Non-Electrophonic Cyborg Instruments: Playing on Everyday Things as if the Whole World were One Giant Musical Instrument

Steve Mann, Ryan Janzen, Raymond Lo, and James Fung
Dept. of Electrical and Computer Engineering
University of Toronto
http://www.eyetap.org/publications/

ABSTRACT

We introduce a new musical instrument in which computation is used to modify acoustically generated sounds. The acoustically generated sounds originate from real physical objects in the user's environment. These sounds are picked up by one or more microphones connected to a camera phone which filters the sounds using filters whose coefficients change in response to subject matter present in view of the camera. In one example, a row of 12 image processing zones is presented such that sounds originating from real world objects in the first zone are mapped to the first note on a musical scale, sounds originating from the second zone are mapped to the second note of the musical scale, and so on. Thus a user can hit a cement wall or sidewalk, or the ground, and the camera phone will transform the resulting sound (e.g. a dull "thud") into a desired sound, such as the sound of tubular bells, chimes, or the like. Note that the instrument is not an electronic instrument (i.e. not an Electrophone in the Hornbostel Sachs sense) because the sound originates acoustically and is merely filtered toward the desired note. This plays upon the acoustic qualities and physicality of the originating media. For example, if we strike the ground abruptly, the sound resembles that of a bell being hit abruptly. If we rub the ground, the sound resembles that of rubbing a bell. We can scrape the ground in various ways to obtain various sounds that differ depending on which of the camera's zones we're in, as well as the physical properties of the ground itself. These experiences can be shared across "cyborgspace" to effectively blur the boundary between the real and virtual worlds. We present an aquatic instrument that plays upon jets of water, where it is the filter coefficients of the transform that are shared. This allows both users to play the instrument in the jets of water of different public fountains but still experience the same musical qualities of the instrument, and share the physical experience of playing in a fountain despite geographic distances.

MM'07, September 23–28, 2007, Augsburg, Bavaria, Germany.

Categories and Subject Descriptors

H.5.2 [**Info. systems**]: Info. interfaces and presentation—*User Interfaces*; I.4.9 [**Computing Methodologies**]: Image Processing and Computer Vision; J.5 [**Computer applications**]: Arts and Humanities

General Terms

Design, Experimentation, Human Factors, Performance

Keywords

Fluid-user-interfaces, fluid sampling, tangible user interfaces, water-based immersive multimedia, hydraulophones, interactive art, cyborg, aquaborg, waterborg, frolic, frolicious, FUNtain (TM), harmellotron, harmelody, urban beach, urbeach, cyborgfountain (TM), cyberfountain (TM)

1. INTRODUCTION AND GOAL OF PAPER: AQUACOUSTIC AND HYPERACOUSTIC INSTRUMENTS

Much of computer music concerns itself with the generation or composion of music in which the resulting computer-based instrument (or computerized hyperinstrumental extension) would rightly be classified as an electrophone (Hornbostel-Sachs 5th category[4], as currently practiced[1]).

However, computers may also be used for digital signal processing as applied to acoustic instruments, without changing the fundamental categorization of the resulting hybrid. Thus, for example, an electric guitar, whether running through traditional analog guitar effects pedals, or digital effects (e.g. software pedals, waveshapers, etc.) is still a chordophone — the fact that the effects are digital rather than analog (as in a traditional guitar effects pedal) does not necessarily change the hybrid computer plus guitar into an electrophone.

In this paper, we wish to employ the computers in this sense, in order to facilitate the creation of new instruments that are not electrophones (i.e. are not in the extended Horbostel Sachs category 5).

We created a computer-assisted musical instrument from an array of wooden blocks in which a separate microphone was used for each note, and routed through a separate processing channel.

Rather than triggering a sample or MIDI note as might typically be done in computer music, we retained the acoustic property of the instrument by simply passing each of the 19 sound signals through a filter having transfer function $H(f)$, where we computed H based on taking ratios of sound recordings made from real tubular bells and our wooden-block instrument.

1.1 Making a bell-like sound from a dull thud

What we are trying to do is get a dull thud from simple ubiquitous everyday found objects like wood to ring out as clear as a bell, while maintaining all the nuances of how it was struck.

To demonstrate this newly invented instrument in a simple way, we set up a version of it using an array of wooden blocks, each fitted with a separate audio transducer (Fig 1).

Note that the range of expression is much more diverse than merely velocity-sensitive triggering of a recording of a bell sound where amplitude varies with strike velocity. For example, rubbing the sticks against the blocks produces a sound similar to that obtained by rubbing sticks against a real bell.

The wooden blocks can be varied in size so they produce the correct note to begin with, or they can all be the same size (as shown).

Optionally, the audio transducers can be mounted in sticks, mallets, or the like, while an overhead camera allows the computer to see which block is struck. This has the advantage of allowing the computer to slightly modify the transfer function depending on where the block is struck, allowing pitch bend, timbral variation, etc..

With an overhead camera, we can eliminate the need for a separate audio pickup in each block, and instead mount an audio pickup in each mallet or stick, thus reducing the required number of pickups from 19 (one for each block) down to 2 (one for each hand-held mallet), as well as reducing the required number of microphone inputs from 19 down to 2 (thus using a standard stereo sound card rather than a specialized multi-channel analog to digital converter).

With an overhead camera, we can also eliminate the separate blocks, and simply use a single surface as the playing surface, as shown in Fig. 2. The result is a glockenspiel having continuously variable pitch.

