up by the computer, the doctor infers to have found a bone structure or some softer tissues, the submarine captain to approach a rock or a whale, and the whale without computer help to have found a submarine or a mating partner.

Question 6: Non-destructive tests are performed with ultrasound echolocation to detect invisible fissures in critical surfaces such as aircraft wings. The ultrasounds are reflected by the cracks, which are then displayed on a monitor. Why are objects put under water to undergo such non-destructive tests?

Echolocation can also be used at home. Bose developed the ADAPTiQ audio calibration system, which is said to analyze and fine-tune the sound to get optimal quality everywhere in the room. What you hear depends on how your room is furnished. Curtains, boarding, sofa, and the listener's position have effect on how you hear what. Moreover, many surround sound systems have a 'sweet spot,' the place where you get the best sound. If you do not have the money to buy such an advanced home theater system, there is software that employs test sounds to calibrate the speakers yourself (e.g., Ovation's Avia and Video Essentials). You can also use DVDs that produce test signals but placing the speakers right and finding the proper volume is a job that asks a lot of skill. Additionally, you need extra equipment such as a sound-pressure meter. (http://popularmechanics.com/technology/audio/2002/8/for_your_ears_only/print.phtml, Feb. 6, 2003).

Woody Norris of American Technologies invented a sound spotlight that directs ultrasonic waves to a point in space. This way you can send music right to the point where you sit (*Machine Design, 6,* 1997). A military application is that the enemy can be paralized with a directed sound blast without killing the target and without hurting the user. This use of a sound blast resembles the way the snapping shrimp stuns its prey with a sound and light burst of extreme intensity from its claw. (http://stilton.tn.utwente.nl/shrimp/physics.html, http://www-ccrma.stanford.edu/~cc/sjma/shrimp.jpg, Feb. 11, 2003).





Speakers create spotlight of sound with ultrasonics

Most speakers mechanically vibrate thin diaphragms creating sound waves that ripple through the air in all directions. However, a new speaker system directs sound like a spotlight. To transmit a selected sound to specific locations, American Technology Corp., Poway, Calif., developed a system that uses ultrasonic waves that travel in tight paths.

In HyperSonic Sound systems, the ultrasonic emitter is made of piezoelectric crystals joined to an acoustical interface. This device produces a pair of ultrasonic waves — a 200,000-Hz reference signal and a signal that varies from 200,200 to 220,000 Hz. When the reference and variable frequency signals interact, they produce an audible difference signal.

Ultrasonic emitters also have super-high impedance which allows low current in power amplifiers, making them cheaper, lighter weight, and less bulky.

The sound system works in everything from hearing aids and personal computers to movie-theaters — with better sound and lower cost components than in high-end speakers. **Circle 404**

A HyperSonic Sound system, by American Technology Corp., uses ultrasonic emitters to create audible sound by combining two high-frequency signals. A reference signal mixes with a variable signal to produce notes with a frequency equal to the difference between the two. For middle C, or 263 Hz, a reference signal of 200,000 Hz mixes with a variable signal of 200,263 Hz.

80 MACHINE DESIGN MARCH 6, 1997 http://www.penton.com/md

Exercise 3: Explain in words to a visually handicapped person how a ball bounces in sand and how a ball bounces from a brick wall to a concrete floor. Do this by using acoustic analogies.

Music

All musical instruments make the air vibrate. By modulating frequencies and amplitudes, you get something of a melody. Timbre or resonance of an instrument depends on the way the air vibrates. Humans perceive a simple vibration as a pure tone (e.g., Medieval recorder). Air that vibrates in complicated ways is perceived as a full, grinding sound (e.g., death metal). A musical sound is composed of several waves. The basis is an undertone, which is combined with overtones. The wave with the longest wavelength is the undertone, which humans perceive as a deep sound. The overtones have higher frequencies, increasing from the first overtone to the second, etc. The ratios among the different overtones give an instrument its own what the producer calls 'sound.'

Question 7: In the 1950s, Robert Moog developed a synthesizer that played one note at a time. Nowadays, computers can generate far more complicated frequency compositions. Compare an acoustic instrument with an electronic synthesizer. Why are synthesized sounds (≠ samples) yet often perceived as more 'flat' and with less 'individuality?'

Gaffurius (1492). *Theorica Musices*, Milan, Italy.

If a number of sounds or notes is played together, this is a chord. Certain sound combinations humans judge as harmonious and other sound combinations are perceived as disharmonious or discordant. Pythagoras (about 550 BC) thought that beauty came from harmony, which in his case, was always related to mathematics. He discovered a relationship between the height of a tone and the length of a string or pipe. Cutting a string in half leads to doubling the frequency of the undertone. The classic Western music scale consists of eight notes (therefore, *oct*ave). An octave is the musical interval between the first and the eighth note of a scale and the highest note has exactly twice the frequency of the lowest note. All notes are at equal frequency distances from each other. However, what is judged as the 'right' chord or the 'proper' music scale also depends on the musical system you cling to. Schönberg's 12 tones system accepts quite different 'chords' than Medieval or oriental systems or Pythagoras. In other words, what is 'good' or 'beautiful' music depends on social and cultural norms, whether you want to be identified with a social group and a set of rules you think is right as well as on individual taste (cf. Pythagoras).

However, culture, social behavior, and taste are not the only important factors. The brain develops from being sensitive to absolute tones (baby) to being sensitive to harmony (as an adult). This process takes place in the auditory cortex, a strip of millions of brain cells between two lobes near the ear. The music center is near the right ear and mirrored by the language center at the left ear. There are 'music cells' that respond to certain sound frequencies and cells that react to particular intervals. The more a child is exposed to music and sound, the more specialized these 'music cells' become in reacting to complex musical patterns.

Young people tend to prefer music with fast rhythms (house, techno) whereas old people enjoy slow tunes better. This preference differentiation has a biological background. Through ware and tear, the pulsation in the brain, that is the firing frequencies of the neurons carrying information, become slower with age. Therefore, babies prefer rhythms of about 160 beats a minute. For adults, this is 80. We like the rhythms that synchronize with the pulsations in our brain. Brain scans revealed that rhythms have effect on the cortex (for the analysis of complex rhythmic structures) but also on the basic ganglia, where body movement is coordinated. No wonder rhythm and dance are so much interrelated. Therefore, a designer of an interactive system with sounds or music in it not only should be aware of the culture and group psychology of the system's stakeholders but also of their developmental state.

Sound and music recordings are based on copying the sound waves. Analog recordings do this by continuous representations that by and large resemble the original waveform (e.g., the relief of a vinyl record, the magnetic ordering of metal particles on a tape). Digital recordings divide the signal in discrete units that are represented by numbers. The waveform is translated into decimal numbers, each number representing the height of the wave at a given time. These numbers are translated into binary code and then translated into smaller or larger holes burned on a plastic disk (CD-ROM), which are separated by smaller or larger flat areas. A laser beam scans the holes (diffusion) and flats (reflection) and this light pattern is translated into an analog sound signal again (Figure 1).

