Music and Medicine
3(3) 182-188

DOI: 10.1177/1943862111401032
http://mmd.sagepub.com

Dynamic Sonification as a Free Music Improvisation Tool for Physically Disabled Adults

Alan Lem, PhD, RMT[1,2] and Garth Paine, PhD[2]

Abstract

This paper reports on the development and initial evaluation of a video-based, dynamic, sonification device used with 5 physically disabled adults, recent clients of Creative Music Therapy (CMT). Of particular interest was the extent to which the dynamic properties of the technology could assist the participants to engage in a dynamic musical interaction at the level of autonomy available to physically able people. Each participant took part in 8, half-hour sessions utilizing free interactive improvisation. During the study, several sonic algorithms were trialed and adjusted according to each participant's movements and preferences. Informing the sonification design was the concept of dynamic orchestration developed by Paine, and real-time sound synthesis. Results indicated that video-based dynamic sonification systems may be used effectively as free improvisation tools with people who have mild physical disabilities, but that modifications may have to be made for people whose movements are more severely restricted.

Keywords
dynamic sonification, free music improvisation, creative music therapy, physical disability

Introduction

Music has been used as a tool for rehabilitation since the beginning of the 20th century, but it is only in its last 2 decades that technology has enabled participation in musical activity from people with serious physical disabilities. Sonification (see note 1) devices such as Soundbeam[1] provide a medium through which even profoundly physically disabled individuals can become relatively communicative using music and sound. The sense of control, agency, and independence which this provides can be a powerful motivator, stimulating learning, development, and interaction in other areas of life.

However, most sonification devices have been designed for creating melodic and rhythmic patterns using predetermined sounds provided by MIDI (see note 2) and a triggering approach which gives a momentary musical response. These characteristics are useful in many areas of music making, such as composition, but they are not suited to the production of live dynamic sound that would change in response to the velocity, acceleration, and direction of gesture. This limits simultaneously the scope of creative musical expression available to the user and their motivation to move,[2] which is at the foundation of participation in many musical activities.

The connection between movement and creative musical expression has frequently been discussed in the literature, particularly in relation to improvisational music therapy. Pedersen[3] proposed that the understanding of movement allows one to understand the connection between feelings and the body. Trevarthen and Malloch[4] commented that the "...interplay between [...] sounds and bodily movement acts as a medium of the musical relationship."(p11)

The aim of the present study was to develop a dynamically responsive system that would enable people with physical disabilities to engage with virtual musical instruments in a way that accounts for their movement vocabulary and sonic preferences. As dynamic expression was investigated and given the specific needs of the participants, the system was conceived as a free improvisation tool, prioritizing the overall sonic expression, such as the intensity and sharpness of sound, over the more conventional musical structures incorporating the elements of melody, harmony, and rhythm. For the same reason, the system was built based on real-time sound synthesis (see note 3), which, unlike predetermined sonic morphologies (eg, wave-form samples controlled by MIDI), lends itself to the production of dynamic sonic outputs.[5]

---

[1] Golden Stave Music Therapy Centre, University of Western Sydney, Penrith South DC NSW, Australia
[2] School of Communication Arts, University of Western Sydney, Penrith South DC NSW, Australia

Corresponding Author:
Alan Lem, Golden Stave Music Therapy Centre, University of Western Sydney, Locked Bag 1797, Penrith South DC NSW 2751, Australia
Email: a.lem@uws.edu.au

Given the above, video tracking became the technology of choice for gathering movement and gestural behaviour. It is noninvasive, easy to set up, and it provides the sampling rate of 25 to 30 frames per second, which is sufficient to reflect changes in the dynamic quality of most human gestures.

Method

Participants

Participants were 5 adults (2 males and 3 females) in their mid-twenties. All were current clients of Creative Music Therapy (CMT) and were accustomed to participating in musical interaction using acoustic instruments and free improvisation. In line with the practice of CMT, all participants had been previously assessed for specific levels of participation[6] and had individual therapeutic objectives established in areas such as speech, movement, and general development. According to their assessments, all participants had some form of cerebral palsy. Four participants were confined to wheelchairs and had complex physical problems resulting from quadriplegia of spastic and athetoid forms and a degenerative motor disease of unspecified origins, whereas the fifth participant was able to move independently and had a mild form of hemiplegia. In addition to the physical disability, 2 participants had a moderate form and one severe form of intellectual disability.

