A Dynamic Sonification Device in Creative Music Therapy

Dr Alan Lem[1], Dr Garth Paine[2], Jon Drummond[3]
1, 2, 3 - School of Communication Arts, UWS
2 – MARCS Auditory Labs, UWS
1 – Nordoff Robins

INTRODUCTION

Musical expression is one of the most ubiquitous means of creative expression; it evolves in a multi-faceted manner to reflect cultural and historic changes. It is important that the tools for expression evolve to meet the needs of both able and disabled individuals.

Creative Music Therapy (CMT) is a form of improvisational music therapy developed in the fifties and sixties by Paul Nordoff and Clive Robbins who pioneered the use of music improvisation as a means of communication with severely disabled children. The fundamental premise of CMT is that musical response is an intrinsic human trait and that every person, regardless of his/her physical or emotional condition, has the potential to engage in musical interaction with another person. Thus in CMT, the client and the therapist improvise music together building a creative musical relationship in which a therapeutic process takes place and change can be supported. The unique qualities of music as therapy enhance communication and can help people to live more resourcefully and creatively (Nordoff and Robbins, 1977).

The CMT approach is readily accessible to most client groups, but its success largely depends on the client’s ability to use acoustic musical instruments. It follows then that movement is an important component of CMT. Indeed, from a gentle finger tap to powerful beating on the drum, movement defines the dynamics of the music therapy session becoming a powerful reflective measure of its emotional and communicative contents. From this perspective it is easy to understand why CMT has only limited applications with people who have severe difficulties with muscular strength and motor control.

In the conventional approach to CMT, the less the client is able to do due to his/her physical disability, the heavier is the therapist's responsibility for musical guidance. Within an interactive sonic environment, such as Soundbeam (2006); an ultra-sonic sensor capable of detecting movement and translating it into sound, the essence of therapy lies in the internal motivation of the client. In such an environment, even the most severely disabled person has the opportunity for expression, creativity and control. Russell (1996, 1997) listed several areas of development promoted through the use of Soundbeam: Aesthetic Awareness, Imagination, Listening Skills, Choice-Making Skills, Conceptual Skills, Motor Planning Skills, Reflective Cognition, Memory Skills, Spatial Orientation, Language Skills, Exploring a Hypothesis, Social Skills, and Confidence.

Unfortunately, all currently available sonification systems, including Soundbeam, use triggering approaches providing momentary musical responses when, for instance a beam of light is broken, and not on a continuous basis. More importantly, these devices are not sensitive to the velocity, acceleration or direction of movement, and so do not reflect the intentional dynamic qualities of movement, thus limiting the scope of musical and emotional expression. These limitations become apparent in the light of current findings in the area of developmental psychology.

Warren and Coaten (1993) viewed the link between movement and emotion to be of paramount importance to the cognitive and physical development. According to the authors, movement if directed by sensations and feelings can lead to an improved body image and a more effective emotional release.

MacDonald (cited in Payne, 1992) looked at how an underdeveloped vestibular system, which is responsible for spatial orientation, may lead to specific learning and interpersonal difficulties. According to MacDonald problems with recognition and relation to one's own body affect the ability to receive, remember and 'motor plan' or process stimuli. Case studies presented by the author elucidate the role of movement and gesture as a means of emotional expression and communication (p207). For example, she claims that the language of dance can stimulate sensory integration as movements become more expressive and enjoyable instead of being a mere exercise.

Two recent studies (Pederson, 2004; Trevarthan and Malloch's, 2000) explored the connections between movement, emotional expression and the development of musical-therapeutic relationship. Pederson (2004) discussed the notion of psychodynamic movement in music therapy as giving music therapists a deeper contact with their own bodies and feelings, thus enabling them to better understand and assist their clients. In their study of intrinsic human musicality, Trevarthan and Malloch's (2000) found that the "…interplay between […] sounds and bodily movement becomes a medium of the music therapeutic relationship" (p11). Although no particular client base was described, their description of the connection between sound, gesture and the expression of feelings is directly pertinent to improvisational music therapy work with people who have disabilities.

The present study explored the application of an interactive, dynamic sonification system within the framework of Creative Music Therapy. The key objective was to develop a new system that would be uniquely suitable for use with adults who have severe physical disabilities. It was expected that the dynamic nuance of the system would provide intelligible feedback to the client in real-time enabling a continuous monitoring of the quality of the gesture. Of particular interest to the investigators was the extent to which the dynamic properties of the technology could assist in a better understanding of the relationship between movement and emotional expression in people who have severe physical disabilities. The specific aims were:

1. To develop a dynamic and interactive sonification, and visualisation system that would enable physically disabled music therapy clients to dynamically engage with virtual musical instruments.

2. To assess the suitability of the system to track gestures of music therapy clients and extract their characteristic features, leading to the identification of individual vocabulary of movement responses.

3. To increase an understanding of the relationship between the specific characteristics of sound, the emotional intent of the movement and the physical characteristics of different forms of physical disability.

It was also hoped that the outcomes of the study would lead to a better understanding of the role of movement in improvisational music therapy and possibly the psychodynamic patterns of a music therapy session as it unfolds in time.

In light of the above considerations, video tracking became the technology of choice for gathering movement and gestural behavior, because it is a non- invasive, is easy to set up, and provides a sampling rate, which at 25 frames per second, is sufficient to reflect the change in dynamic of most human gestures. A principal consideration was to develop a system that produced sonification outputs that would be relatively unique to each participant. Such an objective excludes the use of any prerecorded sound, which by its very nature has a predetermined and fixed morphology. The development of a system utilizing real-time sound synthesis would provide a framework for producing dynamic morphologies (Paine, 2004), which although constrained within the aesthetic parameters of the synthesis algorithm, could produce sonifications specifically related to the momentary action of each individual client.

METHOD

Participants

Participants were 5 adults (2 Males and 3 Females) in their mid twenties, all of whom had some form of physical disability. They included quadriplegia, paraplegia, hemiplegia, of spastic and athetoid forms, and a degenerative motor disease of unspecified origins. Four participants were confined to wheelchairs, whereas the fifth participant was able to move independently. In addition to the physical disability, two participants had a moderate form and one severe form of intellectual disability.

