EMIC - Compositional experiments and real-time mapping issues in performance

Donna Hewitt

School of Contemporary Arts, University of Western Sydney, Penrith
email: dhewitt@bigpond.net.au

Abstract

This paper discusses the development of a gestural controller interface for contemporary vocal performance and electronic processing called the eMic (extended Mic-stand interface controller). This instrument is a modified microphone stand, custom fitted with an array of sensors and gesture capturing devices aimed at capturing commonly used gestures and movements of vocal performers who use microphones and microphone stands in performance. These common gestures were discussed in an earlier paper prepared for the New Interfaces for Musical Expression Conference 2003 (Hewitt and Stevenson 2003) and it was seen that the gestures form the basis of a well-practiced language and social code for communication between performers and audiences. The microphone itself has become a performance tool of the contemporary vocalist and a means for extending the voice as an instrument. The eMic aims to further facilitate the performer by giving them more flexibility and control over the processing and sound of their voice in a live context.
This paper explores the mapping process, early compositional experiments and the use of the eMic in live performance, identifying the successes and shortcomings of the interface and areas for possible exploration and further development.

1 Introduction

The eMic (extended Mic-stand interface controller) is a gestural controller for contemporary vocal performance with electronic processing. The interface consists of a modified microphone stand fitted with various sensors to capture existing and new physical performance gestures. (Hewitt and Stevenson 2003)

The motivation for developing the eMic was to address some of the problematic technical and aesthetic issues associated with electro-acoustic vocal performance practices. In contemporary music styles such as rock, pop and folk music, vocal performers are often limited in their control over the sound of their voices through the sound reinforcement system. Once the sound enters the microphone, any additional signal processing such as filtering, reverberation, distortion, granulation, delay effects added to the vocal signal are usually carried out by a sound mixer or third party. Often these effects are of an intrinsically musical nature and are closely allied with other vocal production techniques employed by the performer.

The desire for vocalists to harness the available digital signal processing technologies to extend the voice as an instrument, has given rise to a trend in the use of computers in performance to carry out real time digital signal processing. See figure 1.

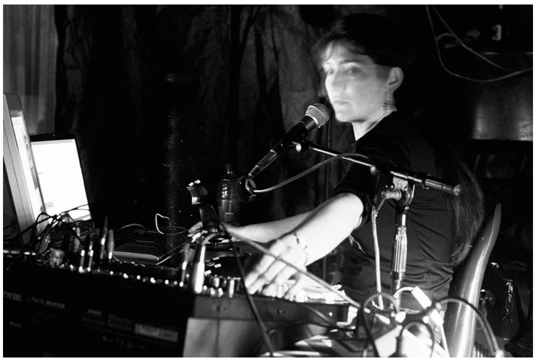

Figure 1. Example of Laptop performance using live vocal input. Donna Hewitt Impermanent Audio 2002.

This practice raises issues concerning the performer's relationship to the audience. The most commonly cited 'deficiency' in laptop performance is that, with the performer seated behind the laptop, there is an inherent lack of gestural communication between performer and audience due to the fact that gesture is so small and often hidden from view. As a result, the performance can have a detached, non-communicative quality (Cascone 2002).

The other perceived limitation of the laptop performances is that the posture of sitting at a computer when trying to vocalise may be physically inhibiting to vocal production.

In summary, the main goals in developing the eMic were to:

1) Increase and improve the control a vocal performer has over the sound of their voice in a sound re-inforcement system

2) Allow for extended vocal technique via electronic processing

3) Improve for audiences the visual/ communicative experience of vocal performances which utilise signal processing

4) Overcome the physical inhibition of vocalising from a sitting position, which occurs in laptop type performances.

1.1 Design Research

Initial studies (Hewitt and Stevenson 2003) identified the most common interactions vocal performers make with the microphone and microphone stand and identified the most effective means of capturing these gestures using available sensing technologies and hardware.

