Musical Interaction with Hand Posture and Orientation: A Toolbox of Gestural Control Mechanisms

Author One
Affiliation
Affiliation
author@one.com

Author Two
Affiliation
Affiliation
author@two.com

ABSTRACT

This paper develops and presents a toolbox of gestural control mechanisms which are available when the input sensing apparatus is a pair of data gloves fitted with orientation sensors. The toolbox was developed in anticipation of a live music performance in which the mapping from gestural input to audio output was to be developed rapidly in collaboration with the performer. The paper begins with an introduction to the associated literature before introducing range of continuous, discrete and combined control mechanisms enabling a flexible range of mappings to be explored and modified easily. The application of the toolbox is then described with an overview of the developed system and performance set up and closed with a reflection on the system with ideas for future developments.

Keywords

Computer Music, Gestural Control, Data Gloves

1. INTRODUCTION

The use of hand tracking for computer interaction has formed a longstanding focus for research and investigation since the emergence of the earliest motion tracking devices of the 1970s [32]. Since then, a variety of approaches have been developed that focus on the acquisition and processing of hand gestures to bring our interactions with electronic devices closer to our natural interactions with non-computerised objects. The range of motion tracking technology available for this purpose can be broadly separated into two categories: methods relying on external apparatus and methods relying on wearable self-contained sensors. External apparatus is frequently required when tracking is performed using optical camera-based approaches [38, 37, 1]. Whereas self-contained methods generally rely on sensing devices which may be worn, often incorporating bend sensors and/or Inertial Measurement Units (IMUs) [29, 19]. As well as enhancing conventional computer interaction, wearable motion capture technology has been widely adopted as a mechanism to enhance aspects of audio and music interaction. Notable examples include [30, 33, 23] with a range of examples reviewed in [25].

This paper develops toolbox of simple control mechanisms for live music performance which are available when a data glove and Attitude Heading Reference System (AHRS) device are integrated to offer a self-contained wearable device monitoring finger flexion and hand orientation. While data gloves have been combined with a range of inertial sensors in prior works, see for example [29, 14, 10, 34, 17, 9, 28], this paper brings together an overview of the associated mathematics, theory and control options. The remainder of this paper sets out the sensing apparatus adopted herein, followed by an overview of the analysis algorithms which are used to extract discrete gestural features from the continuous flow of sensor data. An overview of potential control options afforded by the resulting data is then provided, combining divergent modes of control already extant in the literature with novel combined suggestions to develop a flexible toolbox of control mechanisms which can be employed for live musical performance. The paper is concluded with an example application and evaluation of these control processes within a live musical performance context.

2. GESTURAL MUSIC INTERACTION

The structure of a gestural musical instrument is frequently depicted with the components shown in figure 1. In response to gestures at the system input, audio is produced at the system output. The sensing apparatus produces input data which is processed to produce control parameters before being translated into audio parameters via a mapping layer. The work presented here is focused on the development of a range of data analysis methods resulting in a flexible toolbox of control mechanisms which may be subsequently mapped to audio processing parameters.

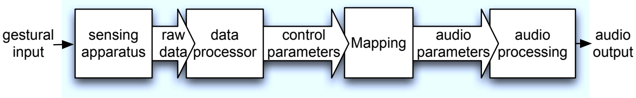

Figure 1: gestural musical instrument structure

3. SENSING APPARATUS

As described by Welch [38] it is the application which determines the appropriate motion tracking approach. A major advantage of wearable self-contained devices, such as data gloves, is that both hands may be tracked in a way that is immune to occlusions with minimal restrictions on the wearer's movements. Since the development of the first data glove in the late 1970s, there have been numerous examples of their utility within musical contexts. For example, the Cyber Composer system [15] has been developed to enable the composition and performance of live music using a vocabulary of hand gestures, which are mapped to construct chord and melody sequences. MusicGlove [14] enables a database of multimedia files to be searched and played back

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
NIME'12, May 21 – 23, 2012, University of Michigan, Ann Arbor.
Copyright remains with the author(s).

using simple hand gestures. Recent examples have seen the mapping of glove-captured gestures for the control of electronic percussion[10] and synthesis [35, 28].