Before being accepted for the study each person, or their carer, was asked to sign a consent form specifying that they understood the objectives of the study and agreed to participate. Each participant had also answered a number of questions concerning their physical ability, musical responses and ways of expressing emotions.

Procedure

Each participant took part in 8 half-hour individual session conducted in the morning at the Golden Stave Music Therapy in the University of Western Sydney. During the session the participants were staying in their wheelchair, except for the fifth participant who was sitting in an ordinary chair or moved around the room. During the trial, a small video camera was pointed to the participant from approximately 2 meters away. In addition to the participant, there were four people were present in the room: the music therapist, the software developer and two music therapy students who were recording the session and collecting the data.

During the session, the music therapist interacted with the participants verbally by describing to them the nature of the technology and the ways to generate the sound, with the emphasis placed on the creative use of the technology. Once it was perceived that the participant started using the system appropriately, the music therapist would often start interacting with him musically using the piano or anther musical instrument deemed appropriate. These interactions were strictly of improvisatory nature.

During each session, several presets were being trialled and adjusted according to each individual's movements and preferences. The adjustments would often involve a conversation between the music therapist and the software developer, which was normally followed by a prolonged period of silence during which the software developer adjusted the system's parameters.

Materials

The Sonification System

Conceptually the sonification software system consisted of two main components — a motion analysis engine and a sound synthesis engine (Figure 1). A video camera was used to capture the participant's movements. The motion analysis engine analysed the realtime video stream, outputting control data streams that were mapped to synthesis control parameters in the sound synthesis engine.

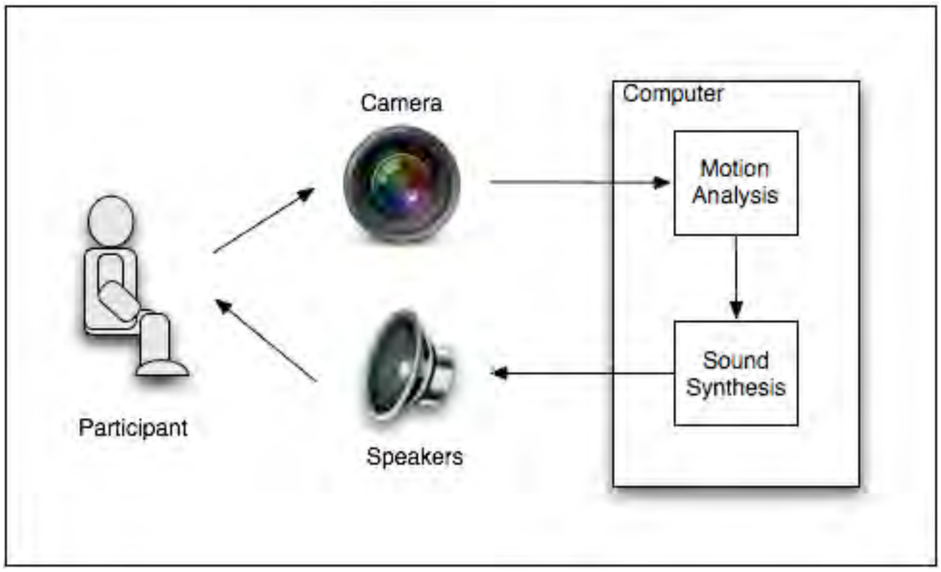

Figure 1 System Overview

Informing the sonification design of the system was the concept of dynamic orchestration (Paine 2007) (Paine, 2002) so that changes in the dynamic of the gesture registered by the system would be interpreted sonically as changes in the voicing of the underlying synthesis model. Thus not only was the amount of movement and location of the centre of the movement mapped to synthesis parameters but also the underling voicing structure of the instrument could be changed dynamically in relationship to changes in the measure of motion detected.

Motion Analysis Engine

To enable the sonification of participant's gestures we used computer video tracking (Dinkla 1994; Packer and Jordan 2002; Paine 2001) techniques to provide temporal measures of the amount of movement and calculated the centre of that movement within the camera's field of view. For this particular application, computer based video tracking provided a number of advantages over other motion sensing techniques, specifically — an easily configurable analysis window altered by simply moving the camera or zooming in or out on a particular location; effective sensing range, data resolution and sampling rate; and no need to attach sensor hardware directly to a participant's body.

A Pointgrey, Dragonfly[1] FireWire (IEEE 1394 interface) video camera was used to capture the video and the softVNS (Rokeby 2002) library of objects for Max/MSP[2] was used to programme the motion analysis software engine. Motion detection was performed using frame differencing (Gonzalez and Woods 2007) with the video window being analysed in regions, as an eight by four grid (Figure 2). Consequently motion detection could be applied individually to each of the resulting sixteen cells in addition to the whole video window. The advantage of this approach is that it allows separation of objects within the space, facilitating individual limb tracking, whilst simultaneously analysing the total overall motion.   This approach also allows for masking of segments of the video image where analysis is not desired, and allows grid positions to be grouped for higher level analysis.

