MEDIATE:
Key Sonic Developments in an Interactive Installation for Children with Autism

Hans Timmermans*, Gerard van Wolferen†, Paul Newland†, Simon Kunath†

*Utrecht School of Music and Technology, Utrecht School of the Arts
hans.timmermans@kmt.hku.nl, gerard.vanwolferen@kmt.hku.nl

†University of Portsmouth
simon.kunath@port.ac.uk, paul.newland@port.ac.uk

http://web.port.ac.uk/mediate/

Abstract
Our paper initially gives a brief overview of MEDIATE, a multi-sensory, interactive environment created with the aim to encourage expressive engagement from children with autism. We then detail the particular innovations required to support such a context when incorporating sonic and tactile interfaces which through a ‘signature analyzer’ gain the ability to both detect and respond to user behaviours within this environment.

1 Introduction
MEDIATE is a 5th Framework EU project funded by the IST (Information Society Technologies) strand undertaken from October 2001 until April 2004. Developed by three European design teams from Universities in Portsmouth (UK), Barcelona (Spain) and Utrecht (Netherlands), constructed by Show Connections and evaluated by Goldsmiths and Kings Colleges, London, the acronym MEDIATE stands for (Multi-sensory Environment Design for an Interface between Autistic and Typical Expressiveness). The project has culminated in the production of a multi-sensory, responsive environment designed to stimulate interaction and expression in children with autism through visual, sonic and vibrotactile means.

1.1 Touch
MEDIATE wished to allow touch and vibration to grow as creative expression - countering the historic trend for poor exploration of this modality. Although a fundamental means of communication, touch and vibratory feedback are usually under played: through work at the School of Art, Design and Media, University of Portsmouth, this modality is encouraged by means of a responsive multi-textured ‘Tune Fork’ and vibrating ‘Impression Wall’.

1.2 Aural
The modality of sound production and creative feedback is discovered by the user through MEDIATE’s capacity to transpose movement via a sensory floor into audio input and adapt a user’s vocalisations. Utrecht School of Music and Technology, Utrecht School of the Arts developing approaches which allowed the user to interactively engage with the generated output - filling the environment with multi-channel aural compositions under the user’s control.

1.3 Visual
Interactive Audiovisual Group, Universitat Pompeu Fabra enhanced an infrared detection system (EyesWeb™) and worked with particle based light compositions projected on two independent environment walls. Drawing the individual into participation these visually exciting surface patterns are animated and influenced by the user’s position, wall contact and full body gesture. Through this control, poetic and flexible interaction scenarios can evolve as the user moves within MEDIATE.

1.4 Signature
Integral to the functioning of MEDIATE is an intelligent inbuilt pattern detection system, ‘signature’, that recognises repetitive behaviour in subjects and reacts by continually adjusting and enhancing its response. Together this multi-sensory interface is designed to facilitate expression through generative, full body interaction and allow children with autism to experience a sense of agency over their surroundings without need for the usual socio-linguistic skills.
2 MEDIATE Architecture

The various sensory research initiatives of the design teams culminated in the environment detailed below:

Diagram 1: Plan of the MEDIATE environment

a). An internal irregular walled hexagon creating the responsive space is held in place by an external Tri-lite rectilinear support frame which is covered (sides and roof) by fire resistant blackout curtains. The irregular hexagon space sits on a tailored floor comprising 27 full or trimmed 1m x 1m sandwich construction double wooden panels covered with a fitted carpet. Twenty of these panels incorporate a piezo-ceramic pick-up, itself sandwiched between two layers of industrial scouring fabric, the signal from a footstep is fed to the tactile and sonic interfaces, both as analogue 'crunch' and midi trigger. The roof consists of a grid of infrared (800-900 nanometres wavelength) lighting. The overall physical footprint of the combined structure is 7.5m x 7.5m with a 3.3 m height.

b). Internal walls 3 and 5 are full coverage back projection screens (3m x 2.5m), whilst walls 2 and 4 carry a touch sensitive, multi-textured branched bas-relief (tune fork) and 3 vibro-equipped floor to ceiling curvi-linear impression tubes respectively. Wall 6 is left plain with a disguised viewing grill (through which parents / carers and psychologists can observe) and wall 1 holds the entrance which is reached by an exterior ramp, semi-shielded by a fabric covered arched passageway.

c). The area behind wall 2 is occupied by technicians and the machinery racks

d). The rear of walls 3 and 5 house the two back projectors, each equipped with a camera capture to detect any silhouette movement at the projection screen faces, i.e., effectively allowing walls 3 and 5 to be ‘touch’ responsive.

(i). Independent for each screen, visual particle generation (typically 5,000 to 20,000) where each particle is trackable, modifiable and testable against the user’s captured properties.