For the computer vision we used the Intel OpenCV image library, but any standard computer vision system, known to anyone skilled in the art, may be used. Improvements to speed of processing can also be implemented using the OpenVIDIA libraries.

We decided to use a stereo wearable camera rig to give the player the option of either hanging the camera rig from a tripod or other mount above a desk, or wearing it. When worn, the player has the benefit of an infinitely large playing area, by simply assigning different transfer functions to a limitless library of real physical objects.

For example, in some of our cyborg street performances we used a vast expanse of sidewalk space to create a giant tubular glockenspiel (Fig 3). The result is a glockenspiel having continuously variable pitch.

We ported our latest version of this software to run on a camera phone, so that, plugging the special stick into the microphone input of the phone, one can use the instrument while listening to headphones (Fig 4).

Figure 2: Making a bell-like sound from hitting a desk: A computer music system that is not an electronic instrument. Sound originates acoustically, and the role of the computer is merely for post-processing (much like a Wah Wah pedal on a guitar). The center frequency of the filter's passband varies with position, as detected by the overhead camera rig. Note the wearable stereo camera rig hanging from a fixed location. The cameras can be mounted to a tripod, or worn by the player.

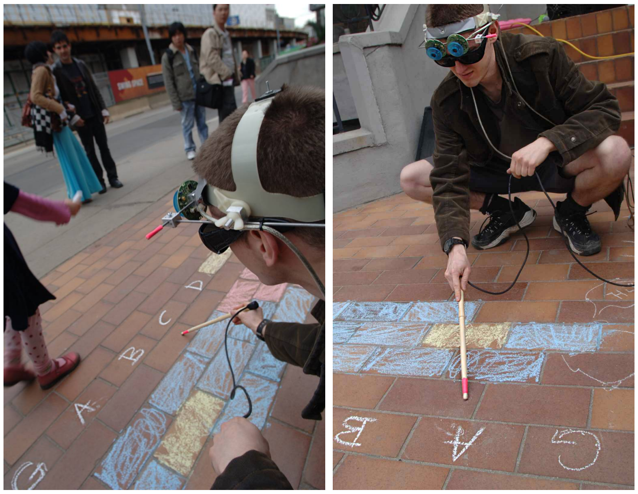


Figure 3: Sidewalk bricks or pool tiles cum tubular bells: Cyborg street performance using wearable camera rig and computer vision to control the transfer function of virtual effects pedals. A Wah-Wah like virtual effects pedal filters the acoustic sound of sticks hitting concrete. Filter transfer functions can be changed to achieve sounds of church bells, glockenspiels, piano, etc., but the sound all originates acoustically, thus remaining in the idiophones (not electrophones) top-level.

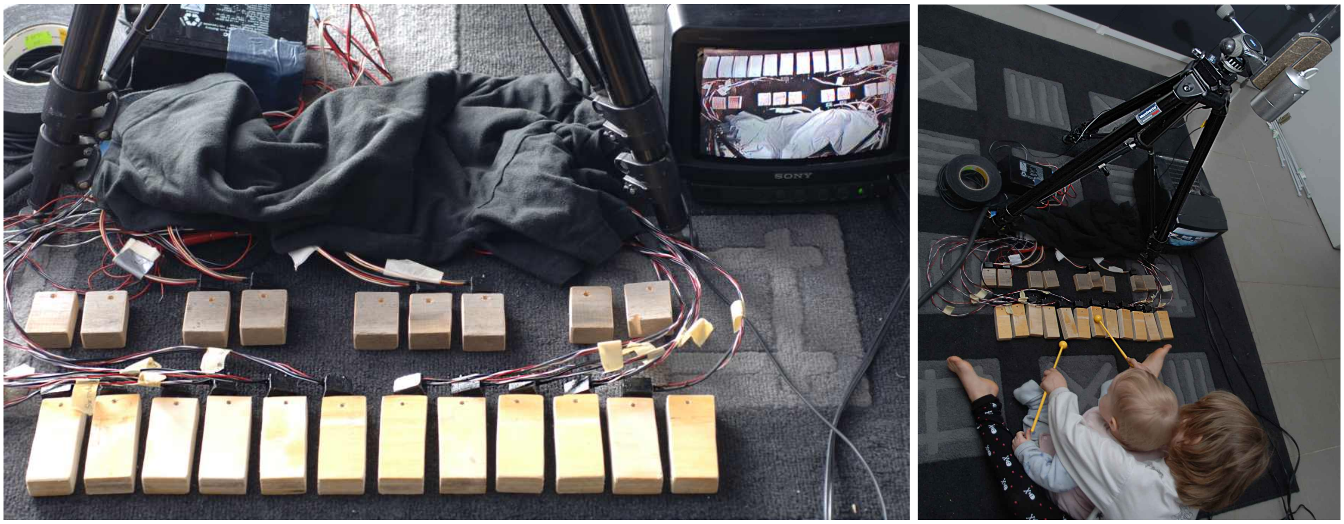


Figure 1: Making a bell-like sound from a dull thud: An array of wooden blocks is setup on a carpet. Each one is fitted with a separate acoustic transducer fed to a separate bandpass filter having transfer function equal to the quotient of the desired bell sound and the sound made by actually hitting the block.