The glove and orientation sensor adopted for this work is shown in Figure 2, comprising two commercially available devices: the 5DT 14 Ultra gloves [3] and the x-io Technologies x-IMU AHRS device [4]. It should be noted that the analysis and control mechanisms presented here are not limited to this hardware, any motion capture apparatus capable of tracking finger flexion and hand orientation may be equally applicable.

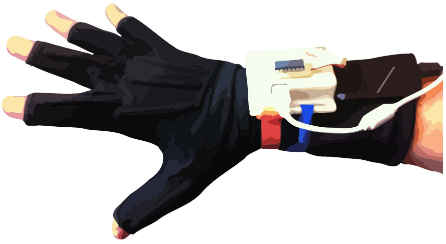

Figure 2: data glove and orientation sensor

3.1 Data Glove

The 14 Ultra device developed by Fifth Dimension Technologies incorporates 14 fibre optic bend sensors. The sensors are positioned at the metacarpophalangeal and proximal interphalangeal joints to measure finger flexion, and between the fingers and thumb to measure abduction/adduction. Frames of 12-bit values for each bend sensor are continuously transmitted at approximately 60Hz via a wired (USB) or wireless (Bluetooth) connection.

3.2 Orientation Device

Attitude Heading Reference System (AHRS) devices are self-contained units able to give an absolute measurement of orientation relative to the Earth coordinate frame. An AHRS device consists of a triple-axis gyroscope, accelerometer and magnetometer and a sensor fusion algorithm to combine the information provided by each sensor into a single estimate of orientation. In recent years, MEMS technology has been rapidly advancing due to the increasing widespread use of inertial sensors within consumer electronics products such as mobile phones, games consoles and other ubiquitous computing devices. This has lead to a new generation of low-cost AHRS products, such as the x-IMU by x-io Technologies. The x-IMU is capable of transmitting instantaneous orientation values simultaneously with the raw sensor data at rates of up to 512Hz via a wired (USB) or wireless (Bluetooth) connection.

3.2.1 Orientation Representation

The x-IMU provides orientation data as a quaternion. Whilst a quaternion is a compact and robust representation of an orientation, meaningful information cannot be directly interpreted. Consequently, quaternions can be converted to alternative representations such as a rotation matrix or an Euler angle sequence.

Euler Angle Representation

An Euler angle sequence represents an orientation as decoupled pitch, roll and yaw angles and is an intuitive representation of orientation. The ZYX Euler angles ϕ, θ and ψ describe the orientation of the sensor frame achieved though the sequential rotations, from alignment with the Earth frame, of ψ around the Earth Z axis, θ around the Earth Y axis, and ϕ around the Earth X axis. While an Euler angles representation is intuitive, it can also be problematic due to the potential of a singularity (or gimbal lock) [36]. The axes of the Euler angle representation are shown in Figure 3

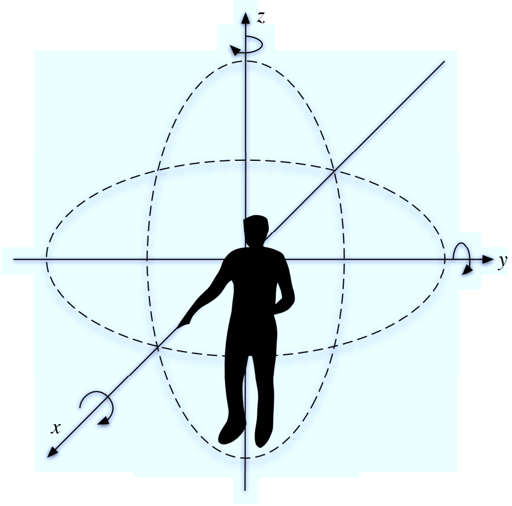

Figure 3: orientation control axes for Euler angles

Rotation Matrix Representation

Orientation data represented as a quaternion can be equivalently represented by a rotation matrix. A rotation matrix is not subject to singularities but still provides orientation data in a way that can be directly interpreted. A rotation matrix is a 3×3 matrix where each column describes a principle axis of the sensor frame as components of a unit vector within an Earth frame axis; the first column describes the sensor X axis as a unit vector in the Earth X axis, the second column describes the sensor Y axis as a unit vector in the Earth Y axis, the third column describes the sensor Z axis as a unit vector in the Earth Z axis.