(ii). EyesWeb (Camurri et al. 2000), user position sensing analysis fed by 7 further cameras (9 in all) capturing the user’s actions at 25 frames a second - the top lit infra-red illuminated user being separated from the non infra-red generating environment projections.

(e). 5 channel microphone (4 walls and overhead) and pressure sensing capture / tactile pick-up (floor, tune-fork and impression wall) with eight speaker sonic and multi-vibration feedback units controlled by MAX/MSP interaction software.

(f). A high-speed network server and client architecture, with bespoke messaging protocols, facilitates rapid communication amongst all modules in all machines.

(g) Fault redundant electrical control and machine management subsystem.

As is appreciated from this overview the environment has combined various recent and novel technologies – we would now like to consider in more detail those facilitating the aural interface and its evolution.

3 Sonic Interface in Depth

The sonic interface is essentially an interactive sound/music system and of course there are many precedents in the field of computer music, (Paine 2003, Rowe 1992, Winkler 1998). Indeed we have used this type of system in our interactive dance projects and interactive installations before. However, in the case of MEDIATE this general type has been adapted to the intended user-profile and advanced by generating sophisticated options to analyze audio. The analysis options are used to feed audio-derived parameters to the rest of the MEDIATE-system; in this way audio can be the multi-channel input for the visual and tactile modalities of the MEDIATE-system (and reflexively for the sonic interface itself).

3.1 Tasks and Components

The sonic interface addresses its tasks through the following components:

- **Input modules**, 5 microphones are used to pick up the sounds produced by the user. The gain of each microphone is controlled to have maximum resolution. (AGC with limiters). The microphone with the loudest signal level is emphasized by the Auto Gain Control.

- **Analysis modules**, audio is analyzed on amplitude, pitch, harmonic content and formants; we used fiddle~ to do this. Only relevant analysis data is communicated to the MEDIATE system; the sonic interface deciding relevance on the basis of particular result parameters - stability, coherence, etc.
Transformation modules, we developed a range of transformations to map the analysis results of all modalities to relevant aspects of sound/music (Winkler, T. 1998, Rowe, R. 1992). Some of the interaction models are used for the other modalities too.

Generation modules, we developed and used some techniques like tunable resonators fed by pulses or (pink) noise to generate sound with acoustical properties and intuitive control, adapted forms of formant synthesis to simulate singing voices (reaching from bass to soprano), modules using pre-recorded samples to generate interactive soundscapes from these can be projected in 3D.

Sound projection and Localization modules, we use dynamic volume panning here to distribute and localize sound (stereo or mono sources) using 8 channels (or more if necessary) to simulate 3 Dimensional Sound.

Output modules to control 8 loudspeakers (or more if necessary), to record the audio-output to hard disk and to record the analysis results.

3.2 Bespoke Developments
Within the framework of MEDIATE we developed:
Communication modules from the basis of TCP/IP connections using modifications of the TCP/IP objects of Open Dragon.
Command modules (constructors and parsers) to communicate on an equitable abstraction level with the MEDIATE system.
Memory modules to store and retrieve generated material and analysis results.
Recording modules to record the data of the memory modules to hard disk and retrieve the material from hard disk.

3.3 Considerations and Decisions
To construct the sonic interface we considered (among others) to develop in C++ starting from scratch and / or to use existing and proven technologies with features that allow extensions and modifications. This resulted in the choice of MAX/MSP with additional development in C++.

3.4 Features and Limitations
The sonic interface is capable of:
• Analyzing 5-channel audio input via the microphones.
• Auto Gain Control is available as is limiting and compression.
• Analyze amplitude, pitch and timbre (spectral analysis related to the fundamental frequency), formants.
• Communicate analysis data and control data to the MEDIATE-system by TCP/IP.
• Generate audio-events and continuous audio-streams controlled by external parameters.
• Transform external parameters into musical relevant parameters.

However, the sonic interface is not designed to:
• Recognize words or sounds.
• Generate compositions of high artistic value.
• Be a sophisticated musical instrument rather it is a limited performance instrument, with some built in intelligence.

3.5 Hardware
The sonic interface runs on two Apple computers running MAX/MSP 4.3 on OSX. Development of externals was done in Metrowerks CodeWarrior 8.3. TCP/IP objects were modified using the source-code of Open Dragon.

3.6 Limitations
As stated in 3.4 the sonic interface is limited in its design to be appropriately responsive to the first time experience by a child with autism. We implemented several sets of instruments each having their own sound design and performance characteristics to provide a palette of possible adaptations of the system to accommodate to specific user interaction. Appropriate choice being informed by our signature analyzer which we now describe in more detail.