Figure 4: A 12-bar idioscope running on a camera phone: One or two drumsticks or mallets with contact microphones plug into the headset input of a standard cameraphone. While listening to earphones, the player strikes an object in view of the camera. There are 12 vertical zones, each defining a separate note on the musical scale. The player can walk down the street and strike street signs, lamp posts, and the like, as part of a live performance webcast in real time. Here the player is locating a fire extinguisher through one of the 12 zones defined in the camera phone view and hitting the extinguisher with the mallet. Whatever pitch is produced by the sound of hitting the extinguisher is filtered and frequency-shifted to the desired note, so that all 12 notes can be produced by hitting this one fire extinguisher or other similar everyday objects.

2. CYBORG IN THE CYBERFOUNTAIN: TURNING ANYTHING INTO A MUSICAL INSTRUMENT!

Our approach enables computing to turn any everyday object that makes any kind of sound into an expressive non-electronic (non-electrophone) musical instrument.

We have seen that our new approach can be used to make any struck object into a perfectly tuned idiophone. Now we show that everyday continuous-sounding objects, such as water jets, can also be turned into musical instruments.

Water fountains make nice soothing sounds that flow continuously. Many such fountains have separate water jets. Consider, for example, Dundas Square (known as “Canada’s Times Square” or “Times Square North”) in Toronto. The main epicenter of Dundas Square is an array of 600 ground spray nozzles located under 20 stainless steel grilles. This is a prominent yet playful public space known as an “urban beach” (Fig. 5). An urban beach, or urbeach, is defined as a space that includes at least one intellectually, artistically, or culturally sophisticated water feature that is also an aquatic play area, free of boundaries, and is located within a culturally or artistically significant area of a city. By being free of boundaries, what is meant is that anyone can wander into the space to play in the water at any time of the day or night, without having to pass through a registration or processing area as with a place where people come specifically for aquatic play, such as a waterpark that operates only during specific hours of the day. Thus most “splash pads” or “spraygrounds” are not quite urban beaches.

An urbeach, being a nice open space, provides an intellectual playground for cyborg peformance. Here the flow of water gives us a unique sonic possibility because we don’t have to hit it to get it to make sound. Rather than hitting the physical object, we simply insert one or two microphones (in this case, we actually used hydrophones, i.e. special microphones optimized for underwater use) into the water jets of a fountain, and suddenly it becomes a new musical in-

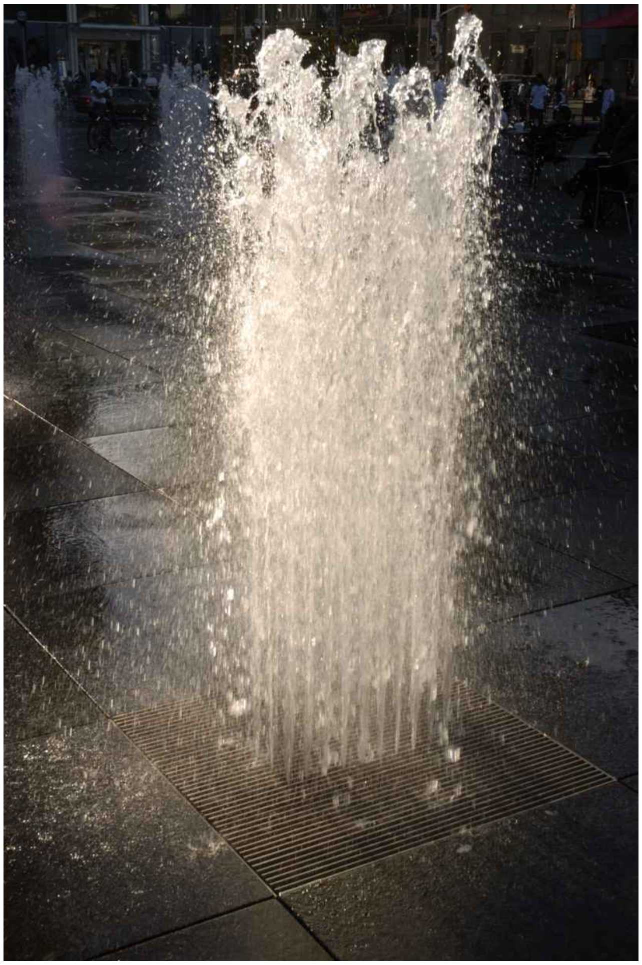
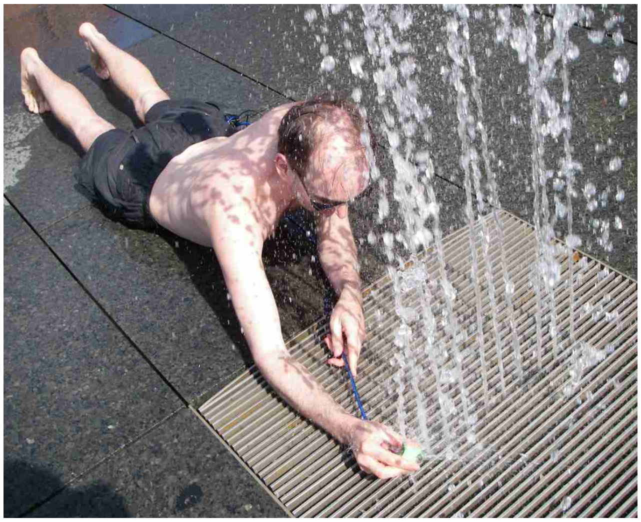


Figure 5: Urbeach (urban beach) at Dundas Square: Like many fountains, Dundas Square consists of various linear arrays of water jets. As an urban beach, these fountains run 24 hours a day, and are accessible for anyone to play in at any time of day. In "cyborgspace" we can see these water jets as "keys" on a musical keyboard.