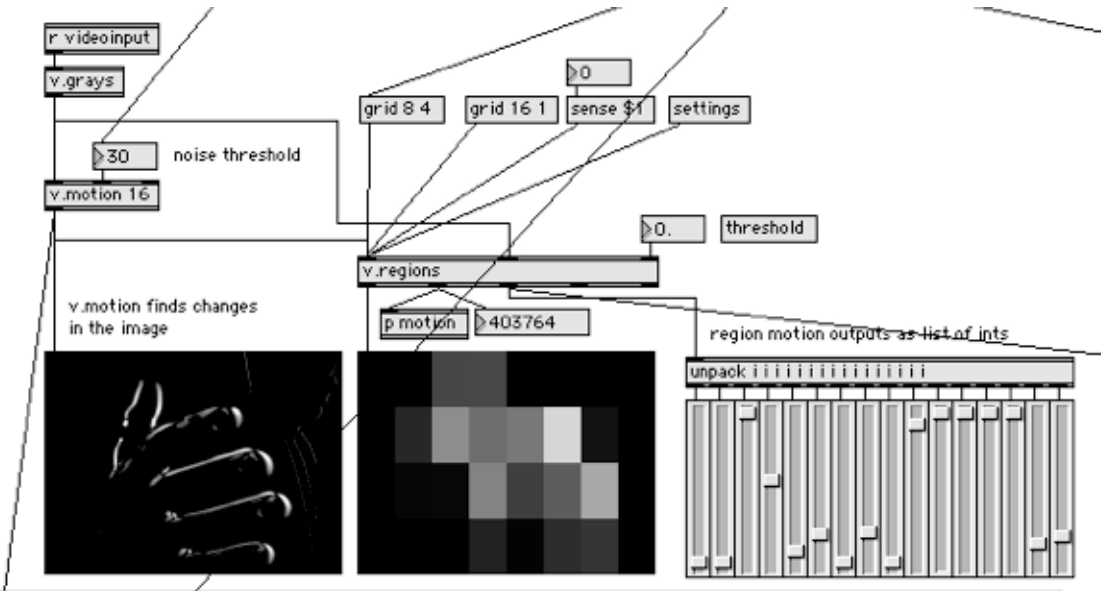

Figure 2 Video window being analysed as a eight by four grid

One of the challenges of developing a system of this nature is to accommodate a wide range of gestural inputs. In order for the synthesis engine to produce a full range of timbres, it requires a consistently full scale range of input data (0 to 1.0). If the camera is maintained at a consistent distance from the subjects, with a fixed zoom range, then the scale of data produced by video analysis will be dramatically different between the movement of an arm, and the movement of a foot. Furthermore, subjects may

[1] See http://www.ptgrey.com/products/dragonfly2/index.asp viewed 7 October 2008
[2] See http://www.cycling74.com, viewed 7 October 2008

move their entire body in the effort to produce a gesture with a single limb, producing movement data throughout the frame. A video tracking system will not easily differentiate between the intended and unintended movement. It was therefore necessary to develop a dynamic mechanism for differentiating between gesture patterns of differing temporal structure, whilst also filtering out unintentional movement, by utilizing a threshold, below which analysis is not carried out, and then to scale the resultant data into the full range required by the sound synthesis algorithm. In addition, a change in the light intensity in the room in which the experiments were taking place, would produce an associated variation in the range of data produced by the video sensing system. It was therefore necessary to develop a mechanism that automatically scaled the output of the video tracking system in such a manner as to produce a full range of input to the sound synthesis algorithm. It was decided that the system would be triggered to auto-scale at the beginning of each client session. This approach allowed the development of a peak detection algorithm, which accommodated both the environmental light levels, and the maximum gestural activity of the subject. This approach calibrated the system to provide the full timbrel range of the synthesis algorithm for each subject, regardless of the range of movement they displayed, or changes in the environmental conditions.

A robust and sophisticated peak detection algorithm was implemented as part of the motion analysis engine, to ensure the appropriate data scaling of large and small gestures while accommodating varying lighting conditions (Figure 3). The overall sensitivity of the system could be controlled by a slider at the bottom of the main screen (Figure 5). The LED to the felt of this slider was red when the system had not auto-scaled, and turned green when the system was calibrated. It was important once the system was calibrated, that the subject was not moved within the video frame, that the video camera was not zoomed in or out, and that the therapist did not enter the video frame. If any of these things occurred, the system required recalibration, carried out by triggering the auto-scale function, by pressing the Recalibrate button on the left of screen.

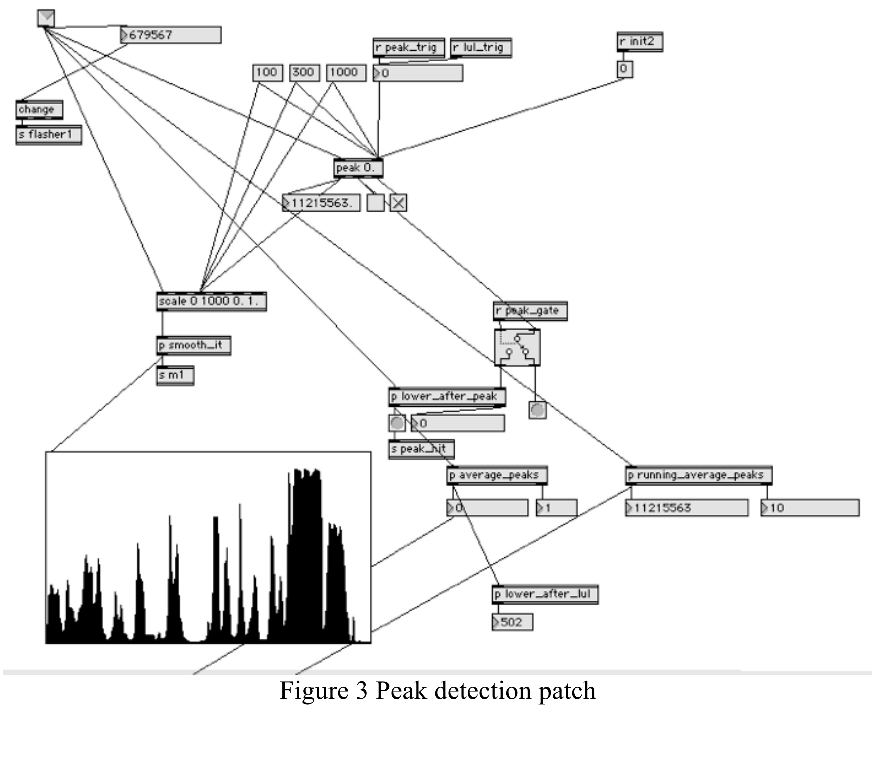

Figure 3 Peak detection patch

Sound Synthesis Engine

The sound synthesis engine was based on a three-voice (layers) model in which the overall mix of the three voices was determined by the amount of movement detected by the system (Figure 4). Each layer could be assigned a unique synthesis algorithm. In general terms, small gestures were be interpreted as layer one while increasing amounts of movement would bring in layer two, with maximum movement bringing in layer three. Each voice also had an independently controllable reverb and spatial mix. A completed instrument in the system consisted of individual synthesis algorithm assignments to the three layers and envelope shapes for dynamically mixing in and out the voices, reverbs and spatial location. An instrument could be further customised for a particular participant by assigning a specific sensitivity level, and selecting active/non-active regions in the eight by four grid for each of the three layers. In the context of the project this overall configuration of an instrument for a particular participant became known as a *preset*.