Before being accepted for the study, each participant, or their carer, signed 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 current physical ability, musical preferences, and ways of expressing emotions.

Procedure

Each participant took part in 8, half-hour individual session conducted mid-morning at a university-based facility. The sessions were conducted in a large room equipped with a grand piano, a selection of drums, and percussion instruments. Apart from the participant, there were 4 people in the room: the primary researcher, a software developer, and 2 students who were recording the sessions on video and transcribing verbal interactions that took place between those present. The primary researcher was a skilled music improviser and music therapist with 22 years of clinical experience.

During the sessions, the participants stayed in their wheelchairs, except for the mobile participant who was sitting in an ordinary chair or moved freely around the room. A small video camera was pointed at the participant from approximately 2 meters away, as shown in Figure 1.

During the first 2 sessions, the researcher would typically demonstrate the sonification system and the ways to generate sound, but he would not suggest to the participant any specific movement responses. Once the participant started using the system independently, the researcher would join him or her in a free music improvisation on the piano or another instrument deemed appropriate. While playing, the researcher's intention was to offer the participant musical support and engagement similar to that promoted in a music therapy session, although here without any specific therapeutic goal in mind. Participants were free to move and create the sound according to their level of ability.

Figure 1. The layout of the room during the trial.

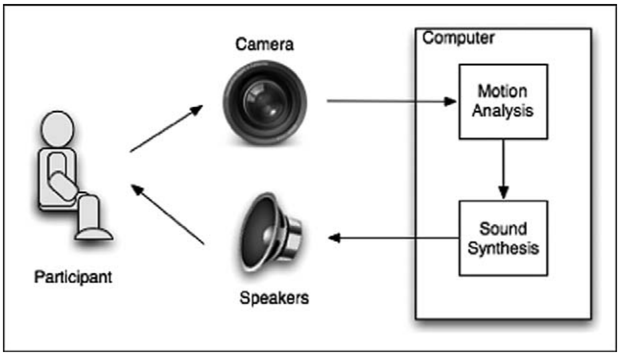

Figure 2. The sonification system overview.

During the course of the study, several sonification algorithms (see note 4) were trialed and adjusted according to each participant's movements and preferences. These adjustments would often involve a conversation between the primary researcher and the software developer, which was typically followed by prolonged periods of silence during which the software developer adjusted the system's parameters.

Materials

Sonification design. The sonification software consisted of 2 components, the motion analysis engine and the sound synthesis engine. The motion analysis engine analyzed the video stream in real time, outputting data streams that were mapped to the synthesis control parameters in the sound synthesis engine (Figure 2).

Informing the design of the system was the concept of dynamic orchestration,[7,8] which allows for dynamic changes in the gesture registered by the system to be interpreted as dynamic changes in the voicing of the underlying synthesis