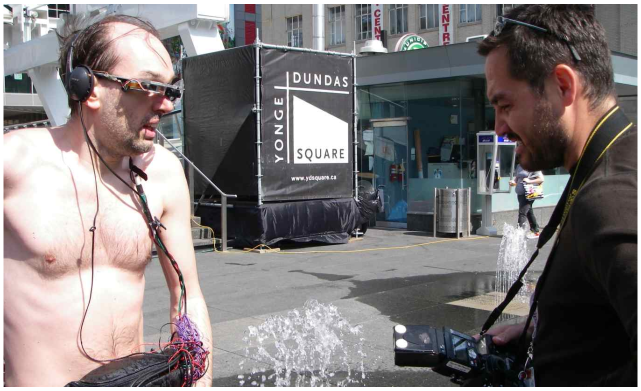


Figure 6: Cyborg in the fountain: Musical performance in a public fountain, with waterproof wearable computer. Interview with Globe and Mail newspaper reporter.

strument in which the sound source originates in water. See Fig. 6 and 7.

3. A FRICTION IDIOPHONE HAVING POLYPHONY COMBINED WITH CONTINUOUSLY VARIABLE PITCH

It was Benjamin Franklin's love of water that led him to invent the glass armonica (sometimes also referred to as glass harmonica), a glass harp consisting of a row of glass goblets all mounted to a single spinning metal shaft.

While playing glass harp underwater, we found that the water imparted some nice attributes to the sound, but we wanted some additional versatility, and the option to have a high Q-factor (less damping) at certain times during our performances. In order to achieve this, we used a spinning cylinder, which produced sound continuously along its entire length.

The sound is picked up by a contact microphone in the cylinder, and transmitted wirelessly to a computer. A computer vision system also connected to the camera takes note of where the rod is touched (positions, orientations, and contact geometry of all fingers in contact with the rod).

Figure 8: Polyphonic friction-idiophone having continuously variable pitch: A spinning aluminum cylinder with a specially textured surface produces sound picked up by a wireless contact microphone inside the cylinder. The sound is fed to one or more (depending on the number of fingers touching the cylinder) bandpass filters controlled by computer vision. The instrument can be used above or below the surface of the water.

This information is used to control the attributes of one or more (depending on the number of fingers touching) bandpass filters. The instrument was used in a variety of public performances (street performances, underwater performances, etc.). See Fig 8.

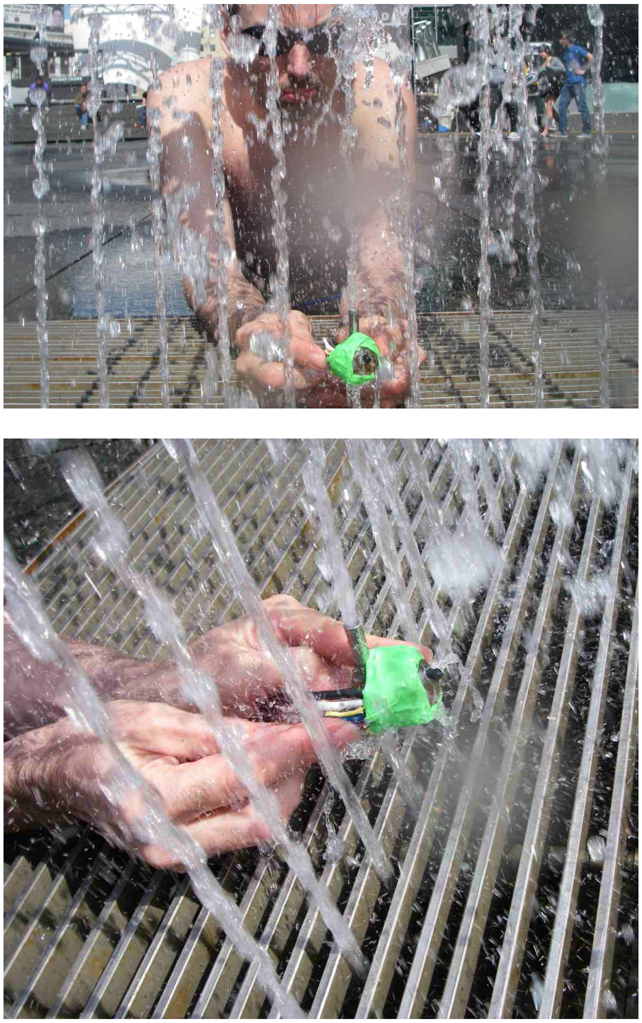


Figure 7: Hydrophone pickups held over water jets: In a computer-mediated world, each water jet is assigned to a different note on a musical scale. The computer vision system (wearable computer and EyeTap) is programmed so that the leftmost jet is the lowest note and the rightmost jet is the highest note. Sound from the hydrophone pickups is post-processed by filters in the computer that select out or emphasize harmonics in the sound that create the desired note, depending on input from the computer vision system. Because sound originates from the water, and is merely processed with computerized effects, this is not an electronic instrument (any more than is, for example, an electric guitar). The instrument affords a full range of musical expression and captures subtle changes in the way the hydrophones are held in the water jets.