4. A TOOLBOX OF GESTURAL CONTROL MECHANISMS

The raw data from the data glove and AHRS sensor can be analysed to extract a set of meaningful control parameters which may later be mapped to audio processing parameters. The control parameters can be extracted from a defined set of gestural control mechanisms which may be continuous, discrete or a combination of the two. The mechanisms which have so far been identified and implemented are expounded throughout the remainder of this section.

4.1 Continuous Control

The data-glove and AHRS sensor provide a continuous flow of instantaneous finger flexion and orientation data which can be extracted directly as control parameters, which can later be mapped to audio parameters.

4.1.1 Flexion Control

To ensure parity between users it is important that calibration is performed to scale the raw sensor to a floating-point value in the range 0.0 - 1.0. To produce continuous control parameters by finger flexion, the sensor value for any chosen joint angle may be interpreted directly as a control parameter. Furthermore, sensor readings may be combined by taking the mean average flexion values for multiple sensor values. For example the mean average flexion for all fingers may be mapped to a parameter where an open hand produces the maximum control parameter value and a closed hand produces the minimum.

4.1.2 Orientation Control

In gestural interaction, deictic or directional gestures referring to a point in space rely on the availability of orientation data. With the AHRS orientation datum set such that when the wearer assumes the position shown in Figure 3, the Euler angles ϕ, θ and ψ are all equal to 0°. From this staring point, the Euler angles decrease and increase with clockwise and anticlockwise rotations around each axes in the range -180° to 180°. This range may be easily converted into any other working range. A useful conversion is as follows:

$$\gamma = 0.5 - \frac{\alpha}{180}, \; \alpha = \begin{cases} -180-\theta & \theta < \text{-}90 \\ 180-\theta & \theta > 90 \\ \theta & \text{otherwise} \end{cases} \tag{1}$$

Which for yaw, (Z-axis) rotation of the sensor, produces values as shown in Figure 4, where pointing west equates to 0.0, east to 1.0 and rotation in either direction is symmetrical within this range.

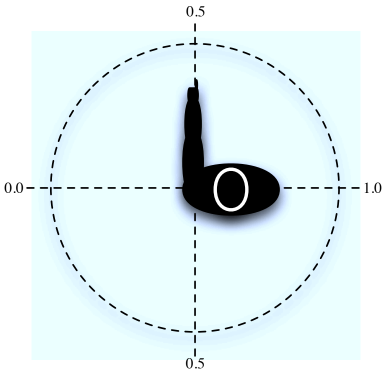

Figure 4: yaw (z-axis) rotation

4.1.3 Positional Displacement

The raw inertial data transmitted by the AHRS device may be processed to provide an indication of relative angular and linear displacement.

Angular

While the Euler angle and rotation matrix representation provide a source of absolute orientation control, relative angular control mechanisms can be derived from the raw gyroscope values. For example, the gyroscope x axis aligns with the user wrist and so measures the angular velocity of the hand as it twists. The angular displacement can then simply be defined by the integrated angular velocity of the wrist. Specifically, the displacement k at time t is defined by equation (2).

$$k_t = k_{t-1} + g\omega_x.\Delta t \tag{2}$$

Estimating Translational Velocity and Position

An AHRS device is only able to provide a measurement of acceleration in the sensor frame directly. This measurement represents the sum of all linear accelerations affecting the sensor and gravity. A translational velocity can only be calculated if gravity is first removed form the acceleration measurement. The direction of the sensor frame Z axis is defined by the third column of the rotation matrix, therefore the direction of the Earth Z axis (gravity) in the sensor frame is defined by the third column of the transpose of the rotation matrix. This 3D value of gravity (of unit magnitude) may simply be subtracted from the accelerometer measurement $\boldsymbol{a}$ (units of g) to yield an accelerometer measurement with gravity removed $\tilde{\boldsymbol{a}}$. This is summarised in equation (3).

$$\tilde{\boldsymbol{a}} = \boldsymbol{a} - \begin{bmatrix} 2(q_1q_3 + q_0q_2) \\ 2(q_2q_3 - q_0q_1) \\ 2q_0^2 - 1 + 2q_3^2 \end{bmatrix} \tag{3}$$