The eMic design aims to support the most common existing gestures and interactions. This was necessary in order to make the instrument as accessible as possible to the large number of vocalists already using microphone and microphone stands. While the eMic does not capture all the gestures and interactions used in microphone performance, decisions were made regarding which ones would be the most useful and ergonomically viable and there is an expectation that new gestures and playing techniques will emerge through use and exploration of the interface.

2 Compositional Approaches & Mapping

2.1 Mapping Definition

The term 'mapping' is generally used to describe the relationship between the performer's input and the associated signal processing parameters, in effect the relationship of the performer's gestures and interactions with an instrument or interface, to the sonic outcomes.

While there is considerable discussion of mapping in existing literature, which has identified various issues and approaches to be considered in creating an effective mapping strategy, every mapping strategy needs to be considered in its unique musical and performative context.

2.2 Software – Choice of Synthesis or Processing Environment

The degree of flexibility of the mapping strategy is largely determined by the choice of software environment. The choice will depend on the technical proficiency of the individual user and/or the desired musical outcome. Composers may be drawn to specific processors available in different software packages that are perceived to have unique characteristics, that is, each implementation of a processor has a different sonic quality and one implementation may be more desirable than another.

On the PC platform, lower level programs such as Miller Puckette's PD provide more data manipulation and more flexible data structure compared to higher level programs such as Audiomulch which may be more user friendly, offer a more intuitive graphical interface and be less time consuming in the creation of patches.

Further, the notion of instrument design and composition in this realm extends to software programming, in that all elements are intrinsically related. This makes the choice of software as critical an issue as hardware design of gesture sensing and in a sense makes computer software skills part of the composer/performers technique.

2.3 Experimental Compositional Approach VS Fixed and Repeatable Approach

Two fundamental approaches to the question of control mapping are those which see mapping as an integral part of the experimental process of composition and those, on the other hand, which identify the requirement for fixed and repeatable mapping of gestural input to system control outcome (Hunt, Wanderley and Kirk 2000).

The desire for a fixed, repeatable approach may exist amongst users with less technical proficiency in software and may therefore be useful in a commercial product, which combines ease of use and repeatability with enough scope and challenge for a performer to become virtuosic. Wessel and Wright (2002) describe this approach as having a 'low entry fee with no ceiling on virtuosity'.

A fixed approach to mapping would see the most common gestures and intentional relationships between the gesture and musical outcome forming the basis for the mapping. This approach, by nature limits flexibility in favour of repeatability and commonality.

Responses from initial demonstrations of the eMic indicate that there is significant interest from vocalists with little or no experience in electronic media, suggesting demand for such an approach.

In an attempt to balance ease of use with flexibility, the eMic mapping embraces elements of both these approaches

2.4 Deterministic Morphological Relationships VS Arbitrary Morphological Relationships

A morphological relationship is the relationship between the physical gesture and the sonic outcome.

A deterministic morphological relationship is one which maintains congruity between the musical intent, the expressive aspects of gesture and the sonic outcome, for example, where delicate stroking of the mic-stand produces intimate or subtle sonic outcomes and where violent movements produce more dramatic, intense outcomes. Davidson (1993,1994,1995), in studies of vocalists and pianists, found that in many contexts the audience relies on physical gesture for much of the information concerning expression and musical intent. This would suggest that deterministic morphological relationships play an important role in both vocal and piano performance. Similarly, Wessell and Wright (2002) contend that in electronic music performance there should be a correspondence between "the size of a control gesture and the acoustic result. Although any gesture can be mapped to any sound, instruments are most satisfying both to performer and the audience when subtle control gestures result in subtle changes to the computers sound and larger more forceful gestures result in more dramatic changes to the computers sound."

Both of these arguments are focused on 'performance as spectacle' or the spectacular aspects of gesture. A major problem with Wessel and Wright's assessment, however, is that it does not address the performer's needs in relation to gesture, which may or may not accord with their subjective (audience focused) judgments. It could be argued that the need for precise control of sonic materials is equally as important as the need for visual stimulation, and that any mapping strategy must offer a high degree of control, alongside visual spectacle.

It could thus be said that *in some contexts* it may be important to strive for a compatibility and a logical relationship between the physical gesture and sonic outcome of that gesture In other creative contexts however, it may be desirable to use more arbitrary, non-correlating mappings which are based on the performer's need to have precise control of materials. Such mappings may bring about unexpected results from the perspective of the audience, or at least obscure the direct correlation between gesture and sonic outcome. It could further be argued that at times it is valid to create tension between what is visible and what is not, as a deliberate performance strategy and this might be an effective way of maintaining audience interest.

2.5 Primary Goals in Mapping and Composing for the eMic

In addressing the goals of the eMic, the mapping strategy so far constitutes a balance between the 'spectacle' of the performance and the performer's need for control over the sonic space. There is a need to have a satisfying communicative relationship from the audience perspective and to create a workable relationship from a performer's perspective, which meets the requirements for satisfactory control of the sound source and allows high-level performance skills to develop. This process of balancing acknowledges both aspects of performance practice as opposed to one alone, with a view to engaging with performance in a more sophisticated fashion.

3 Initial Mapping Experiments

The first composition and mapping strategy for the eMic used a combination of Miller Puckette's PD and Ross Bencina's AudioMulch. All of the audio signal processing occurred within Audiomulch, while PD was used primarily for additional signal conditioning of the MIDI data. The rationale for choosing Audiomulch as the signal processing platform for the audio was its ergonomic interface, the ability to utilise VST plugins along with Audiomulch processors, MIDI capability, automation control, use of prepared material, relative low latency, access to the author Ross Bencina, price and familiarity. Familiarity was important because there was limited preparation time prior to the first performance presentation. Audiomulch lacks extensive signal conditioning capabilities thus PD was used to carry out the MIDI signal processing that was not possible in Audiomulch.

The initial mappings were primarily one to one mappings, that is, one gesture to one parameter, with some additional mappings being one to many, that is, one gesture to numerous parameters. Research by Hunt, Wanderley and Kirk (2000) suggests that mapping strategies that are <u>not</u> one to one can be 'more engaging to users than one-to-one mappings' and they found that these more complex mappings although promising more 'long term potential', 'cannot be learned instantaneously'. The rationale for one-to-one mappings at this stage of the process with the eMic was primarily to make the interface easier to use and learn. Complexity was attained through the use of the Audiomulch's matrix feature to open and close processors along with programmed automation to change the mapping function of the various eMic controls.

Composing for the eMic yielded some surprising results and challenges. Working with heavily processed vocal in live contexts can be challenging for a performer. The voice, unlike other instruments, is the body and we learn to control the muscles of the voice in early childhood. The muscular motor programming of the voice is mediated by aural and other bodily perceptions and there is a very tightly connected feedback system between vocal production and perception. Introducing electronic processing of the voice interferes with this feedback flow and makes control of the voice much more challenging. A common approach in the compositional phase is to record the voice as a sound file, and then experiment with mapping, using the sound file as a substitute for the live vocal. This has the advantage of freeing up the composer from vocalizing while they work on aspects of mapping. A problem arises, however, when one reverts to using live vocal, in that the interference of the processing in the perceptual feedback loop can pose obstacles for the vocal performer.

Assuming that it is not acceptable for the performer to wear headphones (thus hearing no processing) it may be preferable to adopt a different compositional approach, which would integrate the live vocal input into the experimental/ improvisational stages of composing. This would ensure that the vocalist experiences the processed feedback and that the composition accounts for the aural feedback issues. A positive aspect to this limitation is that the body is necessarily re-integrated into the music making process, the body historically having played a “minor role in the creation and performance of electronic music” as identified by Bahn, Hahn and Trueman (2001).

Similarly with the control of the eMic, the musical outcome is likely to benefit from the integration of the body into the compositional process. While the interface is physically independent from the sound source and sound processing engine, the process of mapping and composing are so tightly linked that physical interaction and experimentation with the interface are necessary throughout the composition process, the eMic interface is a “much more gestural or ‘instrumental’ than conventional computer interface devices” (Bahn, Hahn and Trueman 2001) such as a computer keyboard and mouse.

3.1 Audio Control and Signal Network

Figure 2, shows the overall signal flow for the initial performances. The performer is at the centre of the technology both as sound source and controller of sound processing. The eMic controller puts out voltage control messages, which are converted to midi messages via Angelo Fraietta’s Dumb Controller. The signals then either pass through PD unchanged or have some conditioning applied such as smoothing of jitter. PD provided additional midi data control that was not available in Audiomulch, for example, locking off or holding parameter values, see figures 18 and 19. The signal then loops back out through the midi converter and back into Mulch where it controls the various processors. The live vocal signal processing is performed in audio mulch arriving via a small sub-mixer followed by an external USB audio device, the emi2/6.

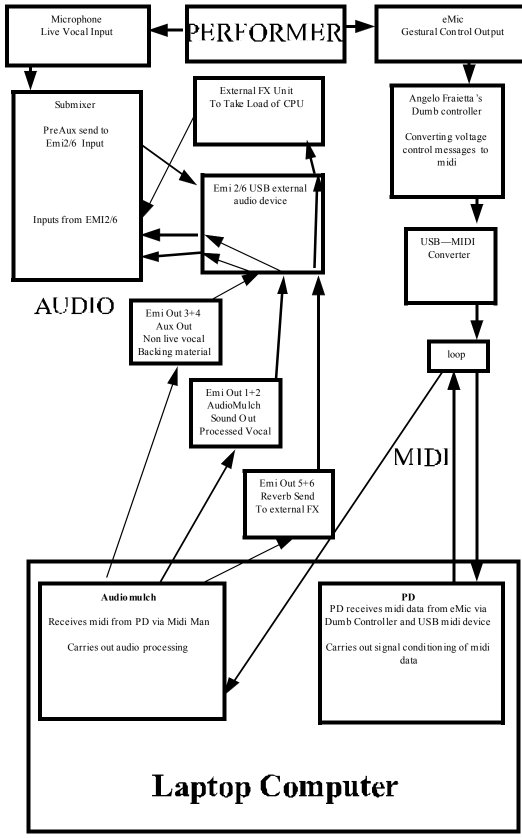

Figure 2. Audio & control signal network.

3.2 Mapping Examples

The following shows some of the mappings that were used.

The right slide controller shown in figure 3, was used to control the amount of signal being processed. It was used to balance between the wet or dry vocal or to also control the amount of processing of either the live vocal or pre-prepared material.

Figure 3. Slide sensor.

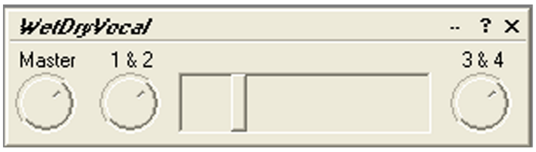

Figure 4. Audiomulch Crossfader mapped to the right slide sensor controlling the amount of signal sent to processors.

The left slide controller was used to control the pitch shift parameter in the GRM Tools Shuffler VST plugin.

The Y axis (left to right) of the joystick shown in Figure 6, was used to control a GRM Tools bandpass sweep (Figure 5.). The joystick X and Y axis were also controlling the GRM tool shuffler fragment and envelope parameters respectively.

Figure 5. Band pass mapped to Y joystick axis.

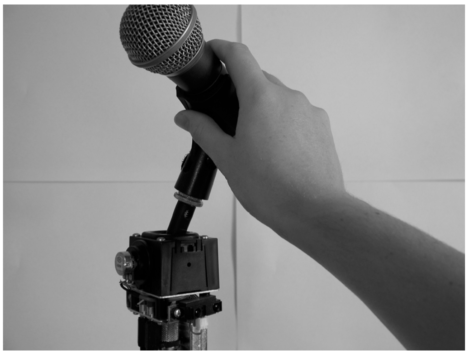

Figure 6. XY Joystick

The foot pressure sensor shown in Figure 7 was used to control the amount of vocal effected by the GRM bandpass.

Figure 7. Foot pressure controller.

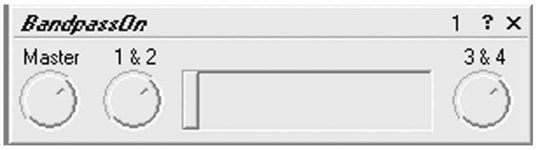

Figure 8. Audiomulch Frosscader mapped to the foot sensor allowing the band pass filter to open.

The front pressure sensor on the microphone clip was used to control the spectral blurring parameter of the *Spectral Monkeyage* VST Plugin. When the sensor is released the freeze parameter becomes engaged.

Figure 9. Grip pressure sensors.

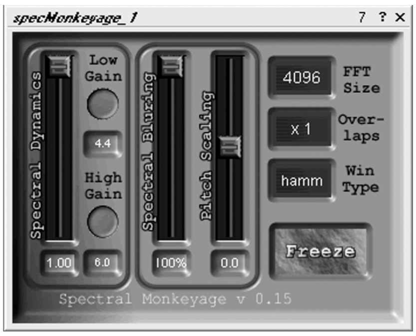

Figure 10. Spectral Monkeyage – Spectral Blurring parameter mapped to front grip sensor.

The rear pressure sensor on the microphone clip was used to control the frequency parameter of the Audiomulch Pulsecomb processor.

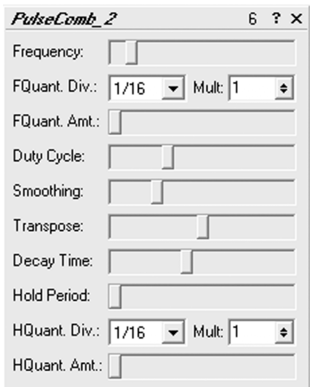

Figure 11. Audiomulch Pulse Comb Frequency mapped to rear grip sensor.

Figure 12, shows the response curve applied to the front grip sensor controlling the frequency parameter. A direct linear relationship in this case was not desirable from a playability point of view.

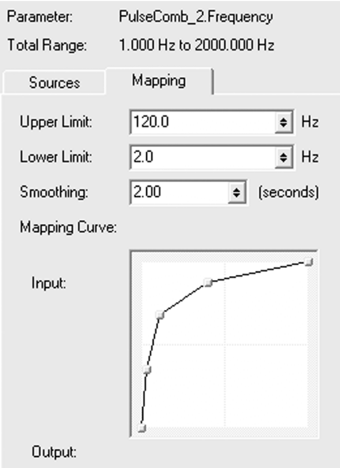

Figure 12. Mapping curve applied to the front grip sensor.

The left and right optical distance sensors on each side of the microphone stand were used to control the Audiomulch Delay parameters. The left hand sensor controlling the send and the right hand controlling the feedback. The non-linear output of the distance sensors along with its' non-tactile nature proved the more challenging in terms of mapping and was not implemented until the second performance, although audience feedback suggests that the mapping employed proved to be a successful and visually satisfying relationship.

Figure 13. Playing distance sensors.

The Y-axis of the tilt sensor (Figure 14) was used to control the velocity parameter of the Audiomulch spatialiser processor shown in (Figure 15). The X-axis of the tilt sensor from an upright position to forward tilting position was used to control the Doppler parameter of the Audiomulch spatialiser (Figure 15). The X-axis of the tilt sensor from an upright position to a backward tilting position (towards the performer) was used to control distortion effects via a VST plugin called Electrofuzz.

Figure 14. Playing tilt sensor.

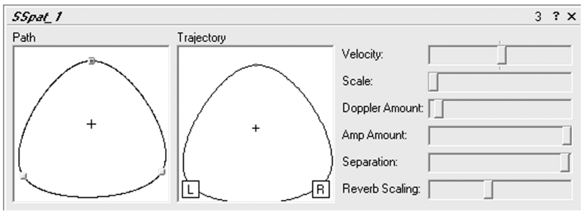

Figure 15. Audiomulch Spatialiser.

The left push button on the joystick encasement (Figure 16) was used to lock off the joystick parameters while the middle and right buttons were used to increment and decrement through the presets of the Audiomulch matrix (Figure 17). The incrementing algorithm was carried out in PD and is shown in Figure 18. The matrix was set up so that moving through the presets would open and close various processors. The transitions between these presets were made smooth by the fade control feature with in the matrix.

Figure 16. Switches on joystick casement.

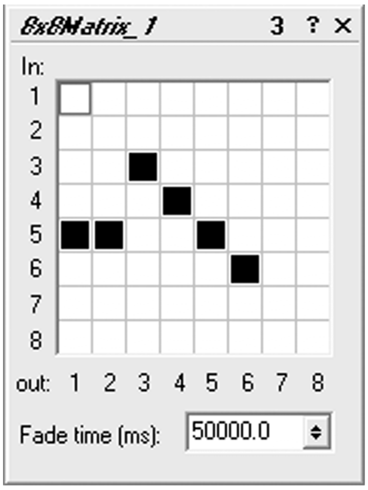

Figure 17. Audiomulch Matrix.

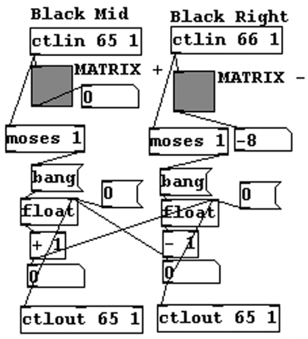

Figure 18. Button increment/decrement PD algorithm.

The front foot switch on the base of the microphone stand (Figure 20) was used to lock off the tilt parameters, while the middle and rear foot switches were used to open and close various processors and files players used for accompaniment. Figure 19, shows the PD algorithm, which enabled the locking off of the tilt parameters and also the smoothing of the tilt sensor data stream.

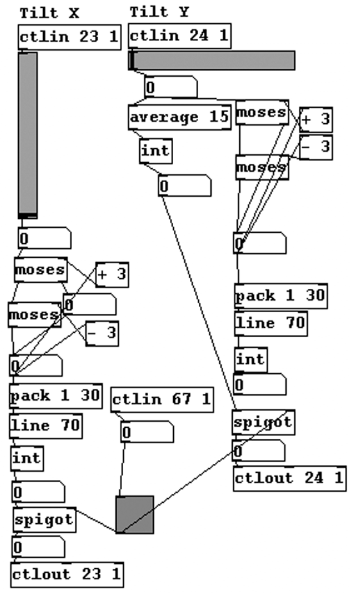

Figure 19. PD patch showing signal conditioning of the tilt sensor's jittery signal and also the switch control allowing the tilt sensor to be locked off.

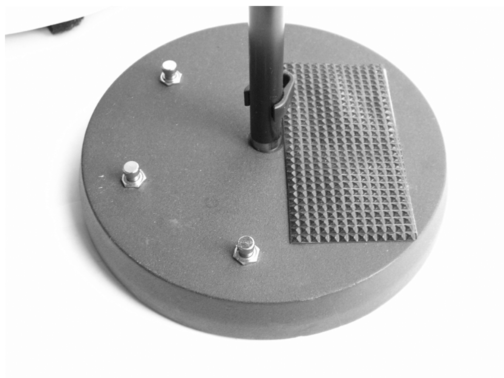

Figure 20. Toggle Switches on the stand base.

4 Initial Performances

At the time this paper was written two performances using the eMic had taken place, the first being '¼ Inch' at the frequency Lab 16th May and the second at the final 'NIME –03' (New Interfaces for Musical Expression Conference) concert, 24th May in Montreal, Canada.

