Northumbria Research Link

Northumbria University
NEWCASTLE

University Library

Running up Blueberry Hill: Prototyping Whole Body Interaction in Harmony Space

Simon Holland, Paul Marshall, Jon Bird, Sheep Dalton,
Richard Morris, Nadia Pantidi, Yvonne Rogers, Andy Clark†

Department of Computer Science
The Open University
Milton Keynes, MK7 6AA, UK
{s.holland, p.marshall, j.bird, n.dalton,
richard.morris, k.pantidi, y.rogers}@open.ac.uk

†Department of Philosophy
The University of Edinburgh
George Square
Edinburgh, EH8 9JX, UK
Andy.Clark@ed.ac.uk

ABSTRACT

Musical harmony is considered to be one of the most abstract and technically difficult parts of music. It is generally taught formally via abstract, domain-specific concepts, principles, rules and heuristics. By contrast, when harmony is represented using an existing interactive desktop tool, Harmony Space, a new, parsimonious, but equivalently expressive, unified level of description emerges. This focuses not on abstract concepts, but on concrete locations, objects, areas and trajectories.

This paper presents a design study of a prototype version of Harmony Space driven by whole body navigation, and characterizes the new opportunities presented for the principled manipulation of chord sequences and bass lines. These include: deeper engagement and directness; rich physical cues for memory and reflection, embodied engagement with rhythmic time constraints; hands which are free for other simultaneous activities (such as playing a traditional instrument); and qualitatively new possibilities for collaborative use.

Keywords

Harmony Space, whole body interaction, embodiment, music

INTRODUCTION

People are well-adapted to dealing with the physical world, as opposed to abstract concepts. This motivates the design and investigation of interactive tools based on embodied metaphors that exploit human physical capabilities so as to facilitate otherwise difficult abstract tasks. Harmony is widely considered to be one of the hardest and most technically complicated parts of music to master. The principles of musical harmony are generally taught in a symbolic manner using a specialised notation and terminology via abstract, domain-specific concepts, principles, rules and heuristics. By contrast, when harmony is visualised, analysed, manipulated and created using

metaphors embodied in the existing interactive desktop tool Harmony Space [1,2] which is grounded in two well-established theories of music cognition and perception [3,4] a new, parsimonious, but equivalently expressive, unified level of description emerges [1,5,6]. This metaphorical level of description focuses on moving objects; locations; centres, trajectories; and moveable 'allowed' and 'forbidden' areas, to be navigated while meeting rhythmic time constraints. This new approach makes it possible to characterise disparate technical musical concepts (such as scales, chords, triads, tonal centres, chord sequences, bass lines, harmonic progressions, modes and modulations) using a single, consistent, unified spatial metaphor.

This paper presents the trial of a prototype version of Harmony Space which uses whole body navigation to carry out a range of musical tasks. Following Papert's notion of body-syntonic learning [7], we were interested in whether participants could exploit their situated sense of space and how their bodies move to gain a deeper understanding of harmonic relationships. A previous 'human-powered' pilot study in Utrecht [5] with no computer-based elements, using physical labels on the floor, a large manually-moved wooden frame, and musicians employed in a Wizard-of-Oz role, demonstrated the potential of whole body navigation of Harmony Space for free composition by musically expert players. In the present study, new goals included identifying possible benefits, both to beginners and accomplished musicians, that a more flexible computer-based whole-body version of Harmony Space might provide.

The central aim of the current trial was to explore design requirements in adapting Harmony Space from a desktop system controlled with a mouse and keyboard to the medium of whole body interaction. This process – taking techniques and practices that work in one medium and applying them in another – can be a productive approach in exploring the affordances of a new medium [8]. We employed a Wizard-of-Oz approach to identify and characterize some of the new opportunities that whole body interaction presents in the context of the principled manipulation of chord sequences and bass lines.

Figure 1: left – the Harmony Space representation projected onto the ground; right – the projector on the top floor of the atrium, with two metal arms holding a mirror for downward projection.

WIZARD-OF-OZ STUDY DESIGN

Physical design

The Harmony Space Twister interface was prototyped in a large atrium, and focused on a floor display area about 6 metres by 4 metres (figure 1, left), illuminated via a mirror by a powerful data projector situated some 10 metres overhead (figure 1, right). White sheeting was taped down over the carpet tiles to provide a high contrast surface for image projection. The principal element of the projected image is a fixed grid of some 12 x 15 circles labeled with note names (musicians interested in the details of note labeling might wish to refer to the Limitations section) Circle labeling may be varied depending on the particular task, and the projection image includes other, dynamically moving features that we will describe later. For the purposes of this design study, a series of predominantly single player tasks were presented as games.