4. ARE THESE NEW INSTRUMENTS ELECTRONIC INSTRUMENTS?

The new instruments such as the idioscope and other instruments derived from physical reality-based sounds (e.g. playing in a fountain as a musical instrument) all derive their initial sound production from real physical processes.

The idioscope uses computer vision to adjust computerized effects that post-process actual sounds from microphones, hydrophones, geophones, or the like. Such a "hyperacoustic" instrument makes it possible to bring subsonic and ultrasonic acoustic vibrations into the audible spectrum and add to the richly physical experience of playing a real acoustic instrument.

Playing in a fountain also gave rise to real physical sound, in which the role of the computer was simply one of filtering to favor desired harmonics, based on computer vision.

Unlike a hyperinstrument[2] in which position sensors, or the like, add synthetic sounds to an acoustic instrument, our hyperacoustic instruments use acoustic sound as their primary computer input, with vision affecting the processing of this sound.

To prove this point, we also constructed some variations of the instrument using mechanical resonators, as well as analog electric resonators (such as a computer-controlled Cry Baby (TM) Wah Wah pedal), to convince even a skeptic of the acousticality of the instrument (e.g. using computer vision to position the setting of an analog guitar pedal connected to a vacuum tube amplifier).

Accordingly, we feel that regardless of whether these post-processing effects are mechanical, analog, or digital, the idioscope, in whole, remains an idiophone, since the initial sound production comes from solid three dimensional physical matter in the real world, also giving a fundamentally tactile and "real" playing experience.

We believe, therefore, that the idioscope is not an electronic instrument, any more so than is an electric guitar or electric piano.

5. CYBORGINSTRUMENTS

We propose a new technology called Cyborglogging for use with idiophones and these new musical instruments, that allows users to share the expressions in real-time. While the instruments themselves are not electronic instruments, the filtering of the sound using a camera phone affords the use of Cyborglogging technology to provide greater utility to the performer.

This work initially grew out of the problems with bringing sheet music into fountains. Even with sheet protectors and lamination, loose pages tend to pose problems in an aquatic play environment, and many public pools even prohibit non-bound reading material, because of possible hazards with loose pages blowing around in an aquatic space.

The 'glog and 'glogger programs implement Cyborglogging (See http://glogger.mobi and http://eyetap.org) and have been used to facilitate communication with remote experts, as well as for social networking. The program runs

Figure 9: Waterproof electric eyeglasses can be used for two purposes: (1) to do computer vision and modify the acoustically generated sounds produced by the water; (2) to replace the mess of loose sheets of music that might otherwise be brought into an urban beach fountain/hydraulophone.

on most modern camera phones and it allows users to share their experience through instant image sharing and commenting. Users from other countries can also connect to the glogger and remotely interact with each other through this application.

5.1 Using EyeTap as Visual Annotation for a musical instrument

Specialized eyewear is already commonly used in aquatic spaces, with the need for improved seeing aids, and the like. Improvements in ruggedization and waterproofing make it possible to build weatherproof electric seeing aids.

A simple outdoor EyeTap, for example, can be used to help musicians communicate with each other, as well as read their music without the need for printed paper. Paperless sheet music thus becomes a possibility in a water fountain. An example of an *aquaborg* is pictured in Fig. 9.

Presently, a number of companies have manufactured waterproof camera phones and waterproof personal computers. Our waterproof EyeTap technology uses a miniature red laser diode to "paint" overlays in red laser-light (http://eyetap.org) for participants wearing our special "reality-mediator" sunglasses. Participants unable to use the special eyeglasses can instead use our 'glog and 'glogger system (http://glogger.mobi).

Audiovisual links are augmented with a live wirelessly sent music score drawn directly from the sensory capabilities of the instrument.

A prime location for such musical collaboration is at the hydraulophone (musical water fountain) in front of the Ontario Science Centre in Toronto, Canada.

Additionally, the existing wirelessly controlled lighting in the park (controlled over the Telus phone network) could function in concert with the music from the hydraulophone water fountain, functioning as a cyber-conductor which is visible to the hydraulophone player (similar to how a performer can see the motion of a conductor's baton out of the corner of their eye while focused on their instrument or sheet music). Both systems–EyeTap and 'glogger–are fully accessible, and are commonly used to assist the disabled (See http://wearcam.org/webramps.htm). The hydraulophone at the Ontario Science Centre is one of the few aquatic play facilities that is fully wheelchair accessible. Additionally, it has received very high ratings for use by the visually impaired. Accordingly, it can also tie into the seeing-eye-people network (another facet of 'glog and 'glogger).

See Fig.10.

5.2 Sharing of the musical experience with the Cyborglogging Technology

Cyborglogging allows a music score that has been played, recorded and transmitted wirelessly to be shared with a community or social networks. Other participants can hear the recordings from the world wide web and facilitate the music learning experience with no geographic limitations. Social networking can be applied to music lessons, as well as the sharing of expertise and remote experiences of science museum spaces.

