Article:
Ng, K. (2002) Sensing and mapping for interactive performance. Organised Sound, 7 (2). pp. 191-200. ISSN 1355-7718

https://doi.org/10.1017/S1355771802002091

Sensing and mapping for interactive performance

KIA NG

Interdisciplinary Centre for Scientific Research in Music (ICSRiM), School of Computing & School of Music, University of Leeds, Leeds LS2 9JT, UK
E-mail: mvm@kcng.org
URL: www.kcng.org/mvm

This paper describes a trans-domain mapping (TDM) framework for *translating* meaningful activities from one creative domain onto another. The multi-disciplinary framework is designed to facilitate an intuitive and non-intrusive interactive multimedia performance interface that offers the users or performers real-time control of multimedia events using their physical movements. It is intended to be a highly dynamic real-time *performance tool*, sensing and tracking activities and changes, in order to provide interactive multimedia performances.

From a straightforward definition of the TDM framework, this paper reports several implementations and multi-disciplinary collaborative projects using the proposed framework, including a motion and colour-sensitive system, a sensor-based system for triggering musical events, and a distributed multimedia server for audio mapping of a real-time face tracker, and discusses different aspects of mapping strategies in their context.

Plausible future directions, developments and exploration with the proposed framework, including stage augmentation, virtual and augmented reality, which involve sensing and mapping of physical and non-physical changes onto multimedia control events, are discussed.

1. INTRODUCTION AND BACKGROUND

Physical movement, gesture and expression play an important role in stage performances, irrespective of the mode of human communications: verbal or non-verbal (audio or non-audio), or the language used. With the advances in electronic and computing technology, there has been increasing interest in new musical instrument design (Livingston 2000) to augment traditional instruments (Schoner, Cooper and Gershenfeld 2000, Nagashima 2001) with new capabilities, for example triggering digital sound and visual output (Paradiso, Hsiao, Strickon and Rice 2000, Nagashima 2001), as well as new interface designs to provide better ergonomics, and/or offer simpler instrumental control to a wider community of users. With such systems, the modes of interfaces, sensitivities and reactions (output) are highly flexible and can be configured or personalised, allowing better access to musical instrument playing with shorter learning time requirement.

There has been considerable research into sensor-based gestural control for interactive performance (Wanderley and Battier 2000, NIME 2002). The developments and advances in these areas have also been heightened with the development of a range of sensors to MIDI interfaces, including I-Cube (Mulder 1995), AtoMIC Pro (Fléty 2002) and others (ISIDM). Examples of sensor-based performance systems include DIEM Digital Dance system (Siegel and Jacobsen 1998, 1999, Siegel 1999), Toy Symphony (see http://www.toysymphony.net) and many more. Camurri, Hashimoto, Ricchetti, Ricci, Suzuki, Trocca and Volpe (2000) presents a comprehensive background survey of related projects, including STEIM's BigEye (BigEye) and Rokeby's Very Nervous System (Rokeby 1997, Winkler 1997, 1998).

The prototypes described in this paper are mainly focused on visual tracking using live video input, as in BigEye, Very Nervous System and EyesWeb (Camurri, Coletta, Peri, Ricchetti, Ricci, Trocca and Volpe 2000). Besides differences in the computer vision and pattern recognition techniques implemented, the prototypes use a distributed design whereby each module (see next section) of the system can be executed on one or more machines with different operating systems, to provide distributed processing capability and effective cross-platform integration.

2. FRAMEWORK DESIGN

The basic requirement of such interaction, at the fundamental level, is an action–reaction model, which maps a specific movement to an audio event. Figure 1 summarises the basic framework of the trans-domain mapping of one creative domain onto another.

The acquisition module consists of a data capture system, which interfaces the framework to the real-world environment. In this paper, a number of acquisition module implementations are presented, including digital video and physical sensors. There is a wide range of sensor and actuator technologies, which could be used to provide input to the framework, including imaging, ultrasound, pressure, motion, temperature, vibration, acceleration, position, range sensor and many more.

Organised Sound 7(2): 191–200 © 2002 Cambridge University Press. Printed in the United Kingdom. DOI:10.1017/S1355771802002091

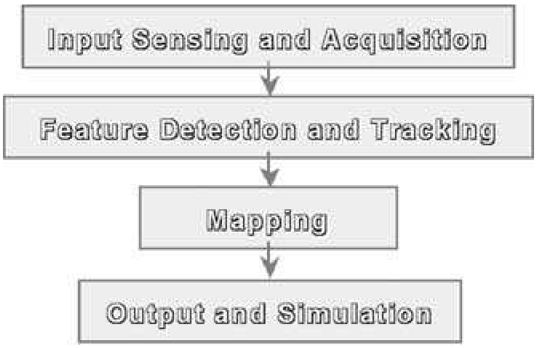

Figure 1. Main TDM modules.

The feature detection and tracking module would contain algorithms to locate and follow certain predefined features in the input data, for example colour, shape and motion. The mapping module is made up of an extensible set of functions which reacts to the detected features by generating an *appropriate* output. The output and simulation module is responsible for multimedia events creation. For example, playing an audio file or displaying an animation.