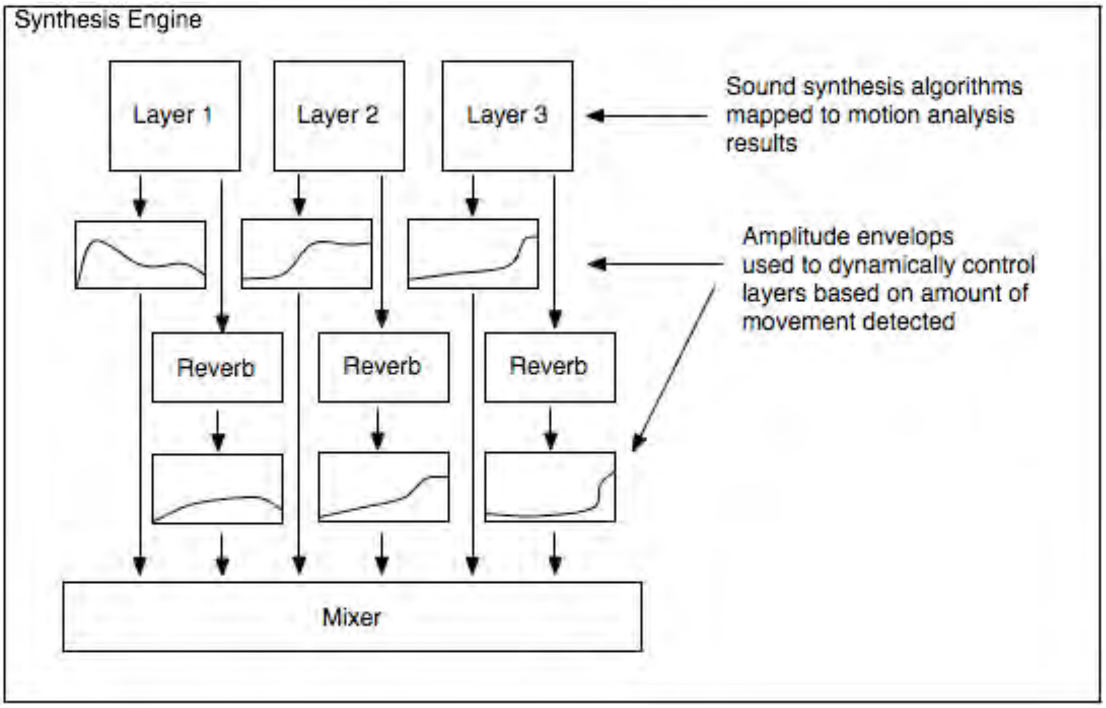

Figure 4 Dynamic orchestration of the three sound layers

Synthesis Algorithms

The system was designed to support a variety of different synthesis approaches. A custom designed granular synthesier based on FOF synthesis (Clarke and Rodet 2003) was incorporated as a voice in many of the presets while a custom designed basic sampler enabled the inclusion of audio files as a layer. The system was also designed to support the use of VST Plugins (Steinberg's Virtual Studio Technology) as voices. Using the VST[3] Plugin architecture we were able to incorporate the Pluggo[4] software

[3] See http://en.wikipedia.org/wiki/Virtual_Studio_Technology viewed 7 October 2008
[4] See http://www.cycling74.com/products/pluggo viewed 7 October 2008

synthesisers and a software emulation of the Korg M1[5] synthesiser as voices in a preset, providing a vast array of sampled and synthesized sounds to build on.

A granular FOF synthesis engine was specifically designed for the system. FOF synthesis (Fonction d'Onde Formatique translated as Formant Wave-Form or Formant Wave Function) uses fragments of exponentially decaying sinusoids, shaped by an amplitude envelope consisting of a cosine shape for the attack and decay with a flat sustain (Eckel et al. 1995). This highly configurable synthesis model provided many control inputs that could be mapped to the outputs of the video analysis engine and in the initial stages of development provided a sonically rich set of possibilities in which to explore mapping gesture to sound. In consultation with the music therapist, the richness and subtlety of the synthesis engines was reduced in order to produce sound variations that were more striking in their difference. We were advised that the clients selected for the study would better perceive the relationships between gesture and sonification through dynamic timbral variations. To some extent this accentuated the use of dynamic orchestration, and reduced the level at which we implemented the dynamic morphology model.

User Interface

The system's operations were configured via a graphic interface operated via a touch screen (Figure 5). Using this control interface the system's main parameters could be configured, specifically the selection of different presets, master volume control, system sensitivity and recalibration. The interface also provided graphic feedback of the system's internal states including a graph of the amount of movement measured, main mix amplitude and alternate direct camera and analysis views.