4.1 Feedback and Audience Responses

Audience responses and feedback can be considered useful research for the development of the eMic interface. Consistencies in the feedback aid in the identification of successes and shortcomings in the area of audience reception and some of these have been outlined below.

Initial experiments suggest that audiences respond positively to moments where real-time vocal input takes place. One interpretation of this may be that in the presence of the microphone and microphone stand there is an expectation or pre-conceived notion that 'singing' will occur, and a corresponding disappointment during moments when the performer focuses on sensor control without live vocal input. This may be due to the fact that audiences are conditioned by existing vocal performance practices. It may also be the case that the connection between the cause and effect (gesture and sonic outcome) are most obvious when there is vocalisation, as this is where the audience may get their "clues as to whether there is any essentially 'live' (human produced) activity" Emmerson (1996).

A number of audience members suggested that in the dense sections of the composition it was not possible to determine what the vocalist was doing and the mapping relationships become obscured to the audience. A consideration of pre-existing contexts may help to interpret this reaction. A voice is traditionally perceived as a solo instrument, which should be heard above a background texture. According to Frith (1996), the microphone has "drawn attention to the place of the voice, to the arrangements of sounds behind and around it. The microphone allows the voice to dominate other instruments whatever else is going on." The role of the live voice in the initial compositions designed for eMic performance however, differs to existing contextual models. The vocal can be transformed into thick textural accompaniments and processing can be so extreme that the voice becomes difficult to recognise. In practice, as the voice becomes more processed, it can start to become a texture as opposed to a line above a texture, thus subverting expectations in relation to familiar musical and performative models.

The composer can decide whether the voice is 'solo' or accompanied by treated vocal or electronic elements. Whatever the compositional choices may be the composer needs to consider the communicative experience with the audience. One approach that was adopted in response to the density issue was to have contrasting sections of the work where the texture became quite minimal and the connections between gesture and sonic outcome more transparent.

Another issue in performance that was raised by the audience feedback, was the necessity for the performer to look at the screen. A developed, well-rehearsed practice would ideally require minimal visual feedback via a computer screen. At this stage the desire to stare at the screen would be attributed to performance anxiety, lack of practice and familiarity with the interface, habit from rehearsing without an audience and a need to ensure the software is functioning correctly. In the interim, that is, until the performer is well practiced with the eMic it might be more useful to situate the screen in a more suitable position, perhaps in front of the performer rather than to the side so the performer does not need to turn their head constantly. As the system becomes more stable and reliable the performer is likely to be more confident and hence independent of the visual feedback.

Responses from the audience suggest that some audience members were making imagined correlations between the gestures they were seeing and the sonic outcomes. This generally seemed to happen in the more dense sections of the piece where direct mappings become obscured. This is obviously a complex issue and requires further investigation as to precisely why this effect is so widely experienced. It may be that due to the unfamiliarity of the gestural interface, the audience are so actively engaged in the process of reading gesture, that they read meaning into non-functional gesture.

In relation to the eMic controls, feedback suggested that the more dramatic gestures such as the tilting of the stand were more satisfying for the audience. Whilst useful for the performer as a visual device, such larger gestures take more time to execute and are less efficient in generating control data. In contrast, some of the smaller, less visible gestures such as the pressure and slide sensors provide much finer control over the sonic material in that they are highly efficient in generating control signals. Audience feedback is therefore useful in testing the balancing of needs described earlier in this paper.

5 Conclusion and Future Work

The initial experiments of mapping and composing for the eMic have been undertaken along with two initial performances. The strategy has been to balance the performers needs with a satisfying visual and communicative relationship for the audience.

Audience responses to the initial performances were overall positive and useful in identifying areas of focus for future research.

Future work will continue the development of mapping strategies for the eMic and see input from other vocal performers and composers. The intention is to build a replica prototype that can be circulated for use by other vocalists who are interested in working with electronic processing.

Acknowledgments

Special thanks to Ian Stevenson for building the eMic and providing valuable advice and support for the project. Thanks also to the Australia Council for the Arts, Sounding Out Initiative for funding assistance to build the eMic, ANAT Australian Centre for Art and Technology for assisting the second performance of the eMic at NIME-03. Thanks to the University of Western Sydney, Julian Knowles and Paul Doornbusch.

References

[1] Bahn, C., Hahn, T. & Trueman, D. 2001. "Physicality and Feedback: A Focus on the Body

in the Performance of Electronic Music." *Proceedings of the Multimedia Technology and Applications Conference*. University of California, Irvine.

[2] Bencina, R. "Audiomulch: Interactive Music Studio" (software) http://www.audiomulch.com

[3] Cascone, K. 2002. "Laptop Music: Counterfeiting Aura in the Age of Reproduction". *Parachute* No. 107, 56.

[4] Davidson, J.W. 1993. "Visual perception of performance manner in the movements of solo musicians". *Psychology of Music,* 21, pp. 103-13.

[5] Davidson, J.W. 1994. "Which area of a pianist's body convey information about expressive intention to an audience?" *Journal of Human Movement Studies*, 26, pp. 279-301.

[6] Davidson, J.W. 1995. "What does the visual information contained in music performances offer the observer? Some preliminary thoughts." R. Steinberg (ed.), *The Music Machine: Psychophysiology and Psychopathology of the Sense of Music* Heidelberg, Germany: Springer, pp. 103-15.

[7] Davidson, J. W. 2001 Fall. "The role of the body in the production and perception of solo vocal performance: A case study of Annie Lennox." *Musicae Scientiae*, Vol. V, No. 2, pp. 235-256.

[8] Emmerson, S. 1996 "Local/field: Towards A Typology Of Live Electronic Music" *Journal of Electroacoustic Music* 9: 10-12, Sonic Arts Network, London UK. http://www.somicartsnetwork.org/ARTICLES/ARTICLE1996EMMERSON.html

[9] Fraietta, A. "CV-to Midi–Midi to CV Midi Controller" http://www.users.bigpond.com/angelo_f/midicontroller/MidiController.html

[10] Frith, Simon. 1996. *Performing Rites: on the value of popular music*. Harvard University Press Cambridge Massachesetts. p. 188.

[11] Hewitt, D and Stevenson, I. 2003 "EMIC – Extended Mic -Stand Interface Controller." *Proceedings of the New Interfaces for Musical Expression Conference* McGill University, Faculty of Music. pp. 122-128

[12] Hunt, A., Wanderley, M. & Kirk, R. 2000 "Towards a Model for Instrumental mapping in Expert Musical Interaction. *Proceedings of the International Computer Music Conference*. San Fransisco. International Computer Music Association. pp. 209-212.

[13] *Impermanent Audio*, Live Performance. Frequency Lab. Sydney, Australia. (August 2002). Photo : Mr Snow. http://laudible.net/impaud/archive/

[14] Puckette, M. *Pure Data* (software) http://www.crca.ucsd.edu/~msp/software.html

[15] Rovan, J., Wanderley, M., Dubnov, S., & Depalle, P. 1997, October 3-4. "Instrumental Gestural Mapping Strategies as Expressivity Determinants in Computer Music Performance." Kansei, The Technology of Emotion. *Proceedings of the AIMI International Workshop*, A. Camurri, (ed.) Genoa: Associazione di Informatica Musicale Italiana, pp. 68-73. http://www.ircam.fr/equipes/analyse-synthese/wanderle/Gestes/Externe/Mapp/kansei_final.html

[16] Wessel, D., & Wright, M. Fall 2002. "Problems and Prospects for Intimate Musical Control of Computers". *Computer Music Journal*, 26:3, pp. 11-22. http://www.csl.sony.co.jp/person/poup/research/chi2000wshp/papers/wessel.pdf