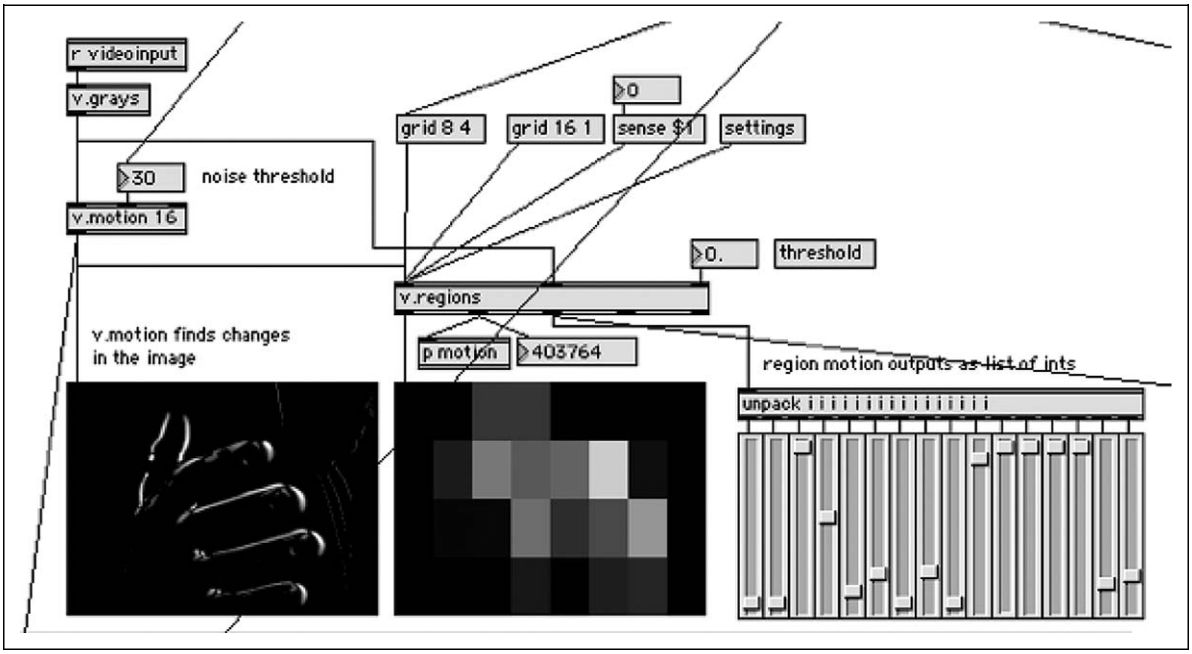

Figure 3. Video window being analysed as an 8 × 4 grid.

model. Here, the amount and the location of movement were mapped out to the synthesis parameters. The voicing structure could be further changed dynamically in relationship to changes in the measure of motion detected.

Motion analysis engine. To enable the sonification of participants' gestures, the researchers used video tracking[9-11] techniques, which provided temporal measures of the amount of movement and calculated the centre of that movement within the camera's field of view.

A Pointgrey, Dragonfly FireWire (IEEE 1394 interface) camera was used to capture the video with the softVNS[12] library of objects for Max/MSP[13] being used to programme the motion analysis software engine. Motion detection was performed using frame differencing,[14] with the video window being analyzed in regions, as an 8 × 4 grid (Figure 3). Consequently, motion detection could be applied individually to each of the resulting 32 cells in addition to the whole video window. The advantage of this approach was that it allowed for the separation of objects within the space, facilitating individual limb tracking, while simultaneously analyzing the total overall motion. It also allowed for masking segments of the video image where analysis was not desired, and grouping the grid positions for higher level analysis.

To ensure the appropriate data scaling of large and small gestures in varying lighting conditions, a peak detection algorithm was implemented as part of the motion analysis engine. This allowed the primary researcher to reset the system at the beginning of each session.

Sound synthesis engine. The sound synthesis engine was based on a 3-voice (layers) model in which the overall mix of the voices was determined by the amount of movement detected by the system (Figure 4), and where each layer corresponded with a unique synthesis algorithm. In general, small gestures were interpreted as Layer 1 while increasing amounts of movement would bring in Layer 2, with the maximum of movement bringing in Layer 3. Each voice also had an independently controllable reverb and spatial mix.

The system was designed to support a variety of different synthesis approaches. A custom-designed granular synthesizer based on FOF (Fonction d'Onde Formatique, translated as Formant Wave-Form or Formant Wave Function) synthesis[15] was incorporated in many of the presets, while a custom-designed basic sampler enabled the inclusion of audio files (eg, piano, drums, etc). This model provided a number of control inputs that were mapped to the outputs of the video analysis engine, resulting in a set of rich sonic possibilities. Following a consultation with music therapy practitioners, the richness and subtlety of the synthesis engine was reduced in order to produce more contrasting sonic variations; it was perceived that participants selected for the study would better respond to the relationship between gesture and sonification through dynamic timbral variations.