Besides being an interface for audio and visual control, motion and gesture analysis is also the main focus for the understanding of the motor processes of human performers to encode and transmit expressions and emotions, for example conductor gesture (Camurri 1998, Marrin-Nakra 2000, Johannsen 2001, Maruyama 2001).

3. MAPPING

In general, mapping strategies can be divided into the following categories (Wanderley 2001a):

(1) one-to-one mapping,
(2) one-to-many mapping,
(3) many-to-one mapping, and
(4) many-to-many mapping (a combination of the above).

For the implementations as described in the paper, the choice of mapping strategy depends on the context of the application and the number of input and output/control parameters. Each model has its purpose and usage. For example, a piano keyboard interface to pitch uses a one-to-one mapping, since each key represents a specific pitch.

This paper describes several implementations of the TDM framework in different contexts and discusses mapping strategies for:

- an interactive stage performance where mapping is taken as a role and design issue of the composition (see section 4.2), and
- a flexible multimedia interface to control audio events using facial expression. In this case, mapping is considered as the design (and configuration) of the instrument (see section 4.5.1).

Reviews and surveys of the development of various sensing technologies and mapping models can be found in Hunt (1999), Hunt and Kirk (1999, 2000), Marrin-Nakra (2000a), Wanderley and Battier (2000) and Wanderley (2001, 2001a). Discussions on complex mapping strategies for expert instrument interaction can be found in Hunt, Wanderley and Kirk (2000).

4. TDM SYSTEMS

This section presents several system designs and implementations using the proposed framework.

4.1. Music via Motion (MvM)

The first implementation using the TDM framework was MvM (Music via Motion). MvM uses input from a video camera to detect and track visual changes in real time, and makes use of the detected visual activities (see figure 2) to generate interesting and *relevant* musical events using an extensible set of predefined mapping functions (Ng 2000).

The prototype is currently equipped with motion- and colour-sensitive modules exploring computer vision techniques (see figure 3). Motion detection and tracking sub-modules include standard frame-differencing and background subtraction. Figure 3 shows the frame-differencing module with an overlaying triangle containing the area of active visual changes. The three vertices of the triangle can be used by one of the mapping functions (discussed later) to define the pitch and volume of an audio output event.

Pixel-wise colour segmentation in RGB space is straightforward and surprisingly effective, but the performance is sensitive to lighting condition. Colour detection is an active research area and many methods have been proposed to model colour spaces, including RGB (Jebara, Russell and Pentland 1998), normalised RGB (Crowley and Bedrune 1994), YCrCb (Chai and Ngan 1998) and others. To provide better colour segmentation, work in progress includes transforming the colour representation (Raja, McKenna and Gong 1998, Drew, Wei and Li 1998), such as HSV (Sobottka and Pitas 1996), to minimise the variance of a colour cluster with illumination normalisation.

Basic mapping functions include a distance-to-MIDI-events mapping, with many configurable parameters, such as scale-type, pitch range and others. Musical mapping can be enhanced with a database of composed musical phrases and several mapping layers can be overlaid in order to produce multi-layered effects. MvM also offers user configurable *active regions* where detected visual activities in certain areas can be mapped onto different MIDI channels.

There has been increasing interest in MvM collaborations from a variety of disciplines. In addition to the

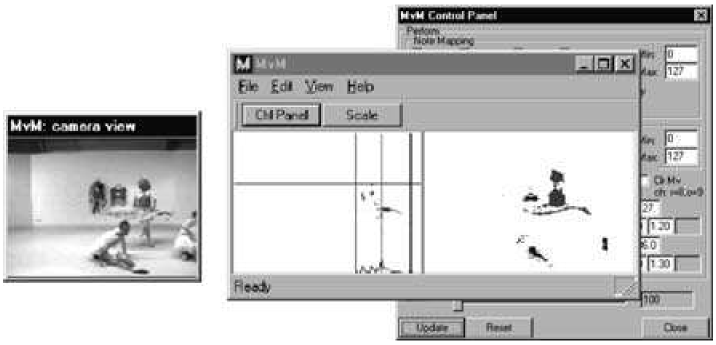

Figure 2. MvM prototype. The left-most camera view window shows the live video input. The main MvM window displays the frame-differencing motion tracker (left), and the colour detection module (right) segments recognised parts of the costumes (red and cyan, in this case).

Figure 3. MvM frame-differencing module.

original intentions for basic multimedia event triggering, choreographers, dancers, composers and artists have found many creative applications for the prototype. There may also be applications for music therapists, to encourage movement, using this motion-sensitive system to provide interactivity and creative feedback. With MvM, the whole body of the user acts as a musical instrument interface, which determines the tempo, volume and audio generation of the performance.

4.2. Coat of Invisible Notes (CoIN)

With the MvM prototype described above, CoIN is a collaborative project designed to bring together multiple creative domains to build special costumes, music and dance within an interactive audiovisual performance interface simulated by the MvM.