The velocity can then be calculated as the integral of acceleration as defined by equation (4) and the position can be calculated as the integral of velocity as defined by equation (5).

$$\boldsymbol{v}_t = \boldsymbol{v}_{t-1} + \tilde{\boldsymbol{a}}_t.\Delta t \tag{4}$$

$$\boldsymbol{p}_t = \boldsymbol{p}_{t-1} + \boldsymbol{v}_t.\Delta t \tag{5}$$

Measurement errors will mean that the calculated velocity will drift, orientation estimation errors and large accelerations may mean that this drift becomes considerable. Bias errors in the calculated velocity mean that errors in the calculated position will rapidly accumulate. A practical solution to this is to use application specific information to indicate when the velocity is known to be zero and reset the calculated velocity value to zero accordingly. This approach has been previously been incorporated into inertial based pedestrian tracking algorithms [41]. In practise positional displacement information may still only be of a useful level accuracy if obtained within a relatively brief time interval of no more than a few seconds.

4.2 Discrete Control Mechanisms

The control mechanisms identified above accommodate many useful modes by which audio parameters may be controlled continuously. However, the continuous data streams may be analysed to identify discrete features within the raw and processed control values resulting from a particular pose or gesture. Successful identification of these gestures leads to the introduction of mechanisms producing discrete control parameters which can later be mapped to control state information.

4.2.1 Posture Identification

When the wearer's hand assumes a particular posture, the flexion values issued by the glove exhibit a unique pattern. Consequently, the problem of identifying a predefined vocabulary of postures from the glove data becomes a pattern matching problem. A range of techniques are available to address this problem but for this work an artificial neural network has been demonstrated to be reliable and robust when gloves are removed/replaced and between different users. Furthermore, the posture set may be easily changed by simply regenerating a new training set for the network. The neural network adopted here is a three layer feed forward structure referred to as the multilayer perceptron [8]. The structure of the network should be configured such that the number of neurones in the input layer is set equal to the number of sensors in the glove (14 for this work) and the number of output neurones is set equal to the number of different postures in the training set. The number of neurones in the intermediate layer should be set to a value somewhere between the quantity found in the input and output layers [8]. With a training set of known input (flexion) values for a set of known postures, the internal connections of the network can be configured using the backpropagation method, as previously described here [26]. Continuous analysis of the calibrated flexion readings enables the identification of postures to act as discrete control parameters.

4.2.2 Segmented Orientation

Division of the sensor orientation range into subregions enables the orientation to act as a discrete control mechanism.

Rotation Matrix Classification

Basic interpretation of orientation data is achieved through the calculation of an orientation as the sensor (hand) pointing either up, down, north, south, east or west. It is not sufficient to classify these cases using Euler angles due to the unpredictable behaviour that may occur when the Euler angle sequence approaches a singularity. An alternative means of classification is available though inspection of the rotation matrix elements. The first row of the rotation matrix defines the sensor X axis within the Earth X, Y and Z axes. If any of these elements is equal to one then the sensor X axis (aligned with low arm and wrist) must be aligned with the Earth axis and the sign of the unit value indicates the direction of the alignment. The classification logic would identify when the user was pointing in a direction if the rotation matrix element was equal to $1 \pm (cos(\theta) - 1)$. The effect is that there are 6 envelopes defined by cones, each with their central axis aligned to a separate axis of the Earth frame; the user is classified as pointing along an Earth frame axis if their arm falls within this envelope.

Euler Angle Classification

If care is taken to avoid issues associated with singularities, Euler angles may also be segmented to enable the identification of discrete angular regions [18]. To prevent unwanted oscillations at segment boundaries due to noise, a null region should be placed between segment boundaries to produce a simulated hysteresis effect. Integer values representing the current angular segment for rotation around the X, Y and Z axis may then be extracted as a discrete control parameter.