User interface. During the sessions, the system's operations were being adjusted by the computer programmer via a graphic interface displayed on a touch screen (Figure 5). This involved the adjustment of the system's main parameters, specifically the different presets, the master volume, sensitivity, and recalibration. The interface also provided graphic feedback of the system's internal states, showing the amount of movement measured, the amplitude of the main audio output, and the alternate camera and analysis views.

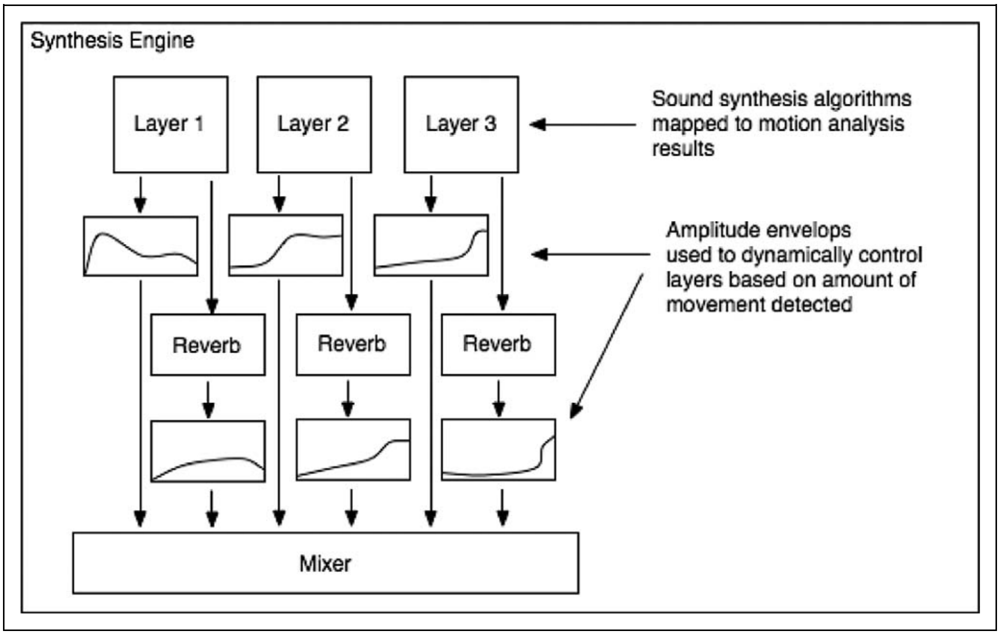

Figure 4. Dynamic orchestration of the three sound layers.

Data Analysis

Each trial session was observed and transcribed by 2 assessors who noted down verbal instructions given to the participants, the characteristic features of their movements and the quality of the corresponding sound. This assessment was conducted twice, during the session and retrospectively, by viewing the video recording. The 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. The characteristic features of participants' movements were identified in the following process:

1. The transcript from each session was read to gain a general sense of movement vocabulary;
2. Key statements describing specific movements were underlined, eg "Made a fast, large circle in the air with the left hand";
3. The underlined key statement was categorized according to the perceived magnitude, length, and intensity of the gesture, similar to the concepts of Labanotation;[16]
4. Each category was given a specific heading, for example "Large Circular Hand Movement."

Following the identification of movement groups, the researchers looked at the corresponding sonic output, which was classified according to the perceived intensity and spaciousness of sound, the closest corresponding sound produced by a conventional musical instrument, and the immediacy of system's responsiveness to the movement. The data were then cross-analyzed to understand if and how specific characteristics of sound produced by the computer might have influenced the participant to move in a particular way and how this might have changed over the period of the study.

Results

The analysis of data obtained from participants with complex physical problems yielded no consistent results. All 4 were observed to engage with a high level of enthusiasm during the first 2 sessions, but the variability of their responses, including the type and the intensity of movement not only remained unchanged, but showed a tendency to decrease over the period of the study. It was also observed that these participants became noticeably exhausted after the first 5 minutes of a session.

The data collected from the mobile participant with hemiplegia were different, and showed that his movements were clearly changing as a result of the sonic possibilities put before him. Generally, it was observed that the incidence of his gross motor responses decreased in favor of more refined, smaller gestures. There were also a number of relationships between the specific character of his gestures and the sonic output of the computer.