For CoIN performances, MvM is configured to detect and track the colour where visual changes are detected. Detected colours are used to control the choice of musical sound and effects. This feature is fully explored and is particularly apparent in a section of the choreography where the dancers are divided into two groups, wearing different coloured costumes. The contrasting movements and interactions between the two groups create interesting musical dialogues with two different musical sounds. The costume designs feature reversible and modular parts, allowing the dancers to reconfigure and reassemble the costumes to create different visual effects, and at the same time, these transformations are detected and reacted by MvM. Hence the visual changes of the costumes can also be used to control the character of the musical responses.

Figure 4 presents rehearsals with the MvM colour tracking sub-module, to trigger special sound effects with colour, and figures 5 and 6 present snapshots from a public performance.

4.3. Mapping strategy in interactive performance with MvM

With the default configuration of MvM, it was found that simple *one-to-one* mapping generates musical output that can be easily related to the dance, but many dancers with independent movements can create complex output that is difficult to follow. Hence, the choreography for the CoIN project started with simple motion to demonstrate the direct correlation of motion and sound, before gradually introducing more complex movements.

Constant *one-to-one* direct mapping of movement can also be tiresome and uninspiring. For the CoIN performance, a background layer of music was specially composed, to provide an overall form and structure, with

Figure 4. Colour detection module tracks the colour of costumes to trigger special sound effects (left), and two groups of dances generating a two (MIDI) channels musical interlude.

Figure 5. CoIN/MvM public performance.

various timed intervals for MvM to perform its solo *cadenza*. Basic expressive features are being added to the MvM prototype. These include an *accent detector* module which keeps a history of the region size of the detected visual changes, the directions and speed of the motion, and their means. Sudden changes in these parameters are used to control factors in audio generation. For example, a sudden increase in speed could be mapped to an accented note played. In this case, the mapping is part of the compositional design and the mapping strategy changes from simple to complex with the progression of the performance, integrating with the background music and choreography.

Figure 7 illustrates the multi-layered mapping functions which were *composed/designed* for the performance, exploring a mixture of simple one-to-one direct sound triggering, special effects and pre-composed musical phrases.

4.4. Music via Sensor (MvS)

With the above-mentioned framework, additional or alternative sensing capabilities can be provided by small and non-intrusive physical sensors (e.g. pressure maps, vibration switches and others) installed on the performance environment, for direct triggering of specific musical events and provide additional dimensions of interactivity. This has been realised recently on two public multimedia-installation projects produced by the final-year music technology class of the School of Music, with an array of pressure-maps on the floor and

Figure 6. CoIN/MvM public performance.

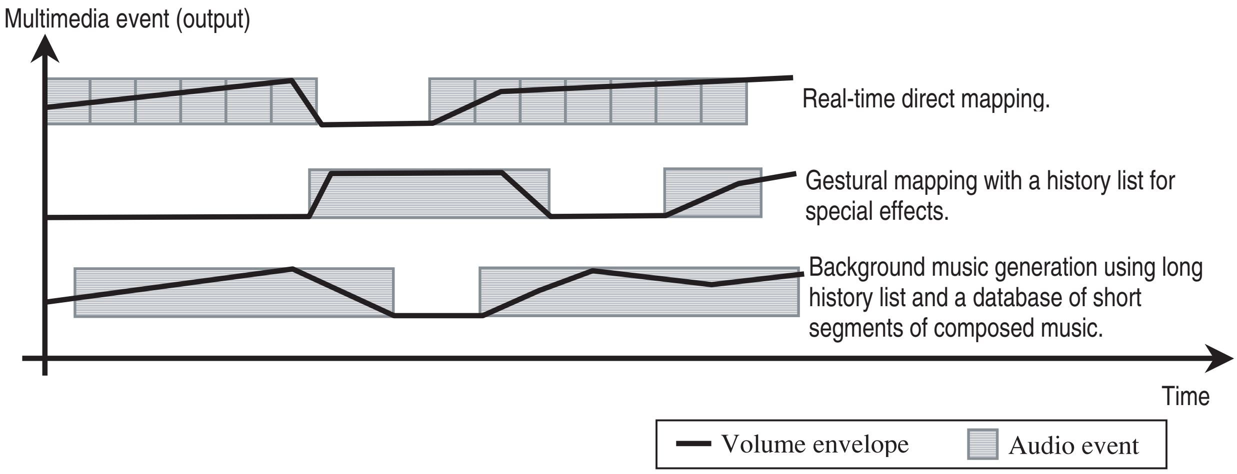

Figure 7. Multi-layered mapping.

wall. The installation expresses a narrative with background music. Special sound effects and short audio segments can be activated by members of the audience who step onto pressure-maps or touch sensors, while pausing to read text or images related to the story.