Cyborglogging technology not only allows sharing of recorded data, but also the sharing of the processes used to create the music. The system allows user defined filter settings for customization of the sounds created from the physical processes to be shared amongst other users. Through the sharing of filters and process, each user in the community can experience the process of creating music, no longer being confined to listening to recordings in order to experience far-away musical instruments. This is an important aspect of the instruments presented here because of the physical nature of the sound generation from real-world objects. With the fountain jet pickups system, the sharing of filter settings allows others to experience the musical framework created by others, while also experiencing the tacticle feel of the water, and splash of the fountain. In this way, the experience of performing in a fountain is shared in the community across geographical boundaries, where each user may experience it in a different fountain, but hear and feel the same musical qualities of the performances.

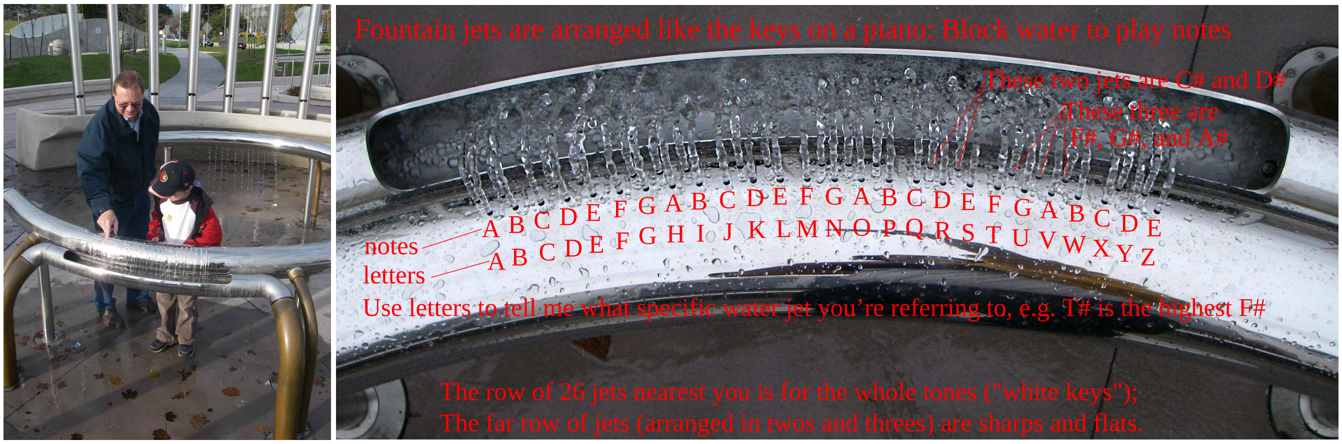


Figure 10: Remote experts and social networking through EyeTap CyberGlasses: An expert system can annotate a player's view of the instrument. The remote expert can coordinate with a multiple participants wearing special sunglasses that we provide, to make the music lessons fun and educational for a teacher or parent as well as a child. A computer-controlled red laser in the sunglasses "paints" annotation over the wearer's field of view in real time while he or she plays. Waterproof EyeTap technology is described in more detail in http://eyetap.org When the special glasses are not available to parents and their children, or after hours, any camera phone on the Telus network can be used as a reality-mediator to facilitate the music lesson.

6. THEORY OF THE "CYBORGSCOPE" (ACOUSTIC INSTRUMENTS IN CYBORGSPACE)

In sections 1 to 3, we presented cyborg instruments which use computer processing based on acoustically-generated audio from one sound source, or several sound sources.

Fig. 1 illustrates how we often use one pickup for each musical note, i.e. one for each sounding mechanism.

In some cyborg instruments (eg. Fig. 2), we only use one acoustic pickup, and use discrete fields in the computer vision to expand the expressivity played through the one pickup across several musical notes.

For the most expressivity, though, we usually use water jets as the medium of interaction, with one hydrophone (underwater microphone) pickup for each water jet.

In general, we take the sound of the water, $w_m(t)$ and process it computationally, with impulse response $h_m(t)$, with separate processing for each of the water signals $m \in [1, M]$.

In the cases when a single hydrophone is used to pickup one common water audio signal, then

$$w_m(t) = w(t) \quad \forall m \in [1, M] \tag{1}$$

The computer vision processing affords us further details on the manner in which a note is being played, and this additional data is used to provide further expressivity in the resulting notes. As this computer vision data originates electronically, not from an acoustic sound source, we refer to the signal as $e_m(t)$, an electronic, non-acoustic modifier on the acoustic sound. For our instruments we use $e_m(t)$ as a simple gain, thereby preserving the acousticality of each instrument.

The set of M processed results are given by

$$y_m(t) = (e_m(t)w_m(t)) * h_m(t) = \int_{\tau=-\infty}^{+\infty} e_m(t-\tau)w_m(t-\tau)h_m(\tau)d\tau \tag{2}$$

Often the non-acoustic modifier $e_m(t)$ is set by the vertical height of a water jet (or mallet, or human hand) seen in the computer's field of vision. One simple way of calculating the location of a water jet, mallet, or hand is doing an optical match, and computing the centre of mass of the match field $\rho_v(x_v, y_v, t)$:

$$x_v^{CM}(t) = \frac{1}{M_v(t)} \int_0^{x_v^{MAX}} \int_0^{y_v^{MAX}} x_v \rho_v(x_v, y_v, t) dx_v dy_v \tag{3}$$

$$y_v^{CM}(t) = \frac{1}{M_v(t)} \int_0^{x_v^{MAX}} \int_0^{y_v^{MAX}} y_v \rho_v(x_v, y_v, t) dx_v dy_v \tag{4}$$

$$M_v(t) = \int_0^{x_v^{MAX}} \int_0^{y_v^{MAX}} \rho_v(x_v, y_v, t) dx_v dy_v \tag{5}$$

where x_v and y_v are coordinates in computer vision. In the most basic implementation, the optical match density can be computed by thresholding the image in terms of how well each pixel matches a certain predefined colour.