The System

The engine driving the system was the 2008 desktop implementation of Harmony Space, which was designed and implemented by the lead author (figure 2). This is coded in Squeak Smalltalk and runs platform independently on Mac and Windows machines. Squeak's internal software FM synthesiser is used to drive the audio. One option was to use the ReacTIVision [9] camera-based tracking system via the existing Squeak implementation of the Tuio protocol [10] to track players' movements - which is now integrated into Harmony Space for that purpose, but for this design study we chose to use Wizard-of-Oz position tracking (i.e. a human operator of the desktop version of Harmony Space tracked the position of players by eyeball). This allowed us to explore the implications and possibilities of different candidate tracking mechanisms (e.g. camera-based head tracking, camera-based foot tracking, shoe-sensors, pressure mats, etc).

The desktop Harmony Space representation used in this study has a number of key features, which are described in more detail in [1,2]. In the present system, these include representing the tonal centre for a piece of music by circling the note in red. Tonal centres were described to the participants in this study as 'home', and as good places to start and finish a song. A second feature noted to
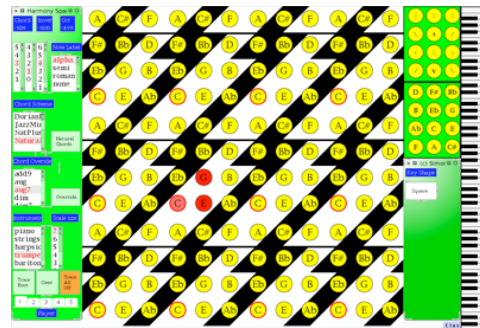

Figure 2: Desktop Harmony Space screenshot showing a C major chord.

participants was the differentiation into white and black areas, referred to as 'allowed' and 'forbidden' areas respectively. It was emphasised that the terms 'allowed' and 'forbidden' were only rough guides. These areas correspond in the simplest case to the white and black keys on a piano – though they move systematically when the key changes. It was finally noted that there are only 12 distinguishable notes in the display. However, it is vital that this basic pattern is repeated like wallpaper to allow the long information-rich trajectories found in tonal music to be more easily seen, played, and analysed.

Interface and task design

In general terms, each task or game focused on a specific song (e.g. Pachelbel's Canon, Michael Jackson's 'Beat It', Fats Domino's 'Blueberry Hill'). The player's task was to navigate over the terrain of the projected surface in such a path as to generate an appropriate bass line (or chord sequence) in time to the playback of the song. In general, the tasks were achievable by musical beginners and experts alike.

We explored different configurations for different tasks. For example, for the simplest tasks (denoted 'bass line only'), when a player stepped on a note circle it was illuminated, and the corresponding bass note sounded. When the player's foot left that circle, the relevant bass note stopped sounding, but the note stayed illuminated so as to leave a visual trace of the path taken. In these tasks, the goal of the player was to learn to perform an appropriate bass line in time with the playback of a song by dancing, walking or jumping in an appropriate path. In exploring different variations of this task, the path was indicated to the prospective player in various different ways. For example, in the 'preview' variation the player could study the whole song, or parts of it in advance by watching the circles illuminating in sequence in time with the music to indicate a correct path. By contrast, in the 'just-in-time' variation, the player was shown which circle to step on next by the lights illuminating just before it was time to move. The interface tended to attract a crowd, so a third way to learn a path was by the 'social' variation, where a coach or peer pointed out, described, or physically demonstrated a possible path.
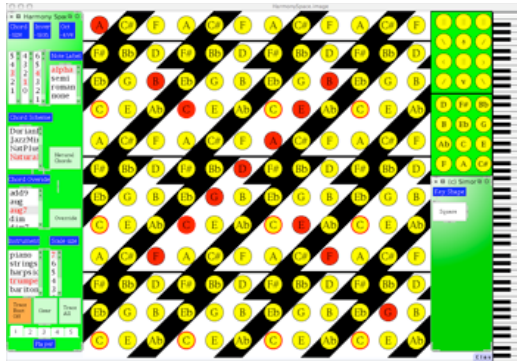
Figure 3: Diagonal paths corresponding loosely to 'Harmonic' ↗ and 'melodic' ↘ movement.