The experimental set-up involved nineteen sensors (a mixture of pressure maps and a variety of switches) to create an eight-note keyboard (installed on the floor) and eleven sensors distributed around a pre-designed pathway to trigger sound effects. In order to provide a simple and low-cost sensor interface to the computer, a PC keyboard controller is used. Each sensor is wired to send a specific character (as type on a keyboard) and a simple key-mapping program is implemented to intercept the key-press and send an appropriate MIDI event or play a WAV file. The set-up is capable of sensing and reacting to simultaneous triggering on differencing sensors (up to eight), and activates a mixture of MIDI and audio output.

Many simple and low-cost switches and sensors, for example flex sensors and vibration and proximity switches, can be used in such installation.

4.5. Multimedia Mapping Server (MMS)

Starting with the simple framework, it was found that many other researches in visual tracking and sensing,

Figure 8. Live video input (left), and the triangulated tracked face shape (right).

Figure 9. Real-time Face Tracker system with spline curves classifying primary face structures.

and the existing system, could be integrated for exploration via the proposed framework. In order to provide seamless integration, data communication between the main modules has been enhanced using sockets (a communication object for applications to send and receive a packet of data across a network) to enable cross platforms and distributed processing. The mapping module has been re-implemented to include a multimedia mapping server, which waits for input data via a stream socket connection on a specific port, and processes the data using the original mapping module.

4.5.1. Interactive Music Head

The Interactive Music Head collaborative projects integrates the MMS with a real-time face (and expression) tracker from an ongoing research project, which aims to create a synthetic *talking head* intended for mediating interaction between humans and machines (Devin and Hogg 2001). Figure 8 shows the triangulated facial expression, and figure 9 illustrates the real-time face tracker system.

The current prototype provides tracking for positions, distances and dynamical changes for the following basic measurements. With the expense of computational complexity and processing time, the shape of these features can also be analysed with pattern recognition techniques to provide additional control parameters and modes of interactions.

(1) Mouth:

- distance between the lower and upper lips,
- width of the mouth, and
- absolute and relative position with respect to another feature;

(2) Eyebrows:

- their absolute and relative position with respect to another feature; and

(3) Eyes:

- absolute and relative position with respect to another feature, and
- height.

With the MIDI output, parameter controls include the usual pitch, velocity (volume) and channel, with their ranges and scales, transpositions and note filter (see figure 11) to provide user-controllable scale-types.

Figure 10 illustrates MMS displaying data tracked by the real-time face tracker system and figure 11 shows part of the graphical user interface (GUI) design. With the GUI, changes of each facial feature (e.g. left-eye-brown, right-eye, mouth opening or location or width)

Figure 10. MMS face reconstructions.

can be mapped onto one or more MIDI channels (or other multimedia events) with independent configurations. In this case, the mapping strategy is implemented as an integrated part of the system which is intended to be flexible and dynamic. Each input feature can be switched on or off and ranked (by weighting) to provide different levels of control and influence. Besides the usual one-to-one mappings for individual event control (for testing and configuration), combinations of input features can also be mapped by

$$\underbrace{\begin{pmatrix}\Delta b_L & \Delta e_L & \cdots & m_V\end{pmatrix}}_{\text{input features}}\underbrace{\begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1j} \\ w_{21} & w_{22} & x_{23} & \\ \vdots & & \ddots & \vdots \\ w_{i1} & \cdots & & w_{ij}\end{pmatrix}}_{\text{mapping coefficients}} = \underbrace{\begin{pmatrix}v_L & v_R & p & \cdots & d\end{pmatrix}}_{\text{output parameters}}$$

where Δb_L is the left eyebrow change (relative vertical position), Δe_L is the left eye change (relative vertical height), m_v is the mouth opening distance (absolute vertical height), w_{ij} are the mapping coefficients, v_L is the left volume, v_R is the right volume, p is the pitch, and d is the duration.

The weighting of each mapping coefficient, which influences the degree of control to certain output parameter(s), is user configurable and updateable dynamically. These coefficients could also be substituted with nonlinear mapping functions, including statistical analysis (Das, Howard and Smith 1999) and modelling to further enhance complexity and abstractions.

Experiments on various different mapping methodologies, using the detected facial expression to influence tonality and note-set, represent work in hand.

5. DISCUSSION

Besides multiple camera input, other sensors and imaging technologies, such as thermal and range imaging, could be integrated into the framework. Future plans also include behaviour modelling (Johnson, Galata and Hogg 1998), and other motion, gestural and expression trackers (Camurri 1998, Camurri *et al.* 2000).

The mapping coefficients matrix, as discussed in section 4.5.1, could be *learned* by training data with known intended motions and outcomes, analysing normal human gestures and expressions in order to simulate natural machine interactions using a reverse engineering approach. For example, we are currently capturing dance motions performed under a set of predetermined expressions (e.g. happy, sad) to analyse the physical differences (acceleration, spatialisation, trajectory, shape, joint angles and so on) encoded in the motion by human dancers in order to convey the expressions.