In our implementation we set up the optical field into 12 regions for 12 musical notes. Finally, the non-acoustic modifier gain is given by the vertical position in the display: $e(t) = y_v^{CM}(t)/y_v^{MAX}$.

The sound processing, in discrete-time, sampled terms, is

$$y_m[n] = (e_m[n]w_m[n]) * h_m[n] = \sum_{\tau=-\infty}^{+\infty} e_m[n-\tau]w_m[n-\tau]h_m[\tau] \tag{6}$$

Note the difference between the above equation and

$$e_m[n]\,(w_m[n] * h_m[n]) \tag{7}$$

which merely modulates the result of the processing. The form in Eqn. 6 is more appropriate because, when sound is ringing out like a bell, we want to modulate the way the bell is hit, not to modulate the reverberance of the bell (by changing its size or shape). We want to create virtual objects in cyberspace which have a virtual playing surface that plays differently depending on where it is hit. In contrast, we are not trying to create virtual objects which are physically morphing their shape over time. Thereby we create a cyberspace environment which is stationary and easy to understand for the user, a philosophy which is grounded in Equations 2 and 6.

6.1 Group delay

We often wish to understand the playing characteristics of instruments, in particular our new instruments in cyberspace. In real life, consider a large gong. The gong, when played, will have a delay approaching a second or more, and in fact should be primed in the seconds leading up to the main strike, by gently doing a mallet-roll on its surface, quietly so the audience doesn't hear. The gong is one example of an instrument with a long playing delay, which affects the technique with which it is played. Playing technique is an important consideration for our cyberspace instruments, where we desire to make training time as quick as possible.

With linear filters used for computer processing, we analyze the group delay of the filters. First we consider the maximum possible group delay. As our implementation is done with digital filters, consider the following pure delay filter:

$$h_{pd}[n] = \delta[n - n_d], \tag{8}$$

with frequency response

$$H_{pd}(e^{j\omega}) = e^{-j\omega n_d}, \tag{9}$$

Let us set up this filter with the maximum possible delay, with only the last tap nonzero, ie. $n_d = N - 1$. In this case, the group delay d_g is simply

$$d_g = -\frac{d}{d\omega}\left[\angle H_{pd}(e^{j\omega})\right] = -\frac{d}{d\omega}\left[-\omega(N-1)\right] = N - 1 \tag{10}$$

which in physical time amounts to a delay of $(N-1)/f_s$ seconds. For our filters of order $N = 5$, with $N - 1 = 4$ time-delay taps, and a sampling rate of 44.1 kHz, a pure delay would be $d_g = 4/(44.1kHz) \doteq 91\mu s$. We observed additional delays due to sampling, conversion, etc. With a more tonally responsive filter (i.e. not the delay filter above), $d_g < (N-1)/f_s$.

More generally, with sampling rate f_s, for any filter m, we can write a group delay

$$d_{g,m}(\omega) \equiv -\frac{1}{f_s} \cdot \frac{d}{d\Omega}\left[\angle H_m(e^{j\Omega})\right]\Big|_{\omega} \tag{11}$$

The pure delay filter was just one example of a linear phase filter, which gives a simple time delay, for the entire frequency spectrum. More general filters do not give the same delay across the entire spectrum. To arrive at an estimate of the time delay perceived by the user, we average the delay function over all frequencies, but weight the average by how much of the signal is passed at each frequency. For one particular filtered-note m out of M notes in the instrument, the delay is

$$d_{gw,m} = \frac{\int_{\omega=-\pi}^{\pi} d_{g,m}(\omega)\left|H_m(e^{j\omega})\right| d\omega}{\int_{\omega=-\pi}^{\pi} \left|H_m(e^{j\omega})\right| d\omega} \tag{12}$$

The delay is now weighted by the amplitude response of the filter. To go futher we might also weight it across the actual input spectrum. Consider firstly the trivial case: a percussive input such as a signal from a mallet strike. A perfectly percussive input, taking the form of a Dirac delta measure input, contains precise timing information, but no pitch information. That is, all input frequencies are heard with the same weighting. In such a case, by the sifting property of convolution, the above delay still applies.

For water flow, on the other hand, the sound of turbulent flow takes on a nontrivial spectrum, in general, and furthermore is better suited to pitch-intensive music rather than timing-intensive music. We then weight the time delay by the mean input spectrum as well:

$$d_{gww,m} = \frac{\int_{\omega=-\pi}^{\pi} \left\{-\frac{d}{d\Omega}\left[\angle H_m(e^{j\Omega})\right]\right\} \left|H_m(e^{j\omega})\right| \left|\overline{W}_m(e^{j\omega})\right| d\omega}{f_s \cdot \int_{\omega=-\pi}^{\pi} \left|H_m(e^{j\omega})\right| d\omega \cdot \int_{\omega=-\pi}^{\pi} \left|\overline{W}_m(e^{j\omega})\right| d\omega} \tag{13}$$

The total delay is

$$d_{inst,m} = d_{a,m} + d_{mic,m} + d_{ADC} + d_{gww,m} + d_{net} + d_{DAC} + d_{spkr} \tag{14}$$

for each note m of M notes in the instrument. $d_{a,m}$ is the acoustic playing delay (eg. gong resonance time), $d_{mic,m}$ is the delay of picking up the sound from microphones, hydrophones, etc., d_{ADC} and d_{DAC} are analog-digital conversion times, d_{net} is the network transmission time for cyberspace instruments, d_{spkr} is any sound output delay.