4 Signature Analyzer
Signature analyzer is a pattern detector that searches for identical sequences in derived descriptions of signals that are produced by a user’s performance. The software development built on the work of van Wolferen (1998), Lenoir (1998) and Maas (2002) and has the following properties:
Detection - all incoming signals (e.g.: sequences of \([x,y,z,\ldots]\) co-ordinates) are processed into descriptive layers and converted to symbolic notation. Every layer can be treated as an input stream. Input signals sometimes are symbolic in origin. In such cases input signals can be processed raw. (Written language for example, can be processed as a sequence of text symbols in order to detect idiomatic expressions but the same input can also be converted into a word type representation sequence in order to detect word type order, author characteristics.)
Storage - detected patterns are stored in databases that are each associated with one stream of incoming data. Streams may be combined (vectorized) to build new streams. For each stream there is one pattern detector running. Numerous streams can be analyzed simultaneously. Analysis processes can be distributed over multiple servers if necessary.
Database structures - each database is optimized for fast access to allow continuous re-writing (sorting). The most frequently occurring pattern, the longest pattern or the pattern with the highest impact (a numerical relation of both) is available to the MEDIATE System at any time during a user session.
From structure to content - this way of analyzing provides us with as many ordered lists of patterns, as there are streams to analyze. With this information every pattern can be traced back to the original performance of the user allowing for monitoring and content analysis tasks.

The meaning of input streams - patterns in the signal derived from sensors in the floor point at the ways the user crosses the space. Patterns from the EyesWeb™ camera analysis point at user specific body movements. Patterns in the audio spectrum analysis, point at similarity of sounds.

In conclusion - signature analyzer is built to behave like an ideal intuition. Ideal in the sense that all pattern occurrences will be shown. Intuitive in the sense that it acts like the sometimes unconscious human awareness of similarity. Also the analysis is objective in the sense that it is not focused to find specific patterns of performance and results from various sessions can be compared to find changes or development in the user’s behavioural repertory.

5 Tactile Interface

The tactile components of MEDIATE comprise three discrete areas - floor, tune fork and impression wall. We detail these here because of their close correspondence with the sonic interface.

5.1 Floor

Each of the twenty active floor panels, the construction of which is detailed in Section 2 (a), when trodden on gives an audio signal that sounds like a footstep on gravel, shingle or pebbles. It was felt that this sound would potentially engage the user most immediately due to its inherent properties of embodiment and dynamism. After familiarization and with increasing interaction, complexity evolves informed by the signature and decision maker modules. The analogue sounds are then morphed into more abstract midi triggered samples through water splashes and shingles to singing voices.

5.2 Tune Fork

Eight different tactile areas are mounted vertically on wall 3 in the rough shape of a tuning fork, consisting of a range of materials - with embedded piezo bi-morph sensors fed to MIDI triggers. MIDI is then passed to MAX/MSP and a sampler. The samples being used at any time are a part of the current interaction again as directed by signature and decision maker’s interpretation of the user’s state.

The bottom half of the tune fork integrates shammy leather, short hair fur, cork, bark and polished wood – the top half studded rubber, felt, crinkled plastic and metal. The range of materials was chosen and ordered along the themes of rough/smooth and organic/synthetic. It was thought that due to the nature of the space, the user would probably start touching at the right side (smooth end) of the tune fork.

Sounds generated from the different areas vary from ‘natural’ sounds related to the materials used to percussive.

5.3 Impression Wall

Fourteen fixed coil moving armature inertial transducers, capable of independent vibration are mounted on a floating structure underneath several layers of medical foam. Sixteen piezo-ceramic sensor inputs are fed to a bespoke interface, the ‘MIDI Podule’, which also forms part of the overall MEDIATE control subsystem.

The foam has the particular property of slow restoration to its normal shape. Interactions can be traced for a little longer than with ordinary upholstering material. This is a simple but effective way to make the surface interactive and interesting. The shape of the elements holding the vibrating actuators was designed to envelop, to be leant against, to be stroked and to invite interaction while also being integrated with the screens (multimodal crossover).

6 Conclusions

In the future we would like to develop the intelligent response of the signature analyzer so that it is capable of more proficiently allowing a self-organisation capacity to arise (Newland and Creed 2003). In the case of the sonic interface this would increase its tolerance for extreme interactions whilst allowing it to manifest more original compositional output. There is also the hope of gaining greater crossover between modalities to improve synaesthetic response – the vibration mode as indicated in 5.3 an obvious transition between tactile and sound.

7 Acknowledgments

Thanks to Norman Jaffe and Thekla Schiphorst for the permission to use the source code of the TCP/IP Objects of Open Dragon.

References