In addition to the video and sensor tracking of human motion for creative mapping, the data could be used to automatically generate statistical models of typical trajectories and motions (Johnson 1998). With such models, realistic behaviours can be generated and applied to control virtual performers simulation (Volino and Magnenat-Thalmann 1999, Badler, Bindiganavale, Rourne, Allbeck, Shi and Palmer 1999), which could interact with human performer/user.

The stage can also be augmented visually (by means of video projection, large display or other technologies) with computer graphics, which could be influenced and animated by the mapping module, and more interestingly, by using a 3D model of a real-environment to transform the stage. However, current VR display technology is relatively limited and expensive, in comparison to a theatre or cinema, in terms of the size of the audience.

Figure 12 shows an example 3D model (in VRML format) captured from real environment (Ng, Sequeira, Butterfield, Hogg and Goņalves 1998, Sequeira, Ng, Wolfart, Goņalves and Hogg 1999). Since the whole scene is represented by graphical primitives (i.e. triangles), the surfaces can be modified and edited. Operations such as transformation, translation, and many other functions are possible. Physically demanding or impossible scenarios can be virtually created, projected onto the stage and dynamically altered by the mapping module depending on the sensor input.

Integrating gesture control and 3D augmented environment simulation could enable real-time control of the

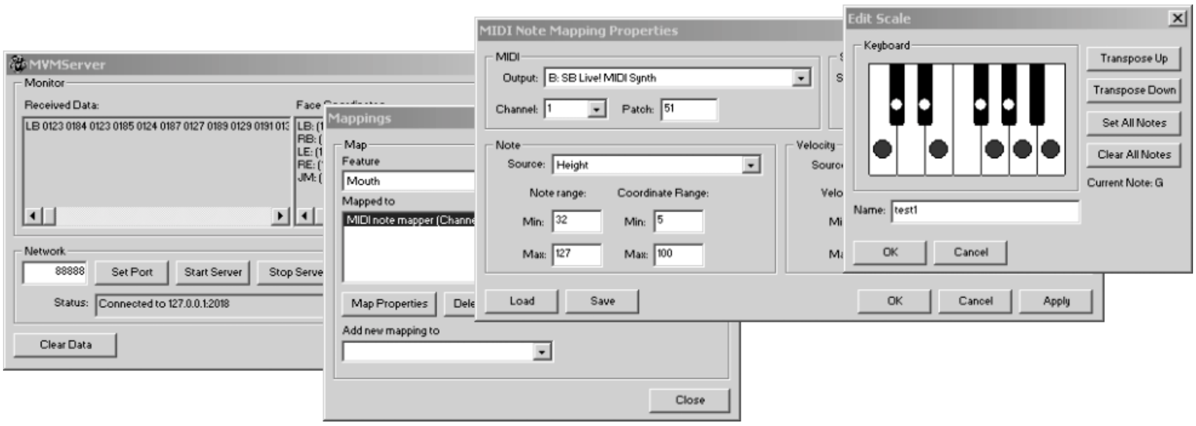

Figure 11. MMS graphical user interface.

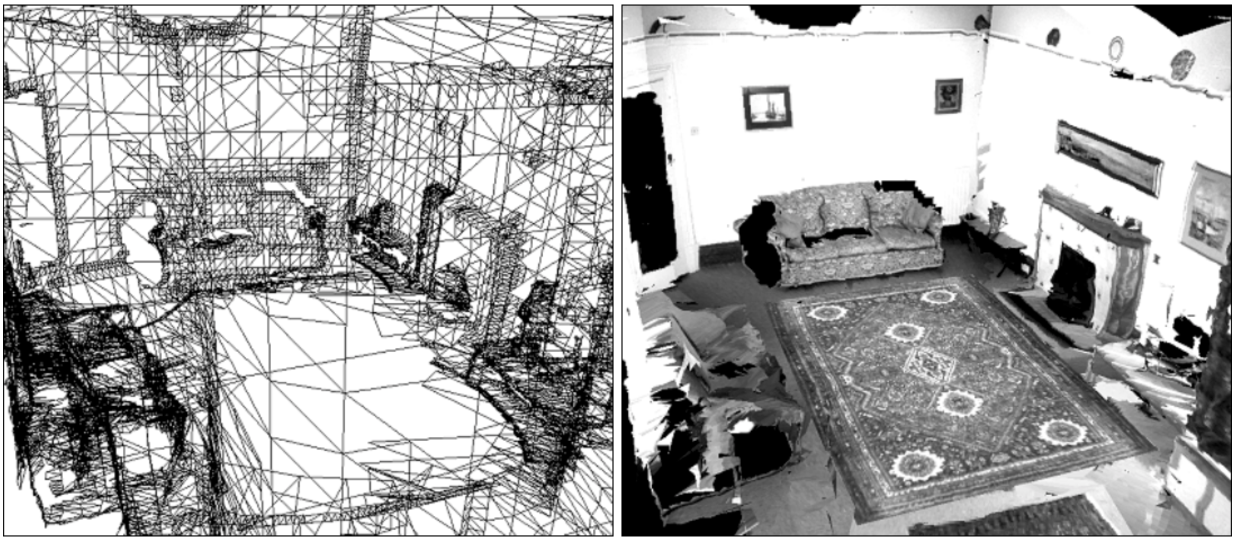

Figure 12. 3D wire-frame model generated from 3D data captured by a laser range finder (left), and for photo realistic visualisation, the model is textured using digital images captured at the real-scene (right).

stage design by the performer, offering another layer of mapping to communicate in the visual domain. Besides changes of scene (backdrop), changes of viewpoint can be easily rendered in a virtual environment to virtually transport the audiences and performers without physical actions. Further discussion can be found in Ng, Sequeira, Bovisio, Johnson, Cooper, Gonçalves and Hogg (2002).