[5] See http://www.vintagesynth.com/korg/m1.shtml viewed 7 October 2008

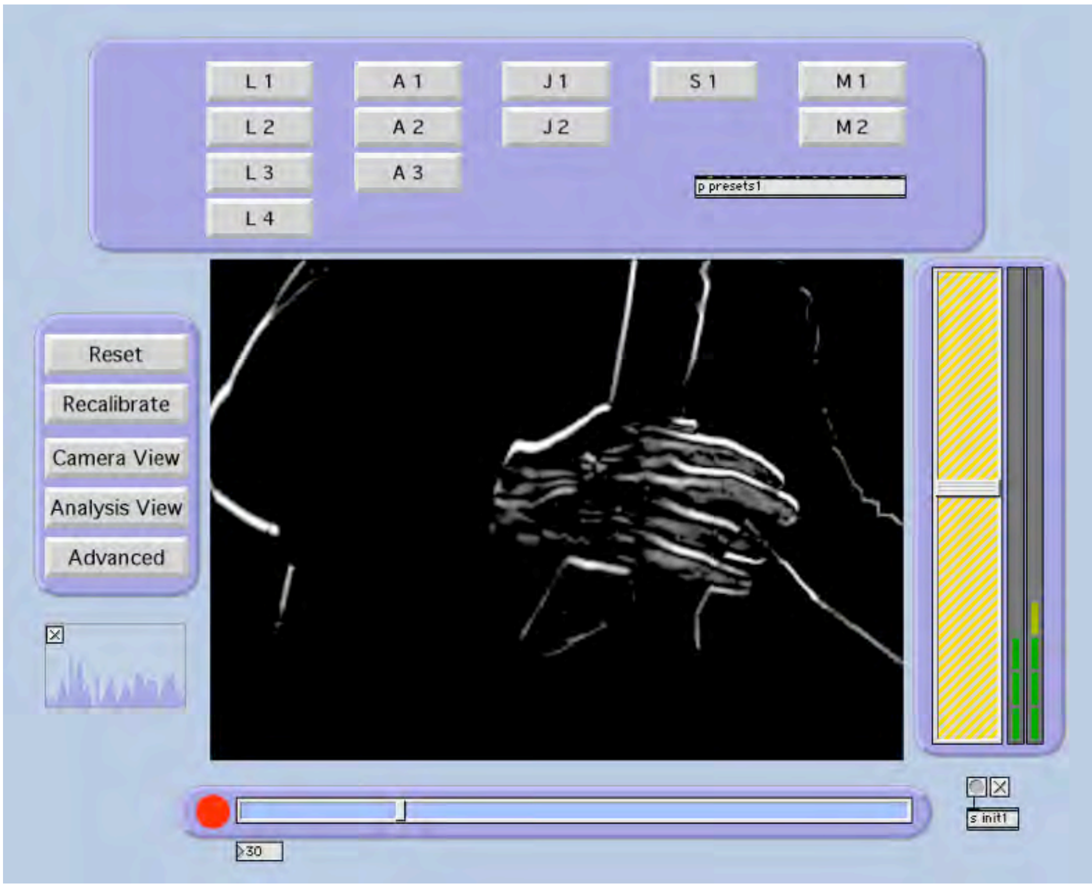

Figure 5 Main Touch Screen Interface

Three different and distinct sonification presets were configured for the first clinical trials session. These three presets were then modified and adapted for the specific participants over the eight sessions, addressing their movement styles and responses to the initial presets. Most of the significant present changes took place in the first few weeks of the experiment with subsequent sessions providing an opportunity for fine-tuning and system refinement. Through this process the initial three presets evolved into a final set of twelve presets, representing a customised group of presets for each individual.

Music Therapy Room

The sessions were conducted in a large music room equipment with a grand piano, a wide selection of drums and both tuned and untuned percussion.

Data Analysis

Each session was viewed and transcribed by two assessors who noted down verbal instructions given to participants, the characteristic features of each movement and the quality of the corresponding sound. This assessment was conducted twice, during the session and retrospectively, by viewing the video recording. Data from each session were collated in a table showing concurrently (1) verbal instructions, (2) the characteristic features of movement and (3) the description of the sound. All patterns of movement observed to repeat consistently were categorised according to their magnitude and intensity. The characteristic features of sound were categorised according to their perceived spaciousness, the closest corresponding sound produced by a conventional musical instrument, and the immediacy of responsiveness to movement.

The data were then assessed for ongoing changes in the vocabulary participants' movement responses and their potential relationships to the quality of the sound and the immediacy of responsiveness of the system.

RESULTS

The analysis of data of the four participants who were confined to wheelchairs yielded no significant results. All four were observed to engage with notable enthusiasm during the first two sessions, but their vocabulary of responses including the variability and intensity of movements not only remained unchanged, but showed a tendency to decrease over the period of the trial. It was also observed that wheelchair bound participants became notably exhausted after the first five minutes of a session.

With participant 5, two important findings should be noted. One is that the most prevalent type of movement was the large, dynamic hand movement (N=40), while the least frequently used was the fine finger movement. The second finding shows that, over the duration of the trial, the incidence of head movements and movements indicative of greater mobility (leg and whole body movements) had both decreased while the incidence of fine finger movements has increased (Table 1).

Table 1. Participant 5: The general categories of movement responses.

	Session 1	Session 2	Session 3	Session 4	Session 5	Session 6	Session 7	Session 8	Total
Large hand movements	3	10	3	4	2	6	5	7	40
Punching hand movements	1	6	2	2	1	2	3	1	18
Head movement	2	1	1	2	5	0	0	0	11
Finger movement	0	0	2	0	2	2	1	4	11
Leg and whole body movement	7	8	3	3	1	0	0	0	22
Total	13	25	11	11	11	10	9	12	102

Each of the above categories of movement was divided into more detailed subcategories:

1. Large hand movements

   a) Both hands
   b) Single left hand
   c) Intensive
   d) Gentle

e) Wide
f) Narrow
g) Fast

2. Punching hand movements

a) Throwing/punching
b) Bowling/throwing
c) Boxing
d) With closed fist
e) Over the head
f) Fast
g) Sudden downwards movement

3. Head movement

a) Single movement
b) Shaking abruptly
c) Nodding
d) Up and down

4. Finger movement

1) single movement

5. Leg and body movement indicative of greater mobility

a) Feet
b) Walking
c) Standing up
d) Stomping with feet
e) Swinging

Table 2 shows a description of 11 computer presets used by participant 5. They have been categorised according to their perceived spaciousness, the closest corresponding sound produced by a conventional musical instrument, and the immediacy of responsiveness to movement. Detailed descriptions of all 11 presets can be viewed in Appendix 1.

Table 2. Description of Individual Presets Used by Participant 5

Preset No	Spaciousness	Sound	Immediacy of response
1	Spacious	Washover; water drops	Slow
2	Spacious	High Bells	Medium
3	Spacious	Piano	Faster

4	**Spacious**	**Marimba; Bell**	**Proportional**
5	**Hard Press**	**Metallic**	**Fast**
6	**Spacious**	**Combination**	**Slow/Fast**
7	**Hard Press**	**Metallic**	**Fast**
8	**Spacious**	**Marimba; Bell**	**Proportional**
9	**Hard Press**	**Metallic**	**Fast**
10	**Normal**	**Strings**	**Proportional**
11	**Normal**	**Saxophone**	**Proportional**

Following the identification of individual characteristics of movement and the sonic output of the computer system, the data was cross-analysed to understand if and how individual characteristics of sounds used might have assisted in motivating the participant to move in particular way and how this might have changed over the period of the trial.

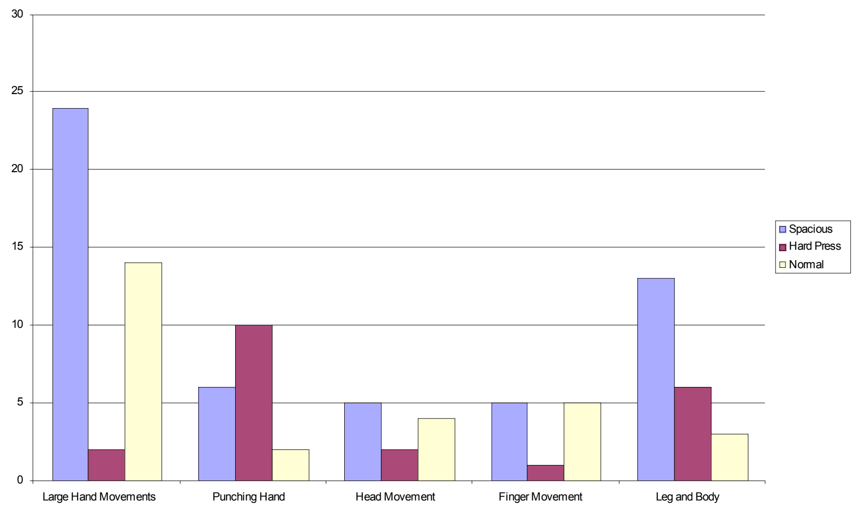

Figure 6. A relationship between the incidence of different movements and the spaciousness of sound generated by the system.

Figure 6 shows ongoing relationships between the incidence of different movements and the spaciousness of sound generated by the system. As can be seen, the highest incidence of large hand and leg movements correlated with the presence of the most spacious sound, whereas punching hand movement seemed to correlate with hard pressed sound of little ambience.

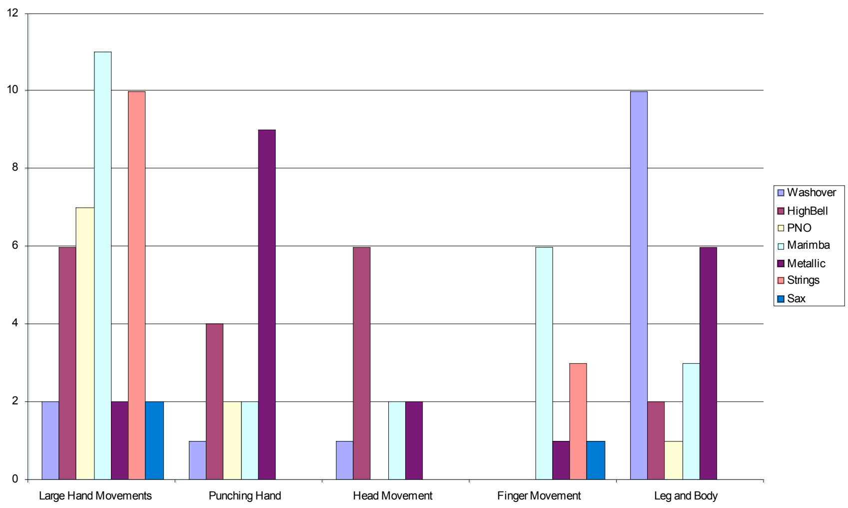

Figure 7. A relationship between the incidence of different movements and the type of instrument simulated by the system.

Figure 7 shows relationships between the incidence of different movements and the type of instrument simulated by the system. As can be seen, large hand movements corresponded with the presence of marimba and strings, punchy hand movements with metallic sounds, head movements with high bells, finger movements with marimba and leg and body movements with a distinctive wash-over sound.

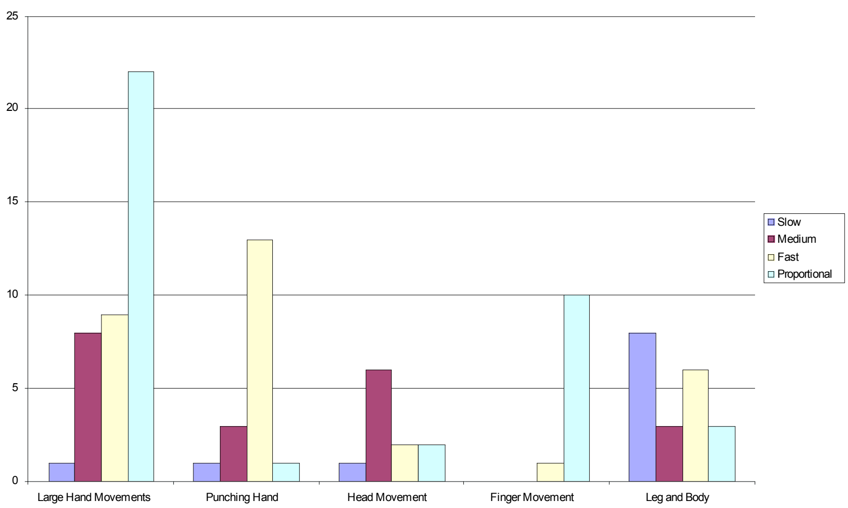

Figure 8. A relationship between the incidence of different movements and the dynamic responsiveness of the system.

The final comparison between the incidence of different movements and the sonic characteristics of the system's output (Figure 8) shows that large hand movements occurred most often when the system responded proportionally to the movement, namely when there was a direct of correlation between the dynamic intent of the movement and the dynamic growth of the unfolding sound, not only in the attack envelope but its perceived sharpness.

The analysis of data presented in Figures 6, 7 and 8 revealed the following relationships between the incidence of different movements demonstrated by Participant 5 and the dynamic responsiveness of the system:

1) The occurrence of large hand movements correlated with the presence of spacious sound imitating marimba and strings, and the responsiveness of the system which followed the dynamic intent of the movement;

2) The occurrence of punching hand movements correlated with the presence of hard pressed sound imitating a metallic object and fast responsiveness of the system;

3) The occurrence of head movements correlated with the presence of spacious sound pressed sound imitating high pitched bells and medium responsiveness of the system;

4) The occurrence of finger movements correlated with the presence of sound of high and normal level of spaciousness, the sound imitating marimba metallic object and the responsiveness of the system which followed the dynamic intent of the movement;

5) The occurrence of leg and movements indicative of greater mobility correlated with the presence of spacious sound described as 'washover', and slow responsiveness of the system.

A significant refinement to the system as a result of the sessions was the creation of a new, duration based interaction model. For participants who performed with more constrained moment, distinguishing large and small gestures became problematic for the computer system. To address the needs of these participants a timer model was developed so that the longer a body movement was sustained, the more dynamic the sound became. This time based analysis function was incorporated into the instrument's preset definition. Subsequently, this type of interaction model was also explored and developed into presets for the more able bodied participants, As it produced a longer-term reward for dynamic activity, in addition to the momentary response attributed to the momentary analysis and sonification algorithms.

DISCUSSION

The present study explored the application of an interactive, dynamic sonification system within the framework of Creative Music Therapy. The objective was to develop a new system that would be uniquely suitable for use with adults who have severe physical disabilities. It was expected that the dynamic nuance of the system would provide intelligible feedback to the clients in real-time enabling a continuous monitoring of the quality of the gesture. Of particular interest was the extent to which the dynamic properties of the technology could assist in a better understanding of the relationship between movement and emotional expression in people who have severe physical disabilities.

The project was unique in that it aimed to develop a new technology with the assistance of disabled participants who would traditionally be asked to use it only after it has been refined and ready to apply in its target role. This was one of the important constraints of the study which resulted directly from its limited funding.

Initial results showed that only one out of the five participants was able to fulfil his role in a satisfactory manner. This participant suffered from hemiplegia but was fully mobile. Responses of the other four participants, who were confined to wheelchairs, indicated clearly that, while they were able to engage with enthusiasm during the first two sessions, their responses, including the variability and intensity of movements not only remained unchanged, but showed a tendency to decrease over the period of the study. It was also observed that these participants became notably exhausted after the first five minutes of a session. It is possible that part of the exhaustion related to the fact that the system was being developed during the sessions, and a large portion of each session was devoted to conversation and computer programming. It was therefore not possible to engage the participant in a spontaneous and fluent music improvisation. Future projects involving sonification technologies may want to consider this particular finding and involve wheelchair bound participants only when the technology is ready to use.

However, data obtained from Participant 5 has pointed to a number of tendencies and correlations. Perhaps the most interesting finding relates to the fact that over the period of the study the incidence of certain gross motor responses has decreased in favour of smaller, more refined movements. This finding is directly supported by the analysis of the sonic characteristics of the computer system, which shows that finer movements occurred at the highest level of incidence when the dynamic responsiveness of the system was proportional to the dynamic intent of the movement. This shows that not only the vocabulary of movements observed but also the sonic possibilities of the technology have undergone a certain refinement during the course of the study.

As far as the potential of the technology to represent the emotional characteristics of movement, there were indications that the spaciousness of the generated sound did influence the spaciousness of the movements and that the two were in direct correlation. It was also found that harsh and rapid sounds tended to stimulate movements which may be typical of certain elements of aggression. Obviously, data obtained from a single participant over the course of only eight sessions has to be treated with caution. Nonetheless, indications are that a dynamic sonification system may be a useful tool in the development of movement responses with certain groups of people who have physical disabilities.

Further development of the video tracking system may include high-level analysis units, reflecting aspects of posture, such as expansion and contraction, as a further approach to constraining sound synthesis outcomes in such a manner as to reflect not only the gesture of a particular limb, but the way in which the generation of that gesture causes variations in the entire body. Such total body movement had to be filtered out of the analysis considers in this project.

The development of a suitably rich sound synthesis engine for a project of this nature is in itself a difficult task. There are many sound synthesis variables which need to be controlled in an interrelated manner. Furthermore the system was designed as essentially a free improvisation system, as task oriented activities were not deemed appropriate, and would not have been applicable across all subjects. It would be possible to produce a gesture follower[6], useing a markov chains approach, produce training sets for a task oriented approach, such as drawing a circle in the air or indicating beats by moving the hand up and down (Bevilacqua, Müller, & Schnell, 2005).

References

Bevilacqua, F., Müller, R., & Schnell, N. (2005). MnM: a Max/MSP mapping toolbox. Paper presented at the International Conference on New Interfaces for Musical Expression, Vancouver, BC, Canada.

Clarke, M. and Rodet, X. (2003). Real-Time Fof and Fog Synthesis in Msp and Its Integration with Psola, In Proceedings of the Music Conference (ICMC2003). Singapore.

Dinkla, S. (1994). The History of the Interface in Interactive Art, In Proceedings of the 1994 International Symposium on Electronic Art (ISEA). Helsinki, Finland.

Eckel, G., Rocha Iturbide, M., and Becker, B. (1995). The Development of Gist, a Granular Synthesis Toolkit Based on an Extension of the Fof Generator, In Proceedings of the 1995 International Computer Music Conference (ICMC1995). San Francisco: International Computer Music Association.

Gonzalez, R. and Woods, R. (2007). Digital Image Processing, 3/E. Prentice Hall.

Nordoff, P. Robbins, C. (1977). Creative Music Therapy: individualized treatment for the handicapped child. NY: John Day.

Paine, G. (2001). Interactive Sound Works in Public Exhibition Spaces: An Artists Perspective, In Proceedings of the 2001 Australasian Computer Music Conference. University of Western Sydney: Australasian Computer Music Association, 67–73.

Paine, G. (2002). Interactivity, where to from here? Organised Sound, 7(3), 295-304.

Paine, G. (2004). Gesture and Musical Interaction: Interactive engagement through dynamic morphology. Paper presented at the New Interfaces for Musical Expression, 2004, Hamamatsu.

Paine, G. (2007). Sonic Immersion: Interactive Engagement in Real-Time Immersive Environments, Scan: Journal of Media Arts Culture, 4 (1).

Packer, R. and Jordan, K. (2002). Multimedia: From Wagner to Virtual Reality. New York: W. W. Norton & Company.

Payne, H. (ed) (1992) Dance Movement Therapy: Theory and Practice, London: Routledge.

Pederson, Inge Nygaard (2004) ‘Psychodynamic Movement-A Basic Training Methodology for Music Therapists’ at (http://www.musik.auc.dk/bib/psychodynamic%20movement.htm - 9/10/2005)

[6] See http://ftm.ircam.fr/index.php/Gesture_Follower viewed 7 October 2008

Rokeby, D (2002) softVNS http://homepage.mac.com/davidrokeby/softVNS.html viewed 1/10/2008.
Russell, K. & Russell, N. (1997). "See that? That's magic": New Sounds and Sights in Music Movement Improvisation - the SoundBeam Experience. ASME, Brisbane, July 1997.
Russell, K. (1996). Imagining the Music, Exploring the Movement: SoundBeam in the Sunshine State. Queensland Journal of Music Education, (4)1, 41-48.
Soundbeam Project (2006). (http://www.soundbeam.co.uk/downloads/leaflets/special-needs-leaflet.pdf, accessed 12 April 2006)
Trevarthen, C. and S.N. Malloch (2000) 'The Dance of Wellbeing: Defining the Musical Therapeutic Effect' in Nordic Journal of Music Therapy, 9, 2: 3-17.
Warren, B. and Coaten, C. (1993) 'Dance: Developing Self-Image and Self-Expression through Movement' in Warren, B. (ed) (1993) Using the Creative Arts in Therapy: A Practical Introduction, London:Routledge.

Appendix 1

Description of Presets Used by Participant 5

Preset 1

This preset uses a diversity of synthesis techniques, including frequency modulation, additive synthesis and modulation synthesis, none of which is dependent on an external synthesiser. There are 3 voices: 1st - a 'wash-over' sound (pad) inspired by Soundtrack (FM frequency modulation), 2nd similar in nature to the 1st one but brighter and with more texture to it, 3st similar in nature to Roland's Ice Rain (an arpeggiated simulation of water drops and electronic repetitive sounds). Smaller movements bring in layer 1, middle activity activates 2nd layer, while large movements bring in layer 3. In this preset a very low threshold was used to allow for a very low level of activity to trigger a sonic response.

Preset 2

3 voices. 1st - low frequency (Jews harp), 2nd – 'magic wand' , 3rd reinforces 2nd layer by adding distinctive high pitched bell-like sounds. More layers are brought in the longer the duration of movement, creating an overall increase in the intensity of sound. When the movement stops, the layers fade out individually in reverse order. This preset has a very low threshold.

Preset 3

The basis of this preset is a piano sound configured as follows: 2 voices. 1st - piano, 2nd - darker richer piano timbre. Smaller movements bring on the first piano layer, middle and larger movements introduce the darker richer piano timbre. The amount of movement also influences the key velocity of the voices and the amount of reverb applied to each layer. The video screen is divided into an 8 X 4 grid, with the vertical lines representing individual notes of a diatonic scale while the horizontal lines control the octaves.

Preset 4

3 voices, 1st - marimba, 2nd aquarium (bell – darker, ‘melancholic’), 3rd - bell ring. The voices are coming in according to the energy level, i.e. the greater the expression, the bigger the sound. Smaller movements bring on the marimba, whilst middle movements generate the mellow bells, and larger movements bring on the brighter bell ring. Reverb is being used in this preset; it increases with the increased intensity of sound. All the sounds are drawn from a Korg M1 simulator synthesiser and are triggered via MIDI interface. The video screen is divided into an 8 X 4 grid, with the vertical lines representing individual notes of a diatonic scale while the horizontal lines control the octaves.

Preset 5

2 voices. 1st - "Hard Press!" pitched metallic attack sound, 2nd - classic orchestral hit sound. Smaller movements bring on the pitched metallic sounds, middle and larger movements introduce the orchestral hits. The amount of movement also influences the key velocity of the voice and the amount of reverb applied to each layer.

Preset 6

3 voices. 1st - marimba, 2nd - muted bell-like sound, 3rd - drum kit. Smaller movements bring on the marimba, middle movements generate the bells with reverb. Larger movements bring in the drum kit.

Preset 7

3 voices, 1st - cymbals, 2 – a hitting, metallic, striking sound, 3rd a combination of cymbals and tom-tom drums. The voices are coming in according to the energy level, i.e. the greater the expression, the bigger the sound. Smaller movements bring on the cymbal sound by itself, mid-range activity activates the metallic hits, while large movements bring on more of the cymbals, drums and much louder metal hits. This preset has no reverb. All the sounds are drawn from a Korg M1 simulator synthesiser and are triggered via MIDI interface. The video screen is divided into an 8 X 4 grid, with the vertical lines representing individual notes of a diatonic scale while the horizontal lines control the octaves.

Preset 8

3 voices, 1st - marimba, 2nd aquarium (bell – darker, ‘melancholic’), 3rd - bell ring. The voices are coming in according to the energy level, i.e. the greater the expression, the bigger the sound. Smaller movements bring on the marimba, whilst middle movements generate the mellow bells, and larger movements bring on the brighter bell ring. Reverb is being used in this preset; it increases with the increased intensity of sound. All the sounds are drawn from a Korg M1 simulator synthesiser and are triggered via MIDI interface. The video screen is divided into an 8 X 4 grid, with the

vertical lines representing individual notes of a diatonic scale while the horizontal lines control the octaves.

Preset 9

3 voices, 1st - cymbals, 2 – a hitting, metallic, striking sound, 3rd a combination of cymbals and tom-tom drums. The voices are coming in according to the energy level, i.e. the greater the expression, the bigger the sound. Smaller movements bring on the cymbal sound by itself, mid-range activity activates the metallic hits, while large movements bring on more of the cymbals, drums and much louder metal hits. This preset has no reverb. All the sounds are drawn from a Korg M1 simulator synthesiser and are triggered via MIDI interface. The video screen is divided into an 8 X 4 grid, with the vertical lines representing individual notes of a diatonic scale while the horizontal lines control the octaves.

Preset 10

3 voices, 1st – string pizzicato, 2nd – a soft strings pad, 3rd - high strings. The voices are coming in according to the energy level, i.e. the greater the expression, the bigger the sound. Smaller movements bring on the pizzicatos, middle activity activates pizzicatos along with the string pad, while large movements bring on the high strings and more reverb. All the sounds are drawn from a Korg M1 simulator synthesiser and are triggered via MIDI interface. The video screen is divided into an 8 X 4 grid, with the vertical lines representing individual notes of a diatonic scale while the horizontal lines control the octaves.

Preset 11

3 voices, 1st - soprano sax, 2nd - alto sax, 3rd - tenor sax. The voices are coming in according to the energy level, i.e. the greater the expression, the bigger the sound. Smaller movements bring in the soprano sax, middle activity activates the alto sax and the maximum activity brings in the tenor sax. This preset has no reverb. The video screen is divided into an 8 X 4 grid, with the vertical lines representing individual notes of a diatonic scale while the horizontal lines control the octaves. All the sounds are drawn from a Korg M1 simulator synthesiser and triggered via MIDI interface.