Return to Index

Augmented Stages for Installation-Arts and Performance

Kia Ng

Interdisciplinary Centre for Scientific Research in Music
Department of Music and School of Computing
University of Leeds, Leeds LS2 9JT, UK

www.kcng.org/mvm mvm@kcng.org

Abstract

This paper presents a multi-disciplinary project, which aims to create an intuitive and non-intrusive interactive audio-visual performance interface that offers the users or performers real-time control of multimedia events using their physical movements. From a straightforward definition of a trans-domain mapping (TDM) framework, this paper reports three implementations and 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. Recent public installations and performances using these systems are also reported. In addition to considerations on interactive technologies using motion and colour tracking, plausible future exploration on stage augmentation with virtual and augmented reality are proposed and discussed.

Keywords: installation, multimedia, vision, sensor, tracking, motion

1 Introduction

Physical movement, gesture and expression play an important role of stage performances, irrespective of the mode of human communications; verbal or non-verbal, or the language used. With the advancements of electronic and computing technology, there has been increasing interests in new musical instrument design to augment traditional instruments (Schoner, Cooper and Gershenfeld 2000) with new capabilities, for examples, triggering digital sound and visual output (Paradiso *et al.* 2000), as well as new interface designs to provide better ergonomics considerations, and/or offer simpler instrumental control to a wider users. With such systems, the mode 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.

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.

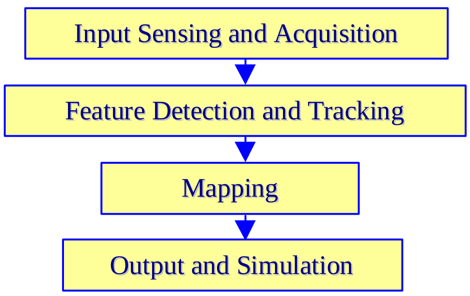

Figure 1: Main TDM modules.

The acquisition module consists of 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. The feature detection and tracking module would contain algorithms to locate and follow certain predefined features in the input data, such as 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.

2 TDM Systems

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

2.1 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).

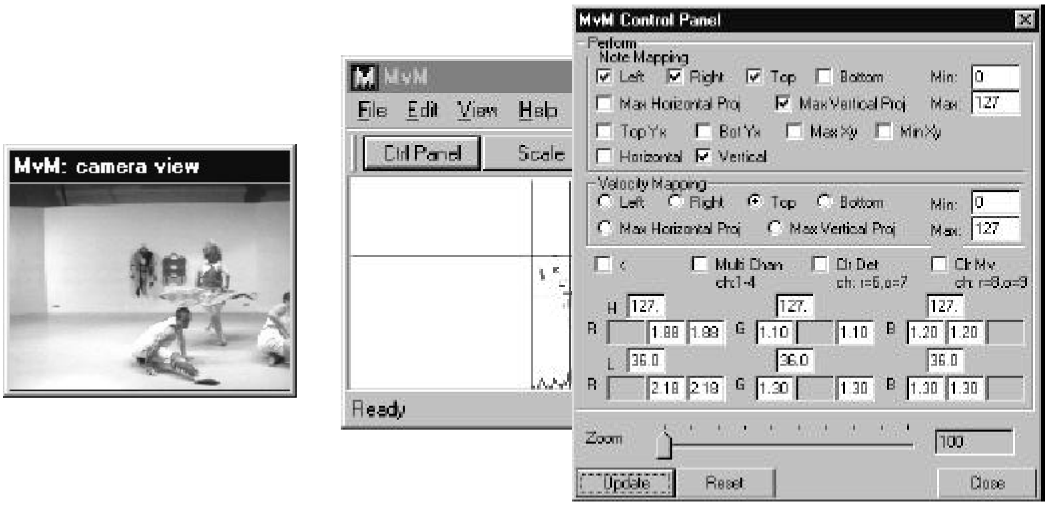

Figure 2: MvM prototype. The left most camera view window shows the live video input. The main MvM window displays the frame-differencing motion tracker (left), and the colour detection module (right), segments recognised parts of the costumes (red and cyan, in this case).

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 function (discuss 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. 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, to minimise the variance of a colour cluster with illumination normalisation.

Figure 3: MvM frame-differencing module.

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 and polyphonic 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 an increasing interest in MvM collaborations from a variety of disciplines. In addition to original intensions 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.

2.2 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 audio-visual performance interface simulated by the MvM.

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.

For CoIN performances, MvM is configured to detect and track the colour where visual changes were 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 costumes in different colours. The contrasting movements and interactions between the two groups create interesting musical dialogues with two different musical sounds (see Figure 4). The costumes designs feature reversible and modular parts, allowing the dancers to reconfigure and reassemble the costume 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 MvM colour tracking sub-module, to trigger special sound effects with colour, and Figure 5 presents snapshots from a public performance.

Figure 5: Snapshots of a CoIN/MvM public performance.

2.2.1 Interactive Performance with MvM

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

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 various timed intervals for MvM to perform its solo *candenza*. Basic expressive features are being added to the MvM prototype. This includes 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 map to an accented note played.

2.3 Music via Sensor (MvS)

With the above mention 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 environments, 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 final year Music Technology class of the Department of Music, with an array of pressure-maps layout on the floor and wall. The installation tells a story line with background music. Special sound effects and short audio segments can be activated by the audiences/visitors who step onto a pressure-map or touch a sensor, while pause to read text or images related to the story.

The experimental setup involved 19 sensors (a mixture of pressure maps and a variety of switches) to create an eight-note keyboard (installs on the floor) and 11 sensors distributed around a decimated pathway to trigger sound effects. In order to provide simple and low cost sensor interface to the computer, a PC keyboard controller is used. Each sensor was wired to send a specific character (as type on a keyboard) and a simple key-mapping programme was implemented to intercept the key-press and send an appropriate MIDI event or play a WAV file. The setup is capable of sensing and reacting to simultaneous triggering on differencing sensors (up to 8), and activates a mixture of MIDI and audio output.

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

3 Multimedia Mapping Server (MMS)

Starting with the simple framework, it was found that many other researches in visual tracking and sensing, and existing system, could be integrated for TDM exploration. In order to provide seamless integration, data communications between the main modules have been enhanced using socket 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 process the data using the original mapping module.

3.1 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 6 shows the triangulated facial expression, and Figure 7 illustrates the real-time face tracker system.

Experimentations on various different mapping approaches, using the face shape contour, represent the works in hand.

Figure 6: Live video input (left), and the triangulated tracked face shape (right).

Figure 7: Real-time Face Tracker system with spline curves classifying primary face structures.

4 Future Direction and Conclusion

Beside multiple cameras input, other sensors and imaging technologies, such as thermal, infra-red, 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 *et al.* 2000).

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 performer (Volino and Magnenat-Thalmann 1999, Badler *et al.* 1999) simulation, 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 graphic, which could be influenced and animated by the mapping module, and more interestingly, by using 3D model of real-environment to transform the stage. However, current VR display technology is relatively limited, comparing to a theatre or cinema, in term of the number of audiences.

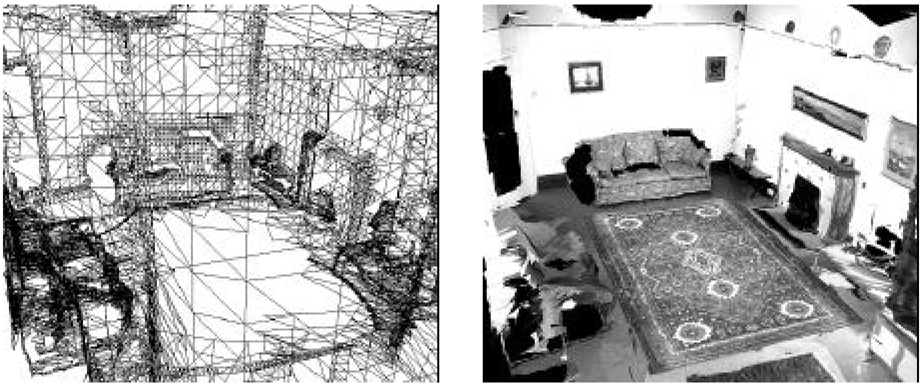

Figure 8: 3D wire-frame model generated from 3D data captured by a laser range finder (left), and for photo realistic visualisation, the model is textured using digital images captured at the real-scene (right).

Figure 8 shows an example 3D model (in VRML format) captured from real environment (Ng *et al.* 1998, Sequeira *et al.* 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.

In short, this paper presents a research framework to explore the *mapping* of one creative domain onto another using computer vision and sensor technologies. Technical details and setup are discussed, and recent performaces, using MvM to integrate music, dance and costume design, and interative-multimedia-installation projects, exploring TDM with physical sensors, are reported. With the advancement in science and technology, it is hope that systems like the MvM will continue to explore the integration of art and science to offer creative human computer interactions for artistic performances (Ng *et al.* 2000, Siegel and Jacobsen 1998, Siegel 1999, Wanderley and Battier 2000).

Acknowledgement

The author would like to thank the RESOLV consortium for the 3D model, and Vincent E. Devin, Aphrodite Galata and David Hogg for the Interactive Music Head collaboration involving the real-time face tracker system. CoIN project was carried out with financial assistance from Yorkshire Arts. Thanks to the CoIN project team, Eddie Copp, Claire Nicholson, 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 Department of Music. Thanks to ICSRiM and Department of Music, University of Leeds, for equipment, and rehearsal and performance space.

References

Badler, N.I., Bindiganavale, R., Rourne, J., Allbeck, J., Shi, J., and Palmer, M. 1999. Real time virtual humans, in *proceedings of the 4th International Conference on Digital Media Futures*, National Museum of Photography, Film & Television, Bradford, UK.

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*, MIT Press, 24(1): 57–69.

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, *Sixth International Conference on Computer Vision*, Narosa Publishing House, pp. 533–540.

Johnson, N. 1998. *Learning object behaviour models*, PhD 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, in *Proceedings IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 866–871.

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, in *proceedings of ISPRS International Symposium on Real-Time Imaging and Dynamic Analysis*, Hakodate, Japan, pp. 356–63.

Ng, K.C. 2000. Music via Motion, in *Proceeding 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. 2000. Playing on a holo-stage: towards the interaction between real and virtual performers, *Digital Creativity*, Swets & Zeitlinger Publishers, 11(2): 109–117.

Paradiso, J. A., Hsiao, K-Y., Strickon, J. and Rice, Peter. 2000. New sensor and music systems for large interactive surfaces, in *proceedings of the International Computer Music Conference (ICMC 2000)*, Berlin, Germany, pp. 277–80.

Raja, Y., McKenna, S. and Gong, S. 1998. Segmentation and tracking using colour mixture models, Asian Conference on Computer Vision (ACCV), Hong Kong, *Lecture Notes in Computer Science 1351*, I: 607–614.

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

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*, Elsevier, ISSN: 0924-2716, 54(1): 1–22.

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. 1999. Two compositions for interactive dance, in *proceedings of the International Computer Music Conference (ICMC99)*, Beijing, China, pp. 56–9.

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

Wanderley, M. and Battier, M. (eds.). 2000. *Trends in gestural control of music*, Ircam - Centre Pompidou, 2000.