4.2.3 Inertial Peak Detection

With access to the continuous flow of raw inertial sensor data, peak detection algorithms can be used to search for value fluctuations resulting from sharp changes in motion. For instance, velocity changes along the X, Y and Z sensor axis manifest as peaks in the corresponding accelerometer readings. Similarly, rotational changes around the X, Y and Z axes produce peaks in the corresponding gyroscope value. Peak analysis algorithms generally adopt a threshold based procedure to signal the identification of maxima [16, 10, 14, 22]. The threshold should be set at a level above the usual range of operation to prevent unintentional peak identification and the process should include a suitable debouncing procedure to prevent single gestures from invoking multiple signals. Gestures stimulating the identification of inertial peaks can be signalled as control parameters, which are particularly useful for time sensitive control such as percussive gestures.

4.3 Combined Control Mechanisms

The continuous and discrete control mechanisms set out above provide a diverse range of control possibilities. However, by combining these mechanisms a further range of gestural control options emerge representing a mixture of discrete and continuous control which may combine both orientation and flexion data. These examples of combined control could be considered more expressive or intuitive as they incorporate the notion of metaphor, representing 'real-world' control interfaces [11]. Moreover, the combination of continuous and discrete control mechanisms enable the development of a state-based control system facilitating one-to-many gestural mappings. In practice this can be achieved by grouping specific audio parameters into modes, enabling a single gesture to be mapped to multiple audio parameters; this arrangement has the added benefits of minimising the likelihood of unintentional interaction and providing space for (unmapped) ancillary/performance gestures.

4.3.1 Ratcheting

Ratcheting is a combined control mechanism which enables a control value to be modified using a process that resembles the operation of a mechanical ratchet. The mechanism works by using the identification of discrete postures as an enabling mechanism for the traversal of a continuous control parameter or discrete control parameter using an estimate of relative angular displacement. For example, a clasping posture (Figure 6b), could be used to engage the addition of rotary displacement around the X axes. To the wearer, this process is analogous to turning a rotary encoder, albeit an invisible one. Alternatively, a two-fingered point posture may be used to enable and initialise the identification of subsequent swipe gestures using an estimation of positional displacement to produce a similar control mechanism.

4.3.2 Selective Orientation

Selective orientation is comparable with ratcheting, with postures used as an enabling mechanism for absolute orientation, rather than relative displacement. The continuous or segmented orientation of the hand produces a control parameter only when a specific hand posture is formed. For instance, a fist posture may be used to enable the control of an application parameter which is scaled to Euler angle of the y axes orientation. This combined control mechanism produces a gestural metaphor for the act of pulling a lever: the hand only controls the 'lever' when a fist posture is assumed, at all other times the wearer is able to move freely.

4.3.3 Segmented Threshold Triggering

The combination of segmented orientation with inertial peak detection enables the orientation of the hand to be taken into account when a sharp change in motion is detected. The control message that is produced as a result of the gesture may be controlled by the angular region occupied by the hand. An obvious example of this combined control mechanism would be for the control of drums [10], where the orientation of the hand selects a drum sound and the 'strike' gesture invokes its playback. Using rotation matrix classification, a bass drum could be triggered when the peak is detected if the hand points downwards, a snare drum selected if the hand points forwards and a high hat selected if the hand points upwards.

5. A MUSICAL APPLICATION

The toolbox of control mechanisms for a data glove and AHRS hand tracking apparatus was developed in anticipation of a four minute performance to be made by the composer/performer Imogen Heap at the TEDGlobal2011 conference in Edinburgh, UK [2]. Due to the narrow timeframe within which the collaborators could meet to build the system, the toolbox was conceived to accelerate the development of a gestural performance system where only the specifics of the hardware was known in advance and the gestural mapping and audio processes were to be developed later. Based upon our experiences developing 'Soundgrasp' [26] it was clear that the most appropriate and efficient mapping/design choices emerge when the performer is involved in the process. Consequently, the system architecture was defined to produce flexible mappings and the rapid development of audio processing algorithms.

5.1 System Overview

Figure 5 shows the major components of the hardware and its associated software. Each hand is fitted with a 5DT data glove, an x-io x-IMU, an LED module and a lavalliere microphone and a third voice microphone is also included. All sensors and audio devices were connected to the performer's Apple Macintosh computer where all sensor/audio data was processed. To act as a primary source of feedback to the performer [35], the auxiliary port on each x-IMU is used to control a set of RGB LEDs.