For a beginner learning the instrument, it is often sufficient to approximate, for the instrument as a whole,

$$d_{inst} \simeq d_{ADC} + d_{net} + d_{DAC} + d_{spkr} + \frac{1}{M}\sum_{m=0}^{M}\left(\begin{array}{r} d_{a,m} + d_{mic,m} \\ + d_{gww,m} \end{array}\right) \tag{15}$$

An instrument-wide playing delay is relevant to performers, and is relevent as well in the creation of the instrument, knowing its overall quickness of response.

Note that the acoustically recorded spectrum, $\overline{W}_m(e^{j\omega})$ is effectively a measure of the turbulence characteristics of fluid flow through water jet m. Different jets produce different spectra (with varying outlet diameter, flow rate, etc.), in particular if each jet is tuned for a specific note. One of our other papers in these proceedings has more detail on this aspect.

6.2 Fluid Sampling

An alternative way of processing acoustically-generated sound to create a cyborg instrument is through fluid sampling. We (Mann and Janzen) propose fluid sampling in one of our other papers to appear in these proceedings.

Envelope detection is the first stage of the algorithm:

$$LPF\{|w(t)|\} \tag{16}$$

In the digital frequency domain we are effectively doing:

$$|W(e^{j\omega})| \tag{17}$$

To arrive at the audio blurring kernal, we then take the derivative:

$$b(t) = \frac{d}{dt}\left(LPF\{|w(t)|\}\right) \tag{18}$$

which leads to

$$B(e^{j\omega}) = j\omega|W(e^{j\omega})| \tag{19}$$

The true delay, not indicated above, comes from the practical filter which is used for envelope detection.

7. FURTHER RESEARCH: GLOBAL VILLAGE FOUNTAIN

Much in the spirit of a town's well, or the "village pump", we envision the fountain as a well spring of community and connectivity.

The fountain symbolizes the fluidity of human efforts across many different areas of study, ranging from fluid mechanics to landscape architecture, to water therapy (health care), music therapy, and the flow of information/ communication.

We've described a fluid, limitless user-interface design strategy that creates fun and playful forms of interaction—such as through frolicking in a fountain—yet also having embodiments serious enough to be the main civic and cultural centerpiece of a landmark architecture site that creates a healing and therapeutic space, open to all members of the public 24 hours a day.

These instruments create a positive, social, fun, educational, and spiritually uplifting space around one of the most basic and comforting elements, such as water.

We also presented the idea that a fountain can be a musical instrument, thus bringing an element of sophistication to aquatic play that makes it accessible to people of all ages (including the elderly), and not just children, as with other known aquatic play features.

In addition to the hydraulophone [3] (http://funtain.ca) we also propose the concept of a Global Village Fountain, consisting of a network of hydraulophones in various cities, connected to each other, so that, for example, a participant can play music and "jam" with (or even have a waterfight with) someone in another city. Blocking water jets in one city can cause the water to spray out in another city, where, by way of a video-conferencing link, we can engage others in playful fun and frolic across geographical boundaries.

8. CONCLUSION

The unique ability of the computer to facilitate the creation of new non-electrophonic musical instruments creates a new kind of very physical real-world experience that can work within "cyborgspace" with a camera phone or wearable computer, in a shared computer-mediated reality space.

Computer processing, digital filtering, and the like may be applied to acoustic instruments, without changing the fundamental Hornbostel Sachs categorization of the resulting hybrid. For example, our idioscopes and other instruments like the fountain, apply digital signal processing and computer vision to real world sounds, and the fact remains that the sound originated acoustically. Tracing back to the original source of sound is true to the spirit of the Hornbostel-Sachs organology, as typically practiced, and cuts to the core of instruments which produce sound from physical processes.

9. ACKNOWLEDGEMENTS

The authors would like to thank Thought Technology Limited and Dr. Hal Myers of Thought Technology Limited. The support of SSHRC, by way of the Anonequity Project at the University of Ottawa, as well as the support of Canada Council for the Arts, Ontario Arts Council, and Toronto Arts Council, is also greatly acknowledged. The recent hydraulophone installations in public spaces were done in collaboration with Chris Aimone. Additionally, Anurag Sehgal assisted in what became the winning entry of the Coram International Sustainable Design Award.

10. REFERENCES

[1] M. J. Kartomi. *On Concepts and Classifications of Musical Instruments.* Chicago Studies in Ethnomusicology (CSE). University of Chicago Press, 1990.

[2] T. Machover. Hyperinstruments: A composer's approach to the evolution of intelligent musical instruments. In W. Freeman, editor, *Cyberarts.* Spartan Books, San Francisco, 1991.

[3] S. Mann. "fluid streams": fountains that are keyboards with nozzle spray as keys that give rich tactile feedback and are more expressive and more fun than plastic keys. In *Proceedings of the 13th annual ACM international conference on Multimedia*, pages 181 – 190, Hilton, Singapore, 2005.

[4] C. Sachs. *The History of Musical Instruments.* Norton, New York, 1940.