An idealised implementation of TDM would explore human senses and beyond, for a full immersive experience, involving not only sight, sound, smell, touch and taste, but also connections with the electronic signals of the human nervous system (Stelarc 1997, Warwick 1999, Mitchell, Bishop, Keating and Dautenhahn 2000, Scorcioni, Nasuto, Krichmar and Ascoli 2000, Wenn, Mitchell and Gabb 2001); transporting the environment into the realm of imagination, allowing artistic creation and communication by thought (Wolpaw and McFarland 1994, Babiloni, Cincotti, Lazzarini, Millán, Mouriño, Varsta, Heikkonen, Bianchi and Marciani 2000), merging real and artificial entities and environments, and diminishing physical distance with networking, telecommunication and telepresent technologies (Zamorano, Jiang, Grosky, Kadi and Diaz 1995, Paulos 2001).

6. CONCLUSION

This paper has presented a framework to explore the *mapping* of one creative domain onto another using computer vision and sensor technologies. Technical details and set-ups have been discussed, and a number of public performances, using MvM to integrate music, dance and costume design, and interactive-multimedia installation projects, exploring TDM with physical sensors, were reported.

The paper has described different mapping strategies under different application contexts. Direct or one-to-one mapping can be uninspiring, however, and under certain circumstances this could be an important mapping model. For example, virtual instrument interfaces, which allow flexible reconfiguration in order to accommodate any physical limitation (Anderson 1999, Taylor and Bishop 1999), require accurate and reproducible control (e.g. pitch); however, other situations require alternative solutions. In order to offer a different degree of abstraction, the mapping interface needs to be dynamic and flexible.

With the advancement of networking and communications technologies, it is hoped that distributed realtime system, telepresence and remote interactive technologies will continue to evolve, and systems like MvM will continue to explore the integration of arts and science to offer creative human–computer interactions for artistic performances, extending and complementing human sensory capabilities, perceptions, communications and understandings.

ACKNOWLEDGEMENTS

The author would like to thank the RESOLV consortium for the 3D model, and Jon Scott, Vincent E. Devin, Aphrodite Galata and David Hogg for the Interactive Music Head collaboration involving the real-time face tracker system. The CoIN project was carried out with financial assistance from Yorkshire Arts. Thanks to the CoIN project team, Eddie Copp, Claire Nicholson, Sita Popat and dancers from the Bretton Hall College for their collaboration and support. Multimedia installation projects were carried out by the final year Music Technology class of the School of Music. Thanks to ICSRiM, School of Computing, School of Music and School of Electronic and Electrical Engineering, University of Leeds, for collaboration, equipment, and performance space. Thanks to David Cooper and Bee Ong for proof-reading this paper. Thanks to the anonymous reviewers and the editors for their helpful comments and suggestions.

REFERENCES

Anderson, T. 1999. Using music performance software with flexible control interfaces for live performance by severely disabled musicians. In *Proc. of the 25th Euromicro Conf. (EUROMICRO '99)*, Italy.

Babiloni, F., Cincotti, F., Lazzarini, L., Millán, J. del R., Mouriño, J., Varsta, M., Heikkonen, J., Bianchi, L., and Marciani, M. G. 2000. Linear classification of low-resolution EEG patterns produced by imagined hand movements. *IEEE Trans. on Rehabilitation Engineering*, Special Issue on Brain-Computer Interfaces 8: 186–8.

Badler, N. I., Bindiganavale, R., Rourne, J., Allbeck, J., Shi, J., and Palmer, M. 1999. Real time virtual humans. In *Proc. of the 4th Int. Conf. on Digital Media Futures*, National Museum of Photography, Film & Television, UK.

BigEye. A real-time video to MIDI Macintosh software. http://www.steim.nl/bigeye.html

Camurri, A. (ed.) 1998. Kansei: the technology of emotion. In *Proc. of the AIMI Int. Workshop*.

Camurri, A., Hashimoto, S., Ricchetti, M., Ricci, A., Suzuki, K., Trocca, R., and Volpe, G. 2000. EyeWeb: Toward gesture and affect recognition in interactive dance and music systems. *Computer Music Journal* **24**(1): 57–69.

Camurri, A., Coletta, P., Peri, M., Ricchetti, M., Ricci, A., Trocca, R., and Volpe, G. 2000. A real-time platform for interactive performance. In *Proc. of the Int. Computer Music Conf. (ICMC2000)*, Berlin.

Chai, D., and Ngan, K. N. 1998. Locating facial region of a head-and-shoulders color image. In *Proc. of the Third Int. Conf. on Automatic Face and Gesture Recognition*, pp. 124–9.

Crowley, J. L., and Bedrune, J. M. 1994. Integration and control of reactive visual processes. In *Proc. of the Third European Conf. on Computer Vision* **2**: 47–58.

Das, M., Howard, D. M., and Smith, S. L. 1999. Motion curves in music: the statistical analysis of MIDI data. In *Proc. of the 25th Euromicro Conf. (EUROMICRO '99)*, Italy.

Devin, V. E., and Hogg, D. C. 2001. Reactive memories: an interactive talking-head. *Research Report Series*, School of Computing, University of Leeds, Report 2001.09.

Drew, M. S., Wei, J., and Li, Z-N. 1998. Illumination-invariant color object recognition via compressed chromaticity histograms of normalized images. *Proc. of the Sixth Int. Conf. on Computer Vision*, pp. 533–40. Narosa Publishing House.

Fléty, E. 2002. AtoMIC Pro: a multiple sensor acquisition device. *Proc. of the Int. Conf. of New Interfaces for Musical Expression*, pp. 96–101. MediaLabEurope, Dublin, Ireland.

Hunt, A. 1999. *Radical User Interfaces for Realtime Musical Control*. D.Phil. thesis, University of York, UK.

Hunt, A., and Kirk, R. 1999. Radical user interfaces for real-time control. In *Proc. of the 25th Euromicro Conf. (EUROMICRO '99)*, Italy.

Hunt, A., and Kirk, R. 2000. Mapping strategies for musical performance. In M. Wanderley and M. Battier (eds.) *Trends in Gestural Control of Music*. Ircam, Centre Pompidou.

Hunt, A., Wanderley, M., and Kirk, R. 2000. Towards a model for instrumental mapping in expert musical interaction. *Proc. of the Int. Computer Music Conf. (ICMC2000)*, pp. 209–12.

ISIDM. ICMA/EMF Working Group on Interactive Systems and Instrument Design in Music (ISIDM). http://www.igmusic.org

Jebara, T. S., Russell, K., and Pentland, A. 1998. Mixtures of eigenfeatures for real-time structure from texture. *Proc. of the Sixth IEEE Int. Conf. on Computer Vision*, pp. 128–35.

Johannsen, G. 2001. Human supervision and control in engineering and music. In *Proc. of the Int. Workshop on Human Supervision and Control in Engineering and Music*, Germany.

Johnson, N. 1998. *Learning Object Behaviour Models*. Ph.D. thesis, School of Computer Studies, The University of Leeds, UK.

Johnson, N., Galata, A., and Hogg, D. 1998. The acquisition and use of interaction behaviour models. *Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp. 866–71.

Livingston, H. 2000. Paradigms for the new string instrument:

digital and materials technology. *Organised Sound* **5**(3): 135–47.

Marrin-Nakra, T. 2000. Searching for meaning in gestural data: interpretive feature extraction and signal processing for affective and expressive content. In M. Wanderley and M. Battier (eds.) *Trends in Gestural Control of Music*. Ircam, Centre Pompidou.

Marrin-Nakra, T. 2000a. *Inside the Conductor's Jacket: Analysis, Interpretation and Musical Synthesis of Expressive Gesture*. Ph.D. thesis. MIT Media Lab.

Maruyama, S. 2001. How is a symphony started by the orchestra? An empirical study for interpersonal coordination between the professional orchestral conductor and player. In *Proc. of the Int. Workshop on Human Supervision and Control in Engineering and Music*, Germany.

Mitchell, R. J., Bishop, J. M., Keating, D. A., and Dautenhahn, K. 2000. Cybernetic approaches to artificial life. *Kunstliche Intelligenz, Special Edition on Artificial Life*, **1**: 5–11. ISSN 0933–1875.

Mulder, A. 1995. The I-Cube system: moving towards sensor technology for artists. In *Proc. of the Sixth Symposium on Electronic Arts (ISEA 95)*. Montréal, Canada.

Nagashima, Y. 2001. Composition of 'Visional Legend'. In *Proc. of the Int. Workshop on Human Supervision and Control in Engineering and Music*, Germany.

Ng, K. C., Sequeira, V., Butterfield, S., Hogg, D. C., and Gonçalves, J. G. M. 1998. An integrated multi-sensory system for photo-realistic 3d scene reconstruction. *Proc. of ISPRS Int. Symposium on Real-Time Imaging and Dynamic Analysis*, pp. 356–63, Hakodate, Japan.

Ng, K. C. 2000. Music via Motion. In *Proc. of XIII CIM 2000 – Colloquium on Musical Informatics*, Italy.

Ng, K. C., Sequeira, V., Bovisio, E., Johnson, N., Cooper, D., Gonçalves, J.G.M., and Hogg, D. C. 2002. Playing on a holo-stage: towards the interaction between real and virtual performers. In Colin Beardon and Lone Malmborg (eds.) *Digital Creativity: A Reader*. Lisse: Swets & Zeitlinger.

NIME. 2002. In *Proc. of the Int. Conf. of New Interfaces for Musical Expression*. MediaLabEurope, Dublin, Ireland.

Paradiso, J. A., Hsiao, K-Y., Strickon, J., and Rice, P. 2000. New sensor and music systems for large interactive surfaces. *Proc. of the Int. Computer Music Conf. (ICMC 2000)*, pp. 277–80. Berlin, Germany.

Paulos, E. J. 2001. *Personal Tele-Embodiment*. Ph.D. dissertation, University of California at Berkeley, USA.

Raja, Y., McKenna, S., and Gong, S. 1998. Segmentation and tracking using colour mixture models. *Proc. of the Asian Conf. on Computer Vision (ACCV)*, Hong Kong. *Lecture Notes in Computer Science 1351* **I**: 607–14.

Rokeby, D. 1997. The construction of experience: interface as content. In Clark Dodsworth Jr. (ed.) *Digital Illusion: Entertaining the Future with High Technology*. ACM Press. ISBN 0–201–84780–9.

Schoner, B., Cooper, C., and Gershenfeld, N. 2000. Cluster-weighted sampling for synthesis and cross-synthesis of violin family instruments. *Proc. of the Int. Computer Music Conf. (ICMC 2000)*, pp. 376–9. Berlin, Germany.

Scorcioni, R., Nasuto, S. J., Krichmar, J. L., and Ascoli, G. A. 2000. Algorithmic extraction of morphological statistics from electronics archives of neuroanatomy. *Society of Neuroscience Conference*, St. Louis, USA.

Sequeira, V., Ng, K. C., Wolfart, E., Gonçalves, J. G. M., and Hogg, D. C. 1999. Automated reconstruction of 3d models from real environment. *ISPRS Journal of Photogrammetry and Remote Sensing* **54**(1): 1–22. Elsevier, ISSN 0924–2716.

Siegel, W., and Jacobsen, J. 1998. The challenges of interactive dance, an overview and case study. *Computer Music Journal* **22**(4): 29–43.

Siegel, W., and Jacobsen, J. 1999. Composing for the digital dance interface. *Proc. of the Int. Computer Music Conf. (ICMC1999)*, pp. 276–7.

Siegel, W. 1999. Two compositions for interactive dance. *Proc. of the Int. Computer Music Conf. (ICMC99)*, pp. 56–9. Beijing, China.

Sobottka, K., and Pitas, I. 1996. Face localisation and feature extraction based on shape and colour information. *Proc. of the IEEE Int. Conf. on Image Proceeding*, pp. 483–6.

Stelarc. 1997. Parasite: event for invaded and involuntary body. http://www.stelarc.va.com.au/parasite/index.htm

Taylor, L. M., and Bishop, J. M. 1999. Computer mediated communication use by the deaf. *Proc. of the World Organisation of Systems and Cybernetics*, pp. 185–9. Brunel, UK.

Volino, P., and Magnenat-Thalmann, N. 1999. 3D fashion design and the virtual catwalk. In *Proc. of the 4th Int. Conf. on Digital Media Futures*. National Museum of Photography, Film & Television, Bradford, UK.

Wanderley, M. M., and Battier, M. (eds.). 2000. *Trends in Gestural Control of Music*. Ircam, Centre Pompidou.

Wanderley, M. M. 2001. Gestural control of music. In *Proc. of the Int. Workshop on Human Supervision and Control in Engineering and Music*, Germany.

Wanderley, M. M. 2001a. *Performer–Instrument Interaction: Applications to Gestural Control of Sound Synthesis*. Ph.D. thesis, University Paris 6.

Warwick, K. 1999. Linking human neural processes directly with technology – the future. *Proc. of the Int. Conf. on Artificial Neural Nets and Genetic Algorithms*, pp. 8–13. Slovenia.

Wenn, D., Mitchell, R. J., and Gabb, M. 2001. Towards the control room of the future. *Proc. of the IEE Conf. on People in Control (PIC 2001)*, pp. 79–85.

Winkler, T. 1997. Creating interactive dance with the very nervous system. In *Proc. of the 1997 Connecticut College Symposium on Art and Technology*. New London, CN: Connecticut College.

Winkler, T. 1998. *Composing Interactive Music: Techniques and Ideas using Max*. Cambridge, MA: MIT Press.

Wolpaw, J. R., and McFarland, D. J. 1994. Multichannel EEG-based brain–computer communication. *Electroencephalography and Clinical Neurophysiology* **90**: 444–9.

Zamorano, L., Jiang, Z., Grosky, W., Kadi, A. M., and Diaz, F. G. 1995. Telepresence and virtual reality in computer-assisted neurological surgery: basic theory and a prototype. *Journal of Virtual Reality in Medicine* **1**: 54–9.