In more complex tasks ('chord sequences'), stepping on a circle elicited not just a bass line but a chord consisting of three or four notes played and highlighted simultaneously. In such cases, a small piece of terminology is helpful - the note stepped on is known as the 'root' of the chord. The shape of the chord produced varies depends on the position of the root within a bounding box known as the key window (the white area, as opposed to black area in figure 3). Differences in chord shape can be seen visually, and heard aurally. This variation in elicited chord shape follows a regular and immediately visually obvious rule – the shape is constrained by the key window, somewhat like water in a container. Two display variants using chords were tested – in one case, 'show all notes', all of the simultaneously sounding notes were shown, but no persistent trace was left. In the other variant, 'root only', the whole chord was sounded, but only the root was illuminated and traced (to make the overall path clear). Generally, the path taken remained illuminated until trace was cleared – except in "show all notes" mode, where the notes of the chord and the notes of the path would have obscured each other.

Participants
The study focused principally on a single participant (participant A) for a two-hour period. However, two watchers of the study (which was in a public atrium space) spontaneously joined in to demonstrate by example alternative paths for some songs to the paths suggested by the coach (one of the authors) or worked out by the participant, so that in effect there were two secondary participants (participants B and C). Participant A did not have any experience in playing a musical instrument, but had done some sol-fa singing at school, and knew the Greek notes names. Participants B and C had some experience of playing guitar and electric bass respectively; previous experience with the desktop version of Harmony Space suggests uses for both beginners and experts.

The trial
As a warm up activity and preliminary orientation, the coach instructed the participant to walk and listen to the paths generated by walking both diagonals through the space, both up and down – on the initial pass stepping in

forbidden and allowed areas equally, and on a second pass avoiding the forbidden areas (shaded black) – see table 1. It was mentioned in passing that the two diagonals are very loosely related to melodic vs. harmonic (or bass) movement. In musical terms, these paths correspond to various banal but useful note sequences which are musically ubiquitous in small fragments (see figure 3 and Table 1). For a more complete treatment of the kinds of harmonic relationships exposed by the Harmony Space representation, see (Holland, 1989).

Diagonal axis	Sense	Areas marked in black	Musical result
'Melodic' ↘	up or down	Use	Chromatic scale up or down
'Melodic' ↘	up or down	Avoid	Major scale up or down
'Harmonic' ↗	up	Use	Real cycle of fifths
'Harmonic' ↗	up	Avoid	Diatonic cycle of fifths
'Harmonic' ↗	down	Use	Real cycle of fourths
'Harmonic' ↗	down	Avoid	Diatonic cycle of fourths

Table 1: Musical effects of some rudimentary paths.

During the brief warm up, other landmark notions about the Harmony Space representation were briefly pointed out, including the root circled in red and the repeating pattern of 12 distinguishable notes. To trial a camera-tracking feature, the participant was asked to use a hand gesture for muting the generation of sound. This involved putting her hands on her head as if covering a fiducial marker [9,10] (in the event, this signal was often given orally to the Wizard of Oz operator).

The participant was then asked to use the whole body Harmony Space representation to play either the bass line or chord sequence of five example songs. The songs used for the trial were chosen to exhibit a variety of clearly distinguishable but thematically related trajectories in Harmony Space. For example, Pachelbel's Canon moves harmonically in a regular zig zag (avoiding the black area outside of the current key) followed by a straight-line trajectory to home (see figure 4).

The differences and similarities of paths between the songs relate directly to their harmonic structure. This reflects the fact that a path in Harmony Space that plays a correct chord sequence (there is generally more than one such path) corresponds at one level of abstraction to a functional harmonic analysis of the piece.
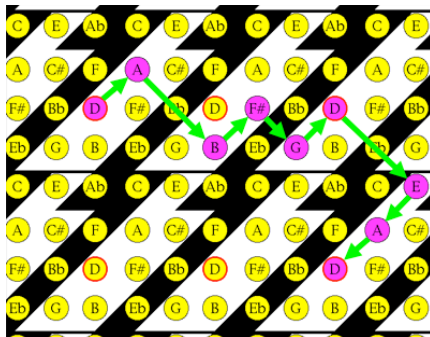
Figure 4: Path of chord roots of Pachelbel's Canon in Harmony Space.

OBSERVATIONS FROM TRIAL

From the Wizard-of-Oz study, we were able to draw out a number of themes that have implications for the development of our planned high-tech version of Whole Body Harmony Space.