Firstly, his large arm movements (Figure 6) occurred with the highest frequency during spacious sounds imitating marimbas and strings, and when the responsiveness of the system was proportional to the dynamic envelope of his gesture.

His second gesture was described as *Punching Movements* (Figure 6). These occurred frequently during loud sounds resembling a crashing metallic wave and a fast system response.

His smaller arm movements (Figure 7) occurred frequently during sounds imitating marimbas and when the system responsiveness followed the envelope of his gesture.

Finally, the mobile participant frequently moved his legs (Figure 7) during a spacious sound described as *Washover*, which corresponded with a slow responsiveness of the system.

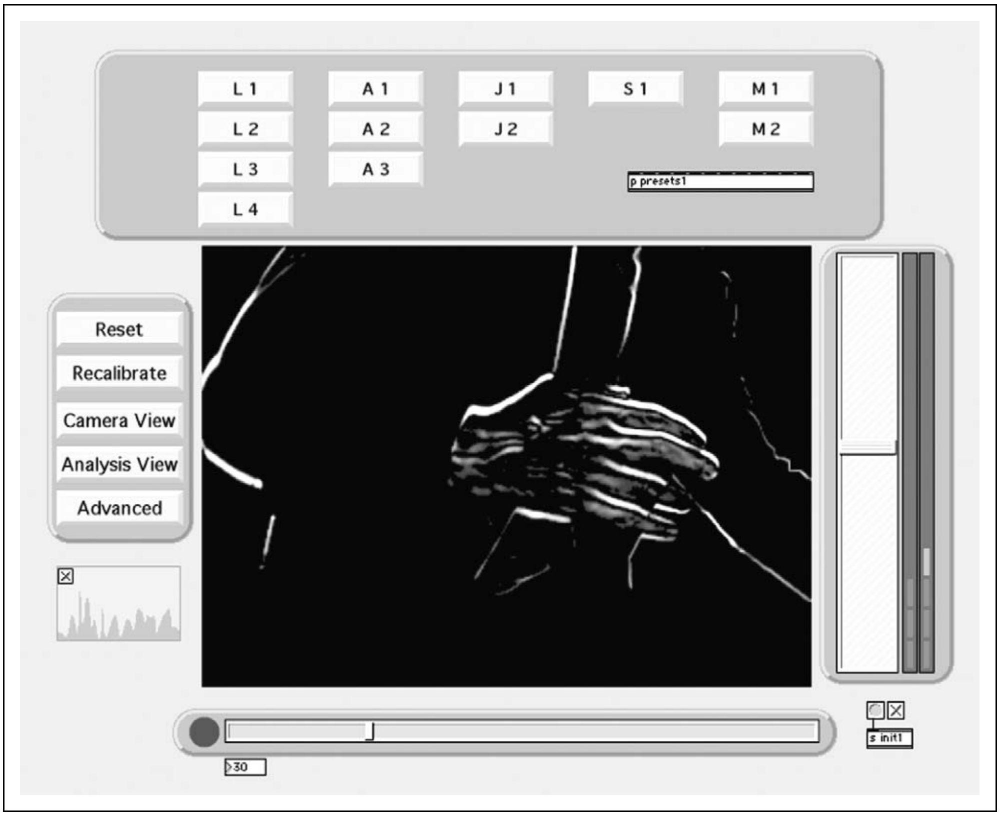

Figure 5. Main touch screen interface.