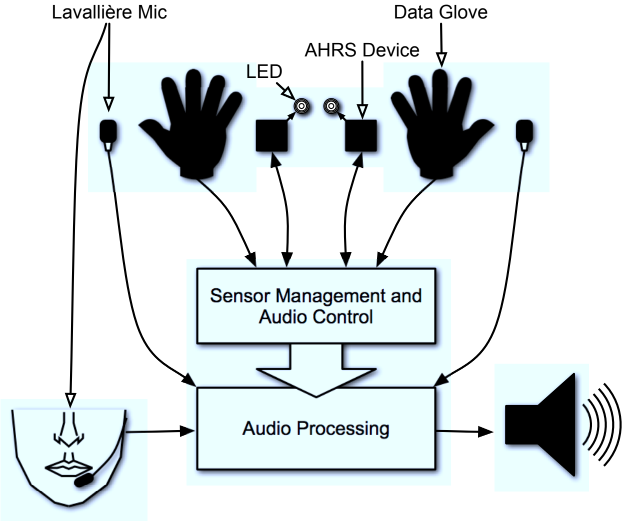

Figure 5: gestural device components

As shown, the software consists of two distinct parts: a sensor/audio control application and an audio processing application. Communication between the applications is made via Open Sound Control (OSC) [40]. The sensor/audio control application is a C++ application written using the libraries Juce [31] and oscpack [6]. This application implements the control mechanisms set out above and transmits commands that control the state of the audio processing application. The environment for audio processing was developed in Max/MSP to enable a range of audio recording, looping, synthesis and modification functions to be prototyped easily.

5.2 Mapping

Most traditional acoustic musical instruments have an intrinsic and rigid link between their control interface and sound-producing elements [21]. In contrast, the control input to digital musical instruments is often decoupled from the associated audio processes and one or more layers are inserted to translate control input into sound output [5]. Within these layers exists a huge design space which raises many questions about the nature of appropriate mappings and the design choices that provide opportunity for virtuosity [39] and lend themselves to engaging and expressive performances [11]. In addressing these questions, one approach is to analyse the natural movements emerging when subjects are asked to gesticulate while listening to music/sound [13, 20]. In other studies, emphasis is placed on the importance of including the performer in the design process [9, 27]. Given the nature of the application in this instance, the performer was central to the development of the mapping.

Prior to the commencement of development, the authors of this paper married the artistic vision for the performance with the available control mechanisms. A specification for the required audio features was established along with the initial plans for their associated gestural control mappings. With the requirement for a wide range of audio control options, it was clear that the audio parameters would have to be organised into modes and accessed via a state based system. The 6 modes are summarised below:

Voice mode enables the mono voice microphone audio to be recorded/overdubbed into the looper.
Wrist mode enables the left and right wrist microphone audio to be recorded/overdubbed into the left and right channels of the looper. This enables the performer to play and record acoustic instruments without the need for external microphones.
Effects mode enables the performer to independently apply reverb, panning and filter effects to the looper audio output and the live vocal microphone.
Synthesiser mode enables the live playback and recording of a two octave synthesiser.
Drum mode enables the live playback and recording of four different drum sounds.
Null mode enables free movement

The development of the control mechanism to audio parameter mapping was an iterative process directed primarily by the performer with input from co-authors and colleagues. The process involved experimentation with different mapping combinations both in the laboratory and rehearsal studios.

Central to the state control of the application was the discrete posture identification control mechanism. As the neural network could be easily configured to accommodate any distinguishable postures, the performer was free to develop their own posture set. For all modes of control, only four postures were required as shown figure 6.

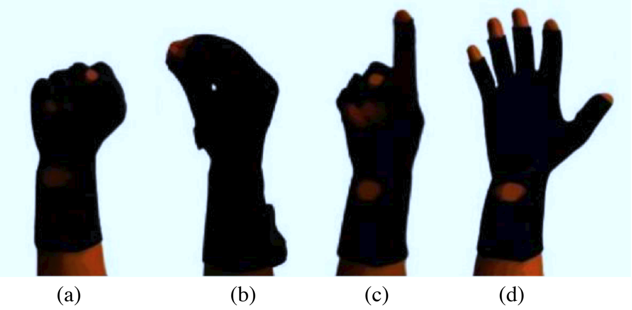
(a) (b) (c) (d)

Figure 6: chosen posture set

With discrete posture control mechanism established, the mode selection gesture is performed using selective orientation control mechanism where the rotation range around the sensor x axis is segmented into six regions, one for each mode as shown in figure 7. Mode selection was performed only when the left hand was open and the right hand formed a fist. This gesture was chosen as it is easy to perform and unlikely to occur incidentally, minimising the likelihood of unintentional mode switching. Each mode was ascribed a colour which was displayed on the left and right hand LED modules to provide feedback to the performer.