Individual differences between participants

There were a number of differences between the four different players (three participants plus coach) who used Whole Body Harmony Space during the two-hour session. Firstly, they differed in their style of moving through the space. The principal female participant used a fluid, dance-based, highly rhythmic form of movement in the faster songs while walking and leaping during the slower numbers. Two male participants used relatively minimal precise planned movements to try to hit the targets at the right times. One male participant used a very exuberant dancing style. In order to suit different group of users, the high-tech version of Whole Body Harmony Space should therefore allow for both precise and more exuberant or dance-like forms of interaction.

The participants' stride length was also observed to vary considerably. This was no problem for any participant making single steps in any of the eight compass directions, but for two particular musical intervals (the whole tone and the tritone) steps of size two are required. Steps of size two on the diagonal in particular required some agility from the somewhat petite principal female participant. This did not create any real problems in the trials, but clearly there is a trade-off between engaged physicality vs. challenges for older, smaller or younger participants. There may also be related trade-offs in precision. One design possibility might be to calibrate the size of the Harmony Space Representation to the physical capabilities of users.

Finally, different participants had markedly different strategies for constructing trajectories for playing the same bass lines. For example, in playing Michael Jackson's "Billie Jean", the coach had envisaged the fragment of the path represented in fig. 5, left which demonstrates the harmonic structure clearly from a music-theoretic point of view. However, the step between Ab and B was physically inconvenient, so participant A used the physically more convenient alternative path shown in fig. 5 (centre). Participant B intervened spontaneously to suggest a third alternative, shown in fig. 5 (right), on the grounds that that this would minimize physical inconvenience. However participant A resisted for two reasons, arguing that this would mean starting each fresh iteration of the pattern in a new position which would be harder to remember and that such a routine would eventually "run out of road" when it reached the side of the display area. Participant B maintained his position, arguing firstly that it was simpler to execute, and secondly that two or more players could collaborate to continue the pattern when one "ran out of road". Apart from differences in their physical and organisational convenience, these three different but equivalent paths may be seen as emphasising different aspects of the harmonic relationships on the song. We suggest that discussions stimulated by such differences of opinion are one of the benefits in translating Harmony Space from a desktop system to a whole body system, encouraging participation and experimentation by players and bystanders (cf. [11]).

Memorability of different harmonic structures

Different songs can have very different paths in Harmony Space. When songs were learned in preview mode or social mode, where the sequences, or parts thereof, were learned in advance, there were clear differences between the physical memorability of different paths. Those songs that were found to be easiest to memorise during the trial were those based principally on simple straight-line trajectories. Those ostensibly simpler patterns where the participants had merely to move single steps away from 'home' in repeated patterns were found to be hard to remember, with participants often forgetting whether to move up or down the harmonic axis. Songs with more than one straight-line trajectory, such as Stevie Wonder's "Isn't she lovely" were found to be of intermediate difficulty. This suggests that when introducing beginners to Whole Body Harmony Space, it makes sense to take advantage of the embodied cognitive economies of straight-line trajectories before moving onto more complex paths.
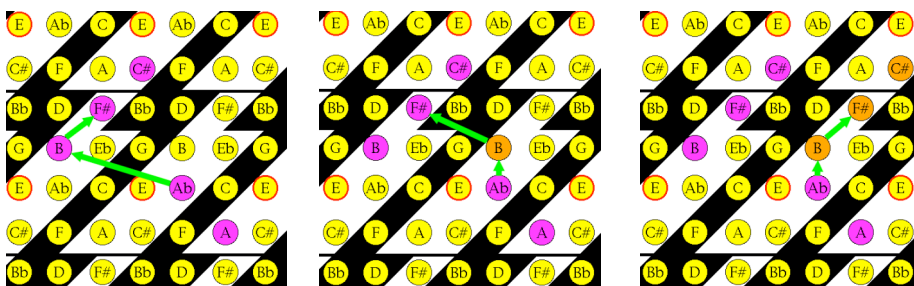

Figure 5: Left: 'expert path'; Centre: first alternative path; Right: second alternative.
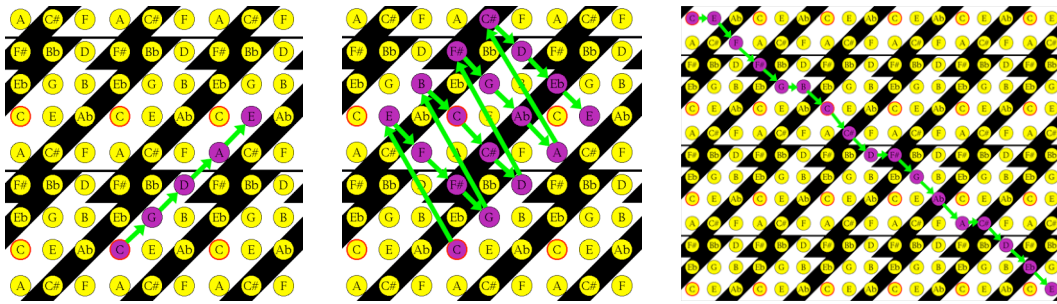
Figure 6: "Hey Joe". Left – underlying trajectory; Middle – with nested chromatic trajectories; Right – 'melodic' path.

Keeping bearings during modulation and changes of trajectory

When the key window and its complementary black area moves (whenever the key is changed) it is relatively straightforward to visually grasp what has happened when using the desktop version of Harmony Space. Subjectively, key changes are at first more disorienting in the whole body version, even when anticipated. However, having experienced this, participant A came up with strategy for avoiding confusion. She reported that she simply fixed her eyes on the note names, and ignored shifts of the black and white areas when working out where to step. One design change that might reduce the disorienting effects of key changes is by smoothly animating rather than jumping the key window to the new location, giving the user more progressive cues to the change.

Similarly, while in the desktop version of Harmony Space, movements in any direction are equally straightforward to perform through the interface, in the Whole Body prototype, the ease of a particular movement depends upon the current orientation of the participant. In particular, most mistakes were made when the next chord or bass note to be played was located behind the participant. This finding demonstrates a trade-off inherent in moving between different media. While the whole body version of Harmony Space may support better memorization of songs and hands free interaction, it also potentially makes movement in the space more demanding. This is analogous to the differences inherent in planning a route using a map and physically walking the route in the real space.

Playability

The speed of bass lines and chord sequences that could be played was limited partly by the speed with which players could move. For example, a fully accurate version of the bass line to "Hey Joe" would nest rapid sideways chromatic trajectories (Figure 6, middle) into the fundamental upwards straight-line trajectory (figure 6, left). Figure 6 (middle) shows the detailed path laid out for maximum clarity of the harmonic structure, but navigation of this version is physically impractical for a single player, due to the speed required for the silent moves to get in position for each chromatic trajectory. However "relay players" could collaborate to achieve this. Alternatively, extended sections of the musically equivalent path (fig 6, right) (which

stresses the melodic aspects of the bass line while de-emphasising its harmonic aspects) are physically workable for a single player, but they would "run out of road" before finishing the pattern. Again, two or more players could execute this path working as a relay. In the trial, the simple path shown in figure 6, left was used. Another possibility for studying and playing pieces otherwise too fast to play faithfully would be to slow the tempo digitally without the altering pitch or harmonic structure.

For purposes of simplicity during the trial, moving to each note circle generally sounded the relevant bass note or chord without adornment – or in some cases decorated with a simple repeated rhythmic figure. There are many other possibilities to give the user more control over rhythm or produce a more pleasing result – for example hand slaps on sensors on the body or repeated foot strikes could be used to control rhythm, or a collaborator could use body movement to modulate the notes produced with different rhythmic figures or melodic figures or accompaniment patterns at different points in the song. Alternatively, automatic accompaniment patterns could be used.

Tracking Issues

We experimented with using hand gestures (e.g. covering the head with both hands) as a way of mimicking the kinds of interaction that might be used in a camera-based tracking system to stop the system tracking the user and mute the music momentarily while moving to another location. Generally, this was found to be too fussy by the participants in the study. Therefore, we suggest that some other mechanism should be used to mute music generation (e.g. a button held in the hand).

Experience in the trial suggests that tracking footfalls, e.g. by pressure mats, shoe sensors or camera, may reflect players' intentions better than tracking head position (e.g. via a fiducial marker on a hat). In some cases, where head position remained intermediate between two note circles, players clearly used footfalls in time with the rhythm to indicate note transitions correctly. On the other hand, it was clear that foot position might sometimes be occluded relative to a ceiling-based camera. A related problem with using a single overhead projector was that players' shadows sometimes occluded the labeling of nearby features. Three possible design changes to address this problem include:

larger note labels; multiple projectors - although this would come at the cost of greater complexity and calibration issues; or labels fixed to the floor – though this would rule out the use of dynamic labeling essential for dealing with more complex harmonic material.

DISCUSSION AND CONCLUSIONS

The purpose of this trial was primarily to explore design requirements in adapting Harmony Space to whole body interaction, and also to gather preliminary evidence on any benefits and key differences compared with the present desktop system. The trial had a number of limitations: position tracking was by Wizard-of-Oz rather than automatic mechanisms; most data gathering focused on a single subject, though spontaneous discussions between the four players were a useful source of information; and the trial was short – just two hours – so that phenomena which may emerge with practice will have been missed. There were also various technical musical limitations; labels reading Ab and C#, for example, should be understood as meaning G#/Ab and C#/Db respectively – a simplification to conserve display real estate; bass lines were treated mostly as root progressions or walking bass; for simplicity, altered chords were mostly treated as decorations associated with particular degrees of the scale.

The trial suggested interesting contrasts with desktop Harmony Space. Players unanimously reported finding the whole-body interface and tasks absorbing, attractive, demanding, and fun. The desktop version attracts broadly similar reactions, but the reaction appeared more marked in the whole body case. Participant A noted that whole-body tasks in Harmony Space were both physically and mentally challenging – a physical and mental workout combined. All players, in different, ways appeared to use their whole bodies to keep track of the harmonic tempo of the song. Participant A moved very rhythmically to the music, and commented that this and the involvement of the whole body helped with knowing when a move to a new circle was due. She was able to remember and discuss aspects of specific paths some days after the trail.

The trial suggests that the whole-body version of Harmony Space offers several new opportunities compared with the desktop version. Key differences appear to be: deeper engagement and directness; rich physical cues for memory and reflection, full embodied engagement with rhythmic time constraints; hands which are free for other simultaneous activities (such as playing a traditional instrument); and qualitatively new possibilities for collaborative use.

As regards further work and design recommendations, the trial made it clear that whole body interface design could be refined to offer several contrasting styles of collaboration as follows; simultaneous players each contributing lines to a polyphonic whole; simultaneous players with heterogeneous roles – e.g. navigating, modulating, inverting, chord alteration, contributing rhythmic elements; players sequentially collaborating to navigate complex paths in relay. To help with knowing when to move, it might be useful for the system to communicate the hypermetre or harmonic tempo (i.e. an appropriate beat) percussively. Variations with an added haptic belt are in preparation.

Because of the importance of footfall locations and timing as opposed to the position of the head or the visual centre of gravity of the body, pressure mats or markers on shoes may be useful technologies to consider.

REFERENCES
1. Holland, S. *Artificial Intelligence, Education and Music*. PhD thesis, IET, The Open University, 1989.

2. Holland, S. (2000) Artificial Intelligence in Music Education: a critical review. In Miranda, E. (ed.) *Readings in Music and Artificial Intelligence, Contemporary Music Studies* Vol. 20. Pages 239-274.

3. Balzano, G. J. The Group-theoretic Description of 12-fold and Microtonal Pitch systems. *Computer Music Journal* 4:4. Winter 1980.

4. Longuet-Higgins, H.C. Letter to a Musical Friend. *Music Review*, August 1962 pp. 244-248.

5. Holland, S. Learning about harmony with Harmony Space: an overview. In Smith, M. and Wiggins, G. (Eds.) *Music Education: An Artificial Intelligence Approach*, 1994.

6. Howard, P., Holland, S. and Whitelock, D. (1994) Keyboard Harmony: some applications of computers in music education. In *Musical Times*, Volume CXXXV No. 1817 pages 467- 471.

7. Papert (1980) *Mindstorms: Children, Computers and Powerful Ideas*. Brighton: Harvester.

8. Rick, J. and Rogers, Y. (2008) From DigiQuilt to DigiTile: Adapting Educational Technology to a Multi-Touch Table. In *Proceedings of Tabletop, 2008*.

9. Kaltenbrunner, M. and Bencina, R. (2007). reacTIVision: a computer-vision framework for table-based tangible interaction. In *Proc. TEI '07*, 69-74.

10. Holland, S. (2007) Open Source Squeak implementation of Tuio. http://mcs.open.ac.uk/sh2/squeakmusic.html

11. Sheridan, J.G., Bryan-Kinns, N. and Bayliss, A. Encouraging Witting Participation and Performance in Digital Live Art. In Proceedings of HCI'07, pp. 13-23.