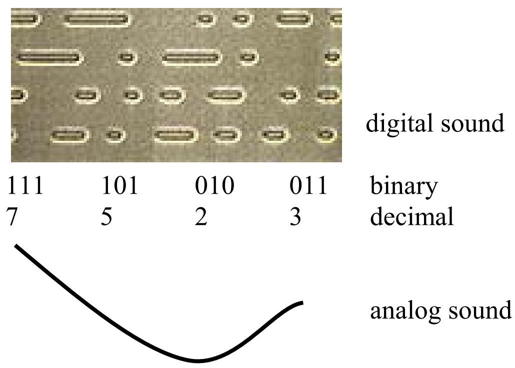
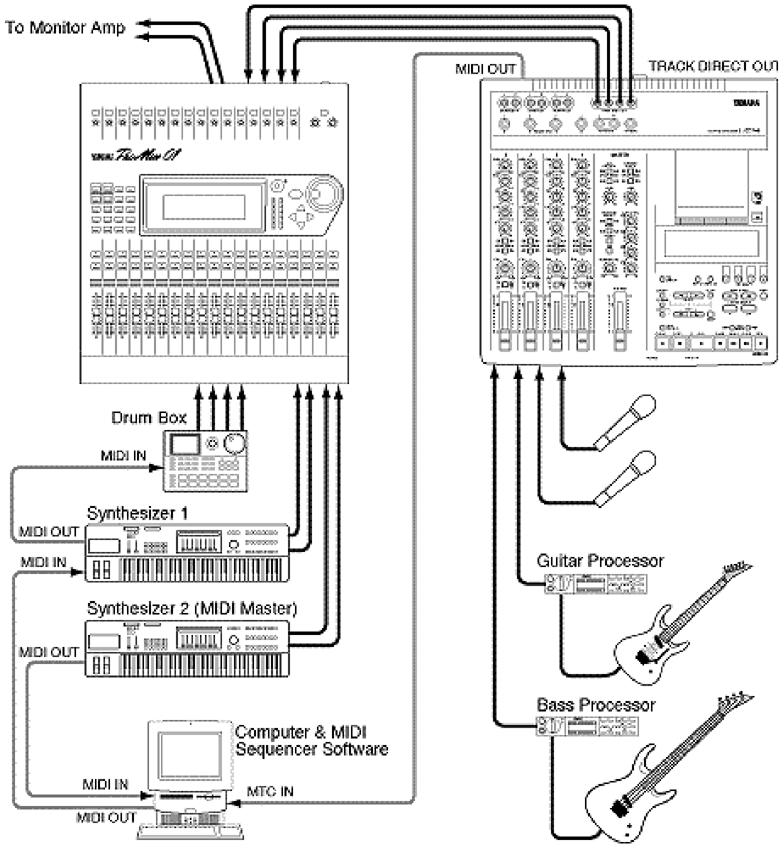


Figure 1. Four representations of a sound wave: Line drawing, decimal numbers, binary numbers, and the relief pattern on a CD.



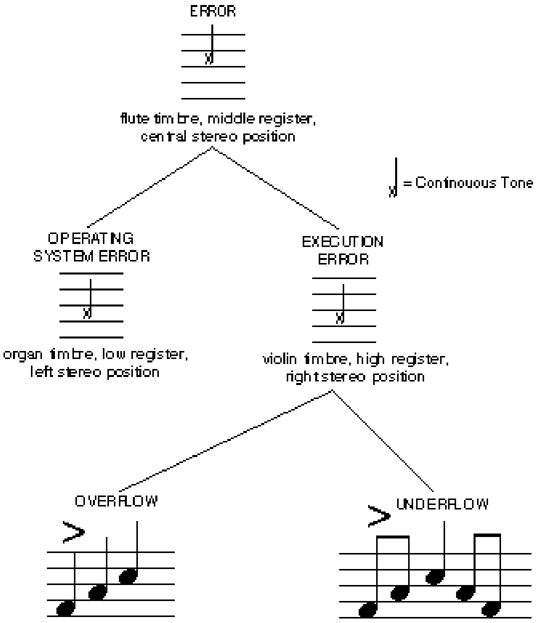
Advantages of using digitized sound, electronic keyboards, synthesizers, etc. in making sounds and music are that you can make completely new sounds and new music. You can increase or decrease the speed, reverse the sounds, and echos can be added easily. With a Musical Instrument Digital Interface (MIDI) you can program a computer to direct and fuse the sounds of multiple electronic instruments (e.g., electric guitar with an electronic drum box). With a sampler, it is possible to record analog sounds (the samples) and insert them into your composition. Because the analog signal is digitized, you can change the numbers to the waveform frequencies and thus, the pitch of the original sound. This way, you can make a breaking glass sing a music scale up to the note that equals the natural resonance frequency of glass, which made it break in the first place.

http://www.yamaha.co.jp/product/proaudio/homeenglish/products/past_pro/multitrack/md4/img/midi.system.gif, Feb. 6, 2003

Stephan Wensveen of Delft TU designs adaptive products on the basis of emotionally rich interactions, that is emotion expressed through behavior. While designing, he keeps three questions in mind: What are the relevant emotional aspects for a context of experience? How can a product recognize and express these aspects? How should the product adapt its behavior to the person on the basis of this information? His approach is illustrated by the design of an alarm clock that plays music in the morning adapted to your mood. Wensveen discovered that the mood of sleeping in is about the same as the mood of waking up but that the number of sleeping hours also has an effect on mood. Therefore, he made a clock with twelve levers, which you should adjust in the evening. By the way you handle the levers (e.g., easy or with force, one at a time or all at once), the device analyzes your mood. The system supposedly recognizes four emotions: Happy, angry, satisfied, or depressed. On the basis of handling the levers and the number of hours you have slept, it chooses the music that wakes you up. If that choice is wrong, you can hit the clock so that it will change its tune. (http://www.delftintegraal.tudelft.nl/info/index.cfm?hoofdstuk=Artikel&ArtID=3996, Feb. 6, 2003).

Auditory Icons and Earcons

Auditory icons and earcons are sound messages that represent certain functions in an interface. An auditory icon uses sound and an earcon music to represent meaning. An auditory icon relates features of the system to things that make that sound naturally (e.g., the dial-up connection represented by a telephone ring). In that way, users can listen to the interface sounds as they would to natural sounds (Gaver, 1997). An earcon is composed of an abstract, nonverbal, synthetic tone or a sequence of tones that represent a message, for example, about how to interact with the computer (Blattner, Sumikawa, & Greenberg, 1989). On the left, you see an illustration of an earcon hierarchy of sounds for representing errors (http://www.dcs.gla.ac.uk/~stephen/earcon_guidelines.shtml, Feb. 7, 2003).



Lemmens et al. (2001) found that it is not indifferent to use earcons or auditory icons, particularly in time-critical applications such as airline control systems or nuclear plants. For example, using a C-major chord (earcon) or a rooster crying (auditory icon) to represent a burning fuse slows down decision time because the meanings of the representation and the event are incongruent. The sound signal then serves as a distractor. Faster reaction times are generated if the representation is meaning congruous, neutral or silent. Thus, if you want to support decisions with an auditory icon or earcon make sure there is meaning congruency (e.g., rooster cry for wake-up call of alarm clock).

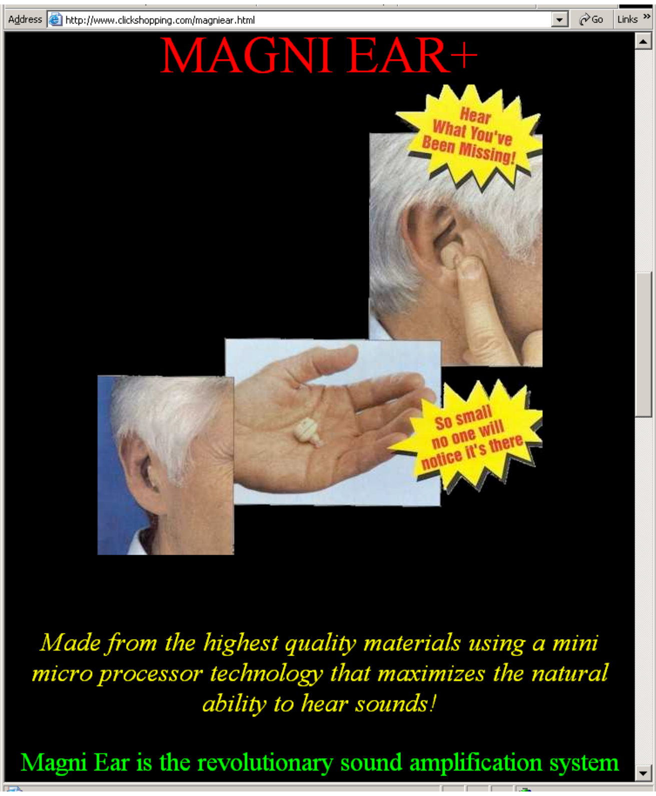
Exercise 4: The ICA 32-bit Windows Client (Version 4.20.727) of Citrix supports a limited scripting language, which can be used to negotiate asynchronous, X.25, or dial-up connections to a server. To configure the scripting, the scripting drive for the ICA Clients supports, among others, the script command BEEP, providing a bell sound as an auditory icon for debugging purposes (http://softserv.murdoch. edu.au/pub/mswin/Citrix/Citrix4.21.779/README1.TXT, Feb. 6, 2003). In all its simplicity, taking a bell sound is a good design choice. Why is a bell sound appropriate both as a bug warning and as related to a dial-up connection? Try to be creative and design an auditory icon or earcon that can be used as a warning and highlights the original meaning of the word 'bug' (the CD cover is a hint).

Wanna know more about digital music making?

The School for Audio Engineering (www.sae.edu) offers courses on Music Production. Here are some of the key words you can use for searching the Web: Digital audio, mixing consoles, simple studio signalflow, studio wiring, multitracking, FX (effects), dynamics processors, patchbays, MDI, Logic, soundmodules, sampling, harddisk recording, Cubase, ProTools, mastering theory, CD, MD, MP3, DVD.

Speech

Speech is the ultimate form of human communication through sound. The human hearing range lies between 20 and 20,000 Hz as a child and deteriorates with age (± 12,000 Hz at 60 years). Yet, our ears are most sensitive to sounds of about 1,000 Hz, which is the frequency area of normal speech (min. 50 Hz - low voice, max. 10,000 Hz - squeaking).



Question 8: MagniEar+ advertises hearing devices with microprocessors that amplify the sounds. If this is the only thing it does, what is the drawback of merely increasing volume? For what sound aspect can a microprocessor be used better to help people with limited hearing range?

In the 1960s, technologists expected that soon we could talk to our computers, have them make appointments with our friends, and let them control the home environment, for example, with a speech-recognizing thermostat. We are not that far yet but certain areas are progressing. For example, providers of telematics products (e.g., NEXIQ Technologies) enter into joint ventures with providers of natural language interfaces (e.g., Fonix) for wireless and mobile devices. Technologies and products, such as Text-To-Speech (TTS) and Automatic Speech Recognition (ASR) are important to, for example, In-Vehicle-Information Systems (IVIS). "IVIS enables vehicle manufacturers to develop, integrate, test, and maintain on-board information and display systems. In addition, it has the ability to manage how information is viewed in the vehicle. It can also use information from various vehicle components as well as from voice, text, and navigation applications to accurately deliver real-time information to the driver." (http://www.mobileinfo.com/News_2001/Issue36/Nexiq_Fonix.htm, Feb. 7, 2003). However, IVIS does not only provide navigational aids and real-time traffic advisories, it also includes applications such as e-mail, news access, and travel information.

Psycholinguist research points out that it makes a difference what type of voice you choose for your talking machine. For example, listeners judge people with a regional accent or dialect as less capable than speakers of the standard language. Dialect speakers are judged as a little more stupid, innocent, and clumsy but also as more social, well-tempered, and more spontaneous than standard language speakers. This stereotype is based on the idea that someone who speaks the language well probably has a higher education. Dialect speakers would have more community values and therefore, they are supposed to be friendlier (Heijmer & Vonk, 2002).

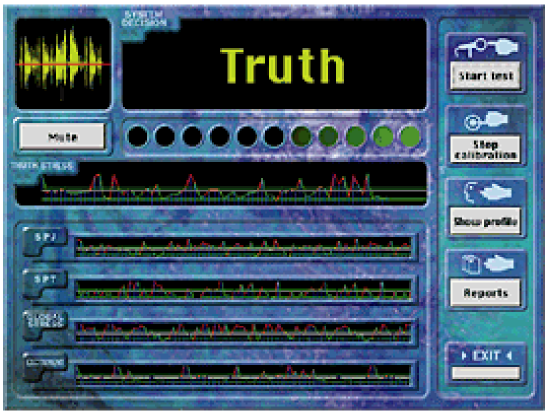
Question 9: Name an application in which you would like a voice that speaks the standard language and one in which you would like a regional accent.

For representing environments under dark circumstances, the vOICe Command Add-On offers speech recognition in blind orientation and mobility. Cameras make images of the environment, which are translated into soundscapes (visible again with spectrographic reconstructions).

> "In mobile applications of The vOICe Learning Edition by blind users, it can be cumbersome to access a keyboard for changing important settings such as mute, fast motion and negative video. In wearable computing applications, a more convenient hands-free user interface is needed. Therefore, The vOICe supports speech input, allowing you to talk to your computer and command it to change settings for a best fit to your changing environment or focus of attention. For instance, when starting to walk you say "speed" into the microphone, and The vOICe automatically switches to a doubled scan rate in order to better perceive movements. When halting to orient yourself, you say "speed" again and The vOICe switches back to its default scan rate for better perception of detail. Moreover, you can say "zoom" to zoom into the central part of the scene, or say "inverse" to better hear dark objects on a bright background via the negative video mode, or say "motion" to hear out moving objects through the automatic motion detection mode. You can even speak "say color" to identify the color of anything in the center of the camera view." (Peter B. L. Meijer, 2003, http://www.seeingwithsound.com, Feb. 11, 2003).

Question 10: As for all mobile applications, you should employ a headset with stereo headphones and a noise cancelling microphone to have better speech recognition in noisy environments. Think of a speech solution for the fact that environmental noise (e.g., someone yelling 'shut up') can be misinterpreted by the system as a command (shut down).

Exercise 5: A major difference with written text is that speech comprehension is even more limited by human memory span. You cannot rewind a spoken sentence without disturbing the natural flow of conversation. Designing representations for people with handicaps is even more difficult. Born deaf people only have visual memory and no linguistic structures available. Their navigation is not conceptual. Blind people mainly have conceptual navigation. Try to design representations for a Web browser for both types of handicaps. Motivate your design decisions.



Question 11: Based on the idea that the voice reflects the mood you are in, Libermann developed a computer program to detect whether someone is lying or speaks the truth. The Israeli border police wanted to use this program for their regular check ups. This lie detection system was available on the Dutch market as well under the name *Truster* (http://www.catchacheater.com/Trust_graph.gif). The system needs a practice set of neutrally spoken sentences of the suspect, which serves as a baseline to detect discrepancies in the test set. The program detects changes in pitch, speaking pace, and trembling of the voice, which supposedly indicate that a statement is true or false. The system was a complete failure. What is the fallacy in calibrating the system to a neutral voice condition? Can you think of other factors that can explain variance in voice pitch, speaking pace, and trembling other than stress by lying? Can you think of certain types of suspects who show no variance in the said variables and yet can lie?

Touch

Actually, the order of discussing sound before touch is the wrong way round but humans are more aware of communication through hearing than through touching. Touch is a word for what are many senses, all based on physical contact. Quite like sound, touch is mediated through pressure on material and conceptualized like this, sound is a form of touching. However, touching is more than that; it is a mixed experience of pressure, pain, warmth, and cold. Touch is extremely important for monitoring the world, including a computer system. If your harddisk is overheating, the only way you can sense this is through touch (feeling the warmth on your skin). While you type on your keyboard and use the mouse, your fingertips are registering the resistance of the keys and buttons and your head calculates how much force is needed and in what direction it should go. Without looking, the small marks on the F and J keys help your orientation in the keyboard layout. Because separate keys represent the letters on your keyboard, the relief of high surfaces and low troughs help you monitor pressing the right key and not two at a time. In fact, if letters are visual representations of sounds, then those visual representations themselves are represented by a haptic system of a relief pattern in a fixed vector space, that is, the keyboard. From the MIT Web site, we learn:

> "In humans, tactile sensing is generally achieved via mechanoreceptor cells located near the surface of the skin, the highest density of which is found in the hand. These mechanoreceptors can perceive vibrations of up to about 300 Hz. Therefore, tactile feedback generally involves high frequency buzzing sensations applied directly to the skin, usually in response to 'contact' between the user and an object. The sensing of forces is more kinesthetic in nature, and is achieved by receptors situated deeper in the body. These mechanoreceptors are located in muscles, tendons and joints and are stimulated by movement and loading. Their stimulus frequency is much lower, lying in the 0-10 Hz range. Therefore, force feedback will consist of artificial forces exerted directly onto the user from some external source.
>
> Thus there are two aspects to the sense of touch; that which provides kinesthetic information and that which conveys tactile information. The kinesthetic information that we perceive about an object are coarse properties such as it's position in space, and whether the surfaces are deformable or spring-like to touch. Tactile information conveys the texture or roughness of the object that we are touching. Both types of 'touching' information must be present in a realistic haptic interface." (http://www.hpcc.ecs.soton.ac.uk/~dtcb98r/vrhap/vrhap.htm#Feedback, Feb. 14, 2003)

Fitts' Law

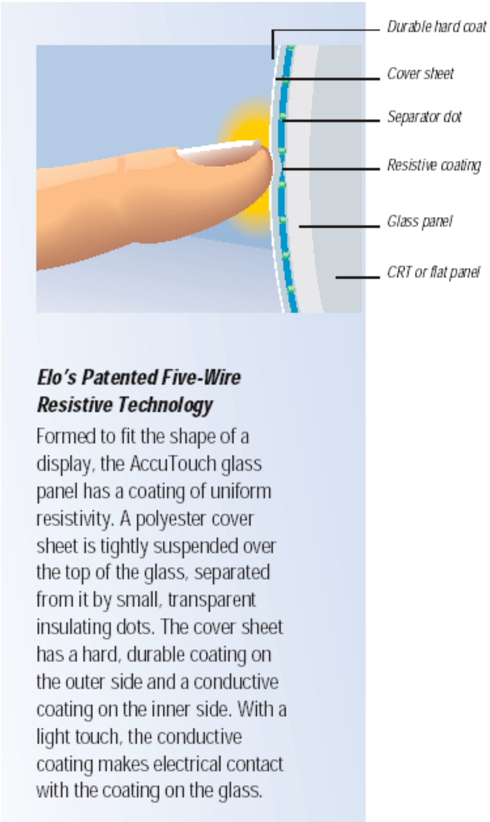
Touching an interface widget requires motion and motion is a function of time and distance. A fundamental principle of interaction design is Fitts' Law (1954), which forms a model of *speeded* human psychomotor behavior, not of slow motion like walking. Fitts showed that the time to reach a target of constant size is a logarithmic function of the *distance* to that target. Additionally, movement time is a logarithmic function of *target size* if distance is held constant: $MT = a + b \log_2(2D/S + c)$, $a + b \sim 100$ msec., where

MT	= time to move to target
a, b	= regression coefficients
c	= constant of 0, 0.5, or 1
D	= distance to or amplitude of target (sometimes referred to as A)
S	= target size (which relates to accuracy)

In other words, the longer the distance the user has to push a pointer to an interface widget (an icon) or a finger towards a touchscreen feature, the more effort it costs. Moreover, the smaller targets cost the most effort. Thus, you could make an interface that positions possible targets in the neighborhood of the mouse location and make them big. With Fitts' Law you can predict the speed, distance, and target sizes that are most convenient for the user's abilities to point. Beware, however, that if you try to make an interface that adapts to the user's needs with features that move around all the time, the user will get confused because his/her mental model of the layout of the interface gets mixed up all the time. Fitts' Law is most applicable to time-critical applications such as distributed multimedia.

Haptic Interfaces

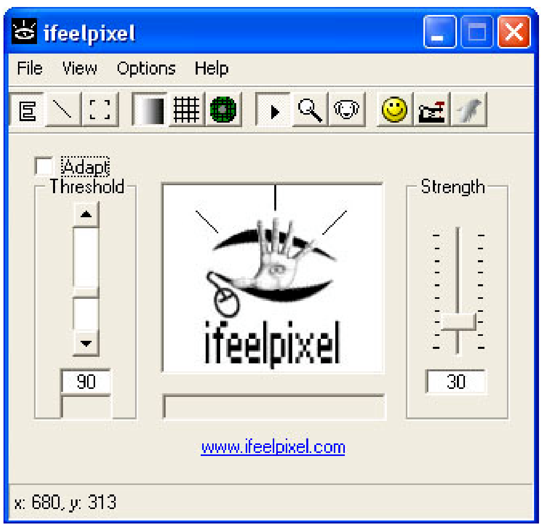
As an alternative to conventional keyboard operation, touchpads, touchscreens, and touch tablets have been developed. Conventional touchscreens use *X* or *Y* axis sensors. AccuTouch has produced a coversheet that works as a voltage-measuring probe to reach a high touchpoint density. This allows recognizing input from fingers, fingernails, a gloved hand, or a stylus. Moreover, even if the surface is scratched, the touchscreen will keep functioning. (http://www.elotouch.com/pdfs/marcom/accute.pdf, Feb. 14, 2003)



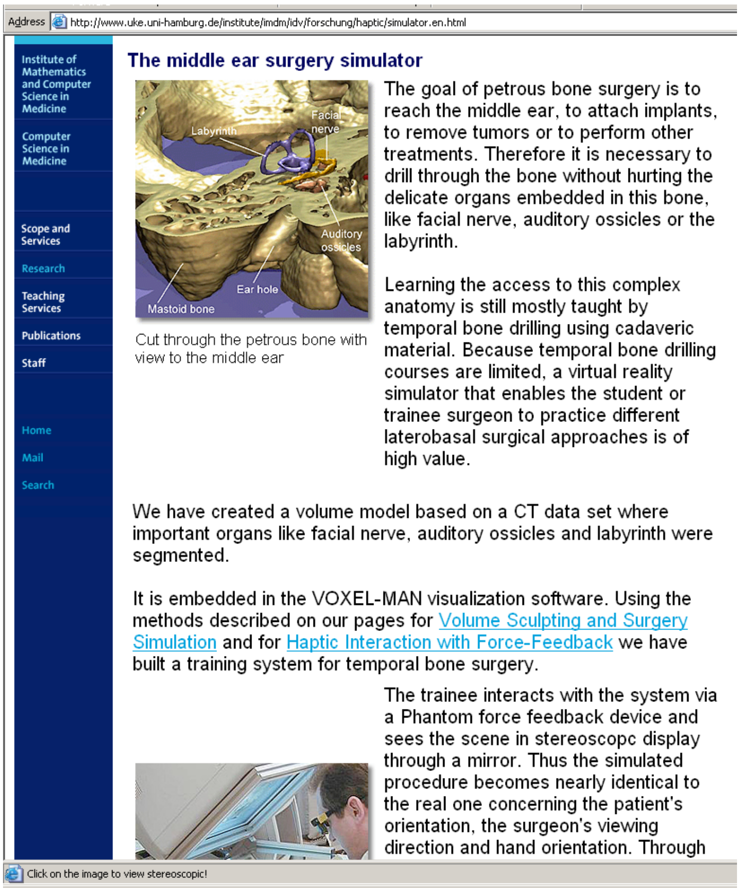
Other technologies such as Smart Board use gas plasma touchscreens. If the user shows, for example, an Excel sheet with the Smart Board, s/he can write on the screen with a whiteboard marker and wipe it out again. However, it is also possible that the handwriting is translated into digital text, which allows the user to fill in the fields by hand. (http://www.smarttech.com/products/plasma/index.asp, Feb. 14, 2003).

In touchscreens and similar devices, touch and position perception are strongly connected. You can only specify positional data by touching the device. Reversely, by touching it you specify a position. Developers at Microsoft saw that if you uncouple touch from position sensing, many new haptic possibilities become available to operate a computer. Hinckley and Sinclair (1999) adapted a Microsoft IntelliMouse Pro and made the Scrolling TouchMouse, which identifies touch in the combined palm/thumb areas and thus knows whether the user handles the machine. Due to five sensors, the mouse is also sensitive to the areas around the wheel. Tapping these areas triggers Page Up and Page Down. This way a new type of interface can evolve (the On-Demand Interface) in which only the necessary toolbars, scroll bars, and other UI widgets appear as indicated by the touch data from the mouse. This saves on occupation of scarce screen space and makes the layout less confusing. A similar device developed by these researchers is the TouchTrackball. (http://research.microsoft.com/ui/touchmouse, Feb. 12, 2003).

Other applications such as Ifeelpixel (http://www.ifeelpixel.com/screenshots/, Feb. 14, 2003) use TouchSense mice to mediate the structures (edges, curves), textures, and colors of the pictures on the display screen. By translating vision into touch, for example, visually impaired users still can have an impression of visuals in a Web site or information kiosk.



Virtual Reality


Address http://www.uke.uni-hamburg.de/institute/imdm/idv/forschung/haptic/simulator.en.html

Institute of Mathematics and Computer Science in Medicine
Computer Science in Medicine
Scope and Services
Research
Teaching Services
Publications
Staff
Home
Mail
Search

The middle ear surgery simulator

Cut through the petrous bone with view to the middle ear

The goal of petrous bone surgery is to reach the middle ear, to attach implants, to remove tumors or to perform other treatments. Therefore it is necessary to drill through the bone without hurting the delicate organs embedded in this bone, like facial nerve, auditory ossicles or the labyrinth.

Learning the access to this complex anatomy is still mostly taught by temporal bone drilling using cadaveric material. Because temporal bone drilling courses are limited, a virtual reality simulator that enables the student or trainee surgeon to practice different laterobasal surgical approaches is of high value.

We have created a volume model based on a CT data set where important organs like facial nerve, auditory ossicles and labyrinth were segmented.

It is embedded in the VOXEL-MAN visualization software. Using the methods described on our pages for Volume Sculpting and Surgery Simulation and for Haptic Interaction with Force-Feedback we have built a training system for temporal bone surgery.

The trainee interacts with the system via a Phantom force feedback device and sees the scene in stereoscopc display through a mirror. Thus the simulated procedure becomes nearly identical to the real one concerning the patient's orientation, the surgeon's viewing direction and hand orientation. Through

Click on the image to view stereoscopic!

Haptic devices also have great use in virtual environments. Previous technology, including keyboards, mice, and data gloves provide only a passive touch interface with the computer. Data Gloves may permit the user to see his or her hand appear in the virtual environment and grasp objects, but does nothing to prevent the user from moving her hand through the object. More advanced data gloves, such as the CyberTouch Glove from Virtex provides a tactile feedback when objects are touched. It accomplishes this by having buzzers that vibrate under the fingertips and palm. This provides the sensation of touching something but it lacks the resistance to motion that is observed when real objects are touched.

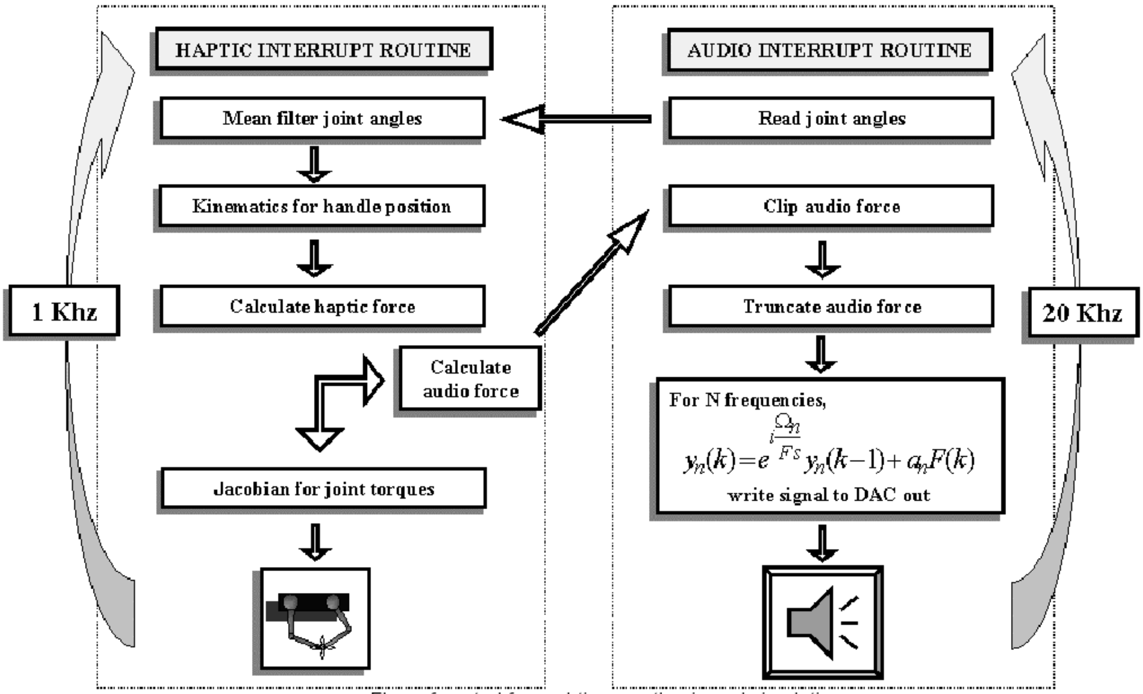
Force feedback is most helpful, for instance, for surgeons at different places who together want to diagnose a fracture of a patient at long distance. You could also think of a sculptor who makes a virtual study before working on a valuable block of marble or a NASA astronaut inspecting the rocky surface of a planet.

To comply with this demand, The PHANToM (Personal Haptic Interface Mechanism) is like an extremely advanced joystick that measures the position of a user's fingertip and feeds a precise force vector back to that finger. Thus, users can manually interact with virtual 3-D objects and the PHANToM can be used for remote control of robots. (http://web.mit.edu/newsoffice/nr/1998/phantomsm.JPG, Feb. 14, 2003)

Yet, Carnegie Mellon University researchers have developed a new type of haptic interface employing magnetic levitation that enables computer users to physically interact with simulated objects and environments on their computer screens. The device is unique because it enables people to not only touch these objects, but to reach in and manipulate them in three dimensions as well. The new system also eliminates the bulky links, cables and mechanisms of current haptic interfaces in favor of a single, lightweight moving part, which floats on magnetic fields. The system has a bowl-shaped floating element containing six levitation coils surrounded by strong, permanent magnets. A protruding handle attached to the bowl is grasped by the computer user, enabling interaction with solid, three-dimensional models graphically depicted on the computer screen. The system is housed in a desktop-high cabinet. (http://www-2.cs.cmu.edu/~msl/haptic/images/berkelhap.jpg, Feb. 12, 2003).

"With a magnetic levitation haptic interface, you can reach into a simulated environment and feel the force and torque of simulated objects," says Ralph Hollis, a principal research scientist at Carnegie Mellon's Robotics Institute. "Early in their development, computers used to be just text. Then came graphics, 3D graphics and speech, involving more and more of a user's senses. Of the last three senses left - smell, taste and touch - the latter will likely be the most useful." (http://www.spie.org/web/oer/september/sep97/magnetic.html, Feb. 12, 2003)

DiFilippo and Pai (2000) go further and argue that users should be able to tap and scrape virtual objects, feel a rough or soft surface and hear the right sound to it. DiFilippo and Pai are developing an audio-haptic interface that makes a tight synchronization between audio and haptic mode (see flowchart). The user should hear the right sound at the right time with the right object. The contact forces the user employs in the simulated environment are fed back by a haptic force-feedback device to the hand and to the ear by the proper contact sound. This is not simply playing back a pre-stored sample or tone but the sound changes while the contact forces change.



Flow of control for real-time synthesis and simulation.

Domotica

You come home from university, and you are sick and tired of the classes you just followed. Probably you can imagine this lifelike scenario quite well and you see yourself let your head hang down while your feet are dragging over the floor. Yet, a magic carpet recognizes this walking pattern as you who is in a terrible mood and it signals the domotica central command system to play some soothing music. Your housemate enters as well and from her walking pattern (pressure areas, pace, motion progression) the carpet recognizes who it is and puts on the television set, switching to CNN, because that is her favorite channel. Another person enters the apartment but this time the carpet does not recognize the walking pattern of the visitor. It sounds the alarm, which is automatically forwarded to the nearest precinct because this might be a burglar. Unfortunately, it is your mom who wanted to welcome you in your new home and brought you a nice gift, the Impressa F90, a lovely designed coffee machine. With you in that terrible mood, the ominous world news on TV, and the police on their way to fetch your mom, you really can use a cup of coffee. The Impressa machine, "touch and go," offers 36 different kinds of coffee, from mild to strong, cappucino to espresso, using different brands and mixes. The sensitive touchscreens guide you step by step through all the settings so that finally you can enjoy the ultimate cup of coffee. Then you can enter your coffee preferences into the PC and the thing will read it out or if you get confused you connect to the Internet for FAQs and help with making coffee with this machine. All these options remind you of the classes you just had where the professor was talking about an article written by Schwartz et al. (2002) in which they explain that the problem of our times is that we have too many choices. Therefore, we keep zapping while watching a program or buy the latest version of a command system merely because it has more possibilities and not because we need them. Schwartz et al. have evidence that certain types of people flourish under such circumstances whereas others suffer. The so-called *satisficers* take the first option that meets their criteria, whereas the *maximizers* only want the best option. Having so many options is problematic to maximizers because you never know whether you selected the best one. Schwartz et al. found that

these people are less happy, less optimistic, less self assure, and are more depressed while suffering more from regrets from which you have to conclude that you must be a maximizer. (http://www.bartolomeo.nl/images/juraf90.jpg, Feb. 12, 2003)

Question 12: Analyze what goes wrong with the magic carpet that uses touch sensors to personalize the settings of the domotica system. If the world of users can be divided into satificers and maximizers are 36 types of coffee sufficient or too much of a good thing? Why? What does that do for the menu structure of the touchscreen displays? What do you think about the conclusion that someone must be a maximizer because s/he has all the emotions a maximizer has?

Rightfully upset by so many bad designs that entered your domestic quarters, you resign from using a word processor and take a pen to write a letter of complaint to a consumer's organization. While writing in anger, the thing starts bleeping and blinking, trying to warn you that you press too hard. This is a sensor pen (Vitelec) that has a muscle tension meter built in and is an ergonomic weapon against repetitive strain injury (RSI). The blinking and beeping has to make the writer aware of excessive muscle tension in shoulder and arm so to reduce the chances to get RSI. Due to the sensor pen, RSI related complaints reduced for 13 out of 19 participants in a test at HM Home Office in the Netherlands (*Aaneen*, Nov. 29, 2002, p. 24).

Reid, J. S., Geleijnse, B., & Van Tol, J.-M. (http://www.foksuk.nl/)

Question 13: Although the sensor pen seems to be successful at fighting RSI, can you think of a situation in which you do not want a blinking and beeping device? What do all these smart and intelligent devices (carpet, coffee machine, pen) that try to pamper their users do for the feeling of independence?

Hapticons

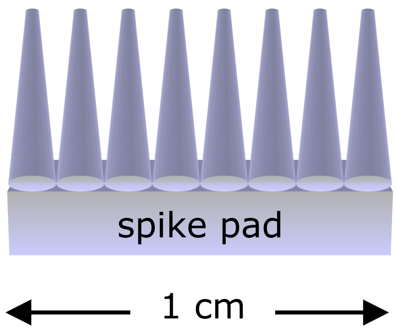
Often the term hapticon designates anything in the interface (a knob, a pointer-attracting visual icon) that allows the *user* to have tactile communication with the computer. To align the word hapticon with the use of visual icons, however, 'hapticon' should be preserved for anything that the *computer* communicates to the user with a fixed tactile sign. Like Braille, a hapticon should be a sign you can touch and that indicates a function of the system.

The creation of haptic icons or hapticons is still in its infancy. Therefore, many of their (im)possiblities are still debated. For example, Dosher and Hannaford (2002) measured the smallest haptic effects that can meaningfully be communicated to a user. They investigated variation in the force displayed to the user, the size, shape, pulse-duration, icon widths, icon waveforms, and static pulse-widths of a haptic feature. These researchers investigated "… the smallest detectable haptics effects with active exploration of saw-tooth shaped icons sized 3, 4 and 5 mm, a sine-shaped icon 5 mm wide, and static pulses 50, 100, and 150 ms in width. Smooth shaped icons resulted in a detection threshold of approximately 65 mN, twice that of saw-tooth shaped icons. In the case of static icons, longer pulse-widths corresponded to slightly higher threshold values."

Exercise 6: Make a haptic sign system (so called hapticons) for an application of your choice while using the indicators high vs. low key resistance, warm vs. cold, and rough vs. smooth surface. Try to make systematic combinations of the touch indicators. As in any language, define a meaning to the representations you created.

In a related study, Dosher (2002) investigated detection thresholds for and use of small haptic effects to improve task performance. The effects of icon representations such as amplitude, shape, and pulse duration on haptic perception were studied with haptic icons that ranged in size from 3 to 5 mm, smooth vs. rough, and static icons of 100 to 150 ms pulse duration. Using Fitts' Law as a measurement of task performance, the effects on user performance with three levels of haptic stimuli between 50 to 300 mN indicated that rough (saw-tooth) haptic icons were more easily detected than smooth (sinusoidal) icons of the same size, by almost a factor of two. Mean subject performance, as measured by Fitts' information processing rate and clicks-per-minute, improved with the size (or amplitude) of haptic stimuli.

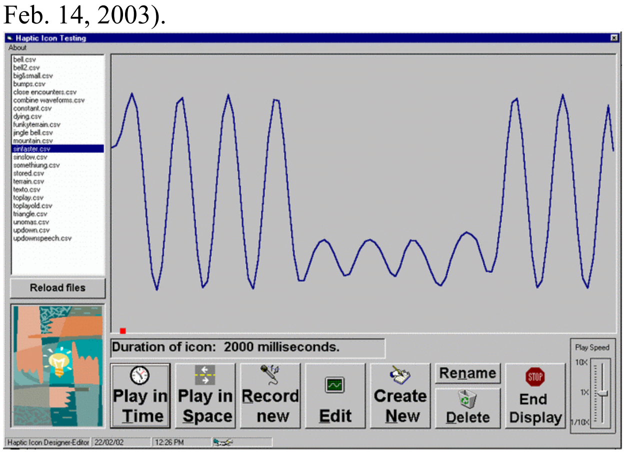
Exercise 7: Invent hapticons as physical signs to facilitate seamless, intuitive communication between a multimedia device and its user. The computer should suggest the functions *rate*, *jump*, *volume*, and *zoom* with a 'spike pad,' which the user can scan with the fingertips. The spikes can go up and down, back and forth, left, right, etc.



Question 14: Apart from hapticons, certain people are busy making smicons. These smell icons are based on smell emission technology (http://web.media.mit.edu/~jofish/writing/smell.as.media.short.paper.pdf, Feb. 14, 2003). Can you think of useful applications and of disadvantages to such signs? To push the limits a little further, shall we ever get to a taste icon (tasticon) in an interface that has different flavors for functions (lick to click)?

Wanna know more about haptic interfaces?

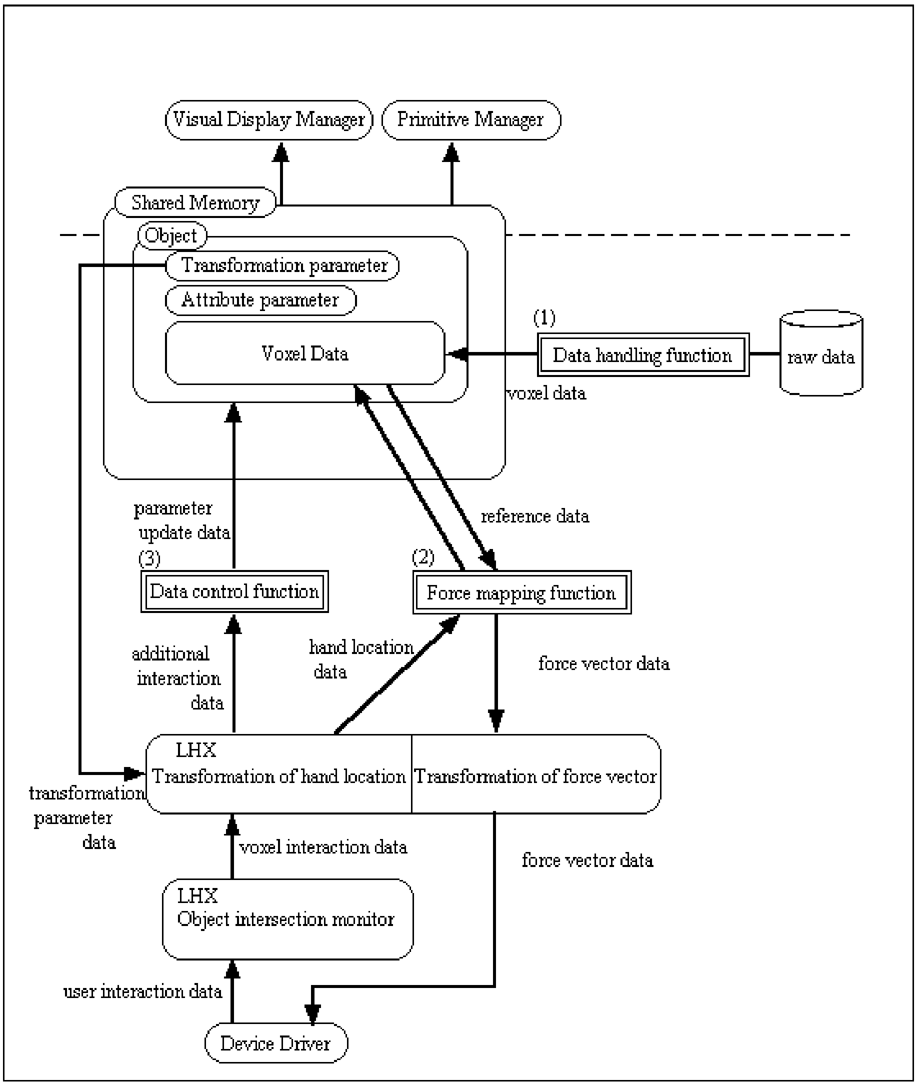
For software to create hapticons, check out http://www.cs.ubc.ca/~danilkis/fusion/projects/icons.html, Feb. 14, 2003).



The Hapticon Displayer-Editor Main Screen

Guidelines for haptic devices are formulated by Oakley, Adams, Brewster, & Gray (2002) at http://www.dcs.gla.ac.uk/~stephen/papers/HCI2002-oakley.pdf, Feb. 14, 2003.

If you want to know more about software infrastructure for haptic interfaces, look in the Library of Haptics, a Versatile Software Platform for Visual/Haptic Environments created by Wataru Hashimoto and Hiroo Iwata at (http://intron.kz.tsukuba.ac.jp/vrlab_web/LHX/icat97.html, Feb. 14, 2003)



Structure of Volume Haptics Library (Hashimoto & Iwata)

References

Blattner, M., Sumikawa, D. and Greenberg, R. (1989). Earcons and icons: Their structure and common design principles. *Human Computer Interaction 4, 1,* 11-44.

DiFilippo, D., & Pai, D. K. (2000). The AHI: An audio and haptic interface for contact interactions. *Proceedings of the 13th annual ACM symposium on User interface software and technology 2000, San Diego, Ca, US* (pp. 149-158). New York, NY: ACM Press.

Dosher, J. (2002). *Detection Thresholds and Performance Gains for Small Haptic Effects.* MSEE Thesis, University of Washington, Department of Electrical Engineering, December 2002.

Dosher, J. L. G., & Hannaford, B. (2002). Detection thresholds for small haptic effects. In M. L. McLaughlin, J. P. Hespanha, & G. S. Sukhatme (Eds.), *Touch in Virtual Environments: Haptics and the Design of Interactive Systems*. Prentice-Hall.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, *47*, 381-391.

Gaver, W. W. (1997). Auditory interfaces. In M. G. Helander, T. K. Landauer, & P. V. Prabhu, (Eds.), *Handbook of Human-Computer Interaction, 1,* (pp. 1003-1041), 2nd ed. Amsterdam: Elsevier.

Heijmer, T., & Vonk, R. (2002). Effecten van een regionaal accent op de beoordeling van de spreker. [Summary effects of a regional accent on the judgement of the speaker.] *Nederlands Tijdschrift voor de Psychologie en haar Grensgebieden, 57, 4,* 108-113.

Hinckley, K., Sinclair, M. (1999). Touch-sensing input devices. In *Proceedings of the SIGCHI conference on Human factors in computing systems: The CHI is the limit 1999, Pittsburgh, Pennsylvania, USA* (pp. 223-230). New York, NY: ACM Press.

Lemmens, P. M. C., Bussemakers, M. P., & De Haan, A. (2001). Effects of auditory icons and earcons on visual categorization: The bigger picture. *Proceedings of the 2001 International Conference on Auditory Display, Espoo, Finland, July 29-August 1, 2001.*

Oakley, I., Adams, A., Brewster, S., & Gray, P. (2002). *Guidelines for the Design of Haptic Widgets.* In Proceedings of BCS HCI 2002 (London, UK) http://www.dcs.gla.ac.uk/~stephen/papers/HCI2002-oakley.pdf

Schwartz, B., Ward, A., Monterosso, J., Lyubomirsky, S., White, K., & Lehman, D. R. (2002). Maximizing versus satisficing: Happiness is a matter of choice. *Journal of Personality and Social Psychology, 83, 5,* 1178-1197.