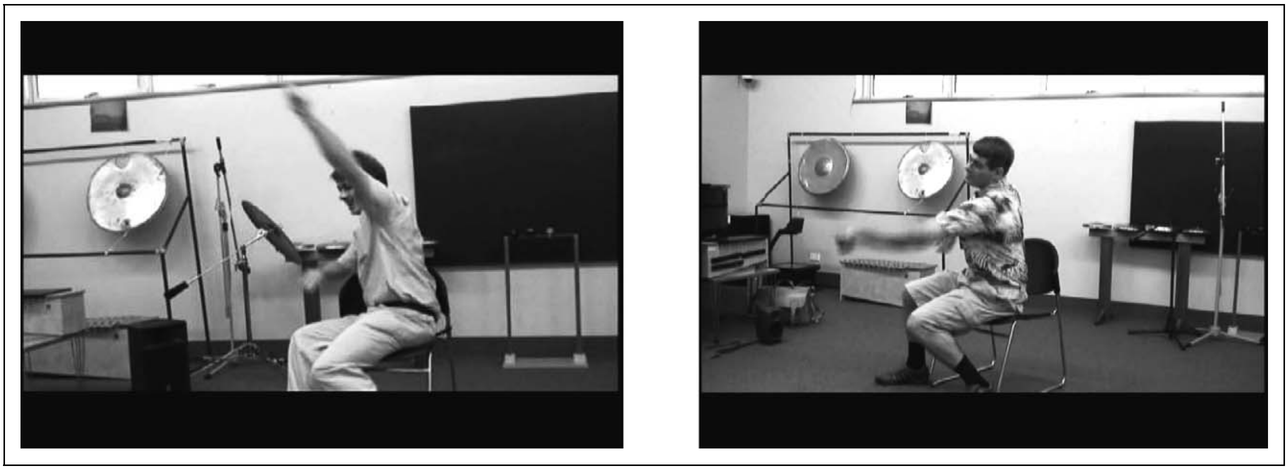

Figure 6. Large arm movements (left) and punching movements (right).

Discussion

The present project was a trial study aiming to develop and evaluate a dynamic, sonification device as a free improvisation tool with physically disabled people. The project was unique in that it aimed to develop a new technology with the assistance of disabled participants who traditionally would be asked to use it only after the system has been refined and made ready for application in its target role. This was one of the most important constraints of the study which resulted directly from its limited funding.

Data obtained from the mobile participant with hemiplegia revealed clear correlations between the sonic output of the computer and 4 categories of movement: large and small arm movements, leg movements, and punching. It also indicated that during the study, the frequency of his gross motor responses decreased in favor of more refined movements, which indicates learning. This observation is directly supported by the analysis of the sonic characteristics of the computer's output, which shows that his finer movements occurred with the highest frequency when the dynamic responsiveness of the

Figure 7. Smaller arm movements (left) and leg movements (right).

system was directly proportional to the dynamic quality of his gesture. This indicates that during the course of the study, not only the vocabulary of his movements but also the sonic possibilities of the technology had undergone a refinement. A video excerpt showing the mobile participant using the system in the final stages of the project can be viewed at Nordoff Robbins Music Therapy Australia.[17]

Data obtained from participants who had complex physical disabilities indicated that their overall level of engagement and their variability of gestures had decreased with time. It was also observed that these participants became exhausted quickly after the beginning of each session. This might have resulted from the severity of their physical disability which, in retrospect, could have been helped by focusing on their smaller gestures. However, it could also result from the fact that a large portion of each session was devoted to conversation and computer programming, due to which it was not always possible to engage participants in a spontaneous and fluent music improvisation.

The development of a suitable sonification technology for a project of this nature is a difficult task and many variables need to be considered in an interrelated manner. The present system was built as a free improvisation tool given that tasks incorporating conventional musical elements, such as melody, harmony, and rhythm, were deemed not appropriate for all the participants. However, considering the level of tiredness observed in participants with complex physical disabilities, it is believed that the design could have incorporated a gesture follower and a task-oriented approach (such as drawing a circle in the air), which might have been easier to relate to by them. Indeed, it is possible that these people would have responded with a higher degree of motivation if extra visual stimulation was available to them. Here, the authors would like to acknowledge that a more unified group of participants would have likely produced a greater transparency of the results. In the final recommendation, the authors would like to propose that future trials involving dynamic sonification technologies invite the participation from people with severe physical disabilities only when the technology is ready to use.

Acknowledgments

The authors acknowledge the participants, Spastic Centre of Australia, Golden Stave Music Therapy Centre, MARC Auditory Laboratories, UWS Research Office, and Nordoff-Robbins Music Therapy Australia.

Declaration of Conflicting Interests

The author(s) declared no conflicts of interest with respect to the authorship and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research and/or authorship of this article: This research was funded by the University of Western Sydney and Nordoff Robbins Music Therapy Australia through a partnership grant.

Notes

1. The concept of sonification relates to the use of audio generated by means other than speech to convey information, for example movement.
2. MIDI stands for Musical Instrument Digital Interface, a protocol designed for recording and playing back music on digital synthesizers. Rather than representing musical sound directly, it transmits information about its velocity and duration, but the sound itself comes from the receiving instrument or computer sound card.
3. Real-time synthesis refers to a protocol by which the sound wave is being generated and modified electronically by the computer during its use rather than relying on predetermined wave-form samples.
4. Here, the word algorithm refers to a mathematical formula used to convert the data points obtained the participant's movement to a sonic output.

References

1. Soundbeam Project. http://www.soundbeam.co.uk/downloads/leaflets/special-needs-leaflet.pdf. Accessed October 18, 2010.
2. Lem A. A dynamic sonification device in improvisational music therapy. Paper presented at Music Technology: Solutions to Challenges Conference; June 12, 2010: London.

3. Pederson IN. Psychodynamic Movement—A Basic Training Methodology for Music Therapists. In: Escjen J, ed. *Analytical Music Therapy.* London: Jessica Kingsley Publishers; 2002:190.
4. Trevarthen C, Malloch SN. The dance of wellbeing: defining the musical therapeutic effect. *Nordic J Music Therapy*. 2000; 9(2):3-17.
5. Paine G. Gesture and Musical Interaction: interactive engagement through dynamic morphology. Paper presented at the New Interfaces for Musical Expression; 2004:Hamamatsu.
6. Nordoff P, Robbins C. *Creative Music Therapy.* 2nd ed. Bristol, UK: Barcelona Publishing; 2007.
7. Paine G. Interactivity, where to from here? *Organised Sound*. 2002;7(3):295-304.
8. Paine G. Sonic immersion: interactive engagement in real-time immersive environments, scan. *J Media Arts Culture*. 2007;4(1). http://scan.net.au/scan/journal/display.php?journal_id=90.
9. Dinkla S. The History of the Interface in Interactive Art. Proceedings of the 1994 International Symposium on Electronic Art (ISEA). 1994: Helsinki.
10. Packer R, Jordan K. *Multimedia: From Wagner to Virtual Reality.* New York: W W Norton & Company; 2002.
11. Paine G. Interactive sound works in public exhibition spaces: an artists perspective. Proceedings of the 2001 Australasian Computer Music Conference. University of Western Sydney: Australasian Computer Music Association, 2001:67-73.
12. Rokeby D. softVNS. 2002. http://homepage.mac.com/davidrokeby/softVNS.html. Accessed October 21, 2010.
13. Bevilacqua F, Müller R, Schnell N. MnM: a Max/MSP mapping toolbox. Paper presented at the International Conference on New Interfaces for Musical Expression; 2005: Vancouver, BC.
14. Gonzalez R, Woods R. *Digital Image Processing.* 3rd ed. Upper Sadddle River, NJ: Prentice Hall; 2007.
15. Clarke M, Rodet X. Real-Time Fof and Fog Synthesis in Msp and Its Integration with Psola. Proceedings of the Music Conference (ICMC2003). Singapore, 2003.
16. Newlove J. *Laban for Actors and Dancers: Putting Laban's Movement Theory into Practice.* London: Nick Hern Books; 1993.
17. Nordoff-Robbins Music Therapy Australia. Music Therapy at the Golden Stave Music Therapy Centre in Sydney–Individual Sessions (4) [video]. Session Recordings; March 2009. http://www.youtube.com/watch?v=VFnsqt_NZFo. Accessed January 25, 2011.

Bios

Alan Lem, PhD, RMT, is a music therapist, supervisor and composer with deep interest in improvisation; he heads the Master of Creative Music Therapy in the University of Western Sydney.

Garth Paine, PhD, is highly regarded internationally as an innovator in the area of interactive music systems and video tracking for gesture based musical performance and is Senior Lecturer in Music Technology at the University of Western Sydney.