In *voice* mode, the voice microphone signal can be recorded/overdubbed into a two channel looper. This mode is controlled using a simple grasping gesture where record is enabled on the identification of an open hand posture (figure 6d) and disabled at all other times, an idea described previously in [26]. The voice microphone input may be recoded/overdubbed into synchronised four or eight beat loops for the left and right hand respectively. In *wrist* mode,

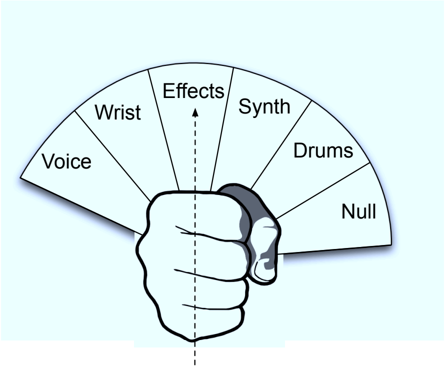

Figure 7: mode selection angles

the audio input received from the left and right wrist microphones can be recorded into a separate stereo looper. This enables the performer to play and record acoustic instruments where the record state is toggled when a rotational inertial peak is detected around the x axis of the right wrist, i.e. a flick of the wrist.

In *effects* mode, continuous control gestures with the right hand applied effects to the output mix of the looper and left hand gestures applied effects to the live vocal input. Reverberation and panning are controlled by the calibrated Euler rotation angles around the y and z axes: lifting the hand towards the sky introduced reverberation and pointing left and right resulted in panning. Furthermore, filtering was applied using the mean average of the finger flexion sensors for each hand. The wrist flick, as described in *voice* mode section, toggled the recording of automation for each of the effects.

The *synthesiser* mode introduced a combined control mechanism in which segmented orientation of the Euler angle rotation around the y axis selected the current note and the posture identification of an open hand was used to trigger note playback on a software synthesiser. Similarly, in *drum* mode, sounds were triggered with the identification of peaks along the z axis of the orientation sensor with the selection of the drum sound using the rotation matrix classification method. In both modes recording was toggled with the formation of a fist with the right hand.

6. APPLICATION EVALUATION

The development of the system was complete and stable and behaved as expected on the day of the performance where it received a positive reaction from the generous audience at TEDGlobal. However, the developmental process highlighted several notable points for consideration and identified areas for future development.

While the toolbox of control mechanisms was sufficient to implement the majority of gestures/mappings requested by the performer, the system as it stands was unable to accommodate them all. For example, the current sensor apparatus is unable to track the positioning of the hands with respect to the body. Consequently, gestures requiring the identification of this information were omitted. By placing additional orientation sensors at multiple points on the upper body, the relative 3D positions of the arms may be tracked, an approach described in [24]. Additionally, the control mechanisms delineated here are almost all derived from the instantaneous sensor data and do not identify temporal gestures requiring the analysis of a history of time frames. For the extraction of these types of gesture, analysis methods such as the hidden Markov model [7] or dynamic time warping [12] would be appropriate.

During the iterative development of the gesture to audio mapping, several areas were found to be problematic. For example, the performer wanted to engage record mode while playing the piano or Array mbira. A control mechanism had to be identified which would not restrict performance and would not occur incidentally. After several attempts with other control mechanisms, rotational peak detection was found to be the most appropriate; and with practice could be executed efficiently. However, on several occasions this gesture would cause the gyroscope sensor to saturate, which over a sustained period would cause the AHRS readings to drift.

Initial plans sought to map only those audio parameters that the performer used within their previous performances. However, with the capacity for gestural control, simple audio parameters were granted a new lease of life. Panning, for example when controlled directly by pointing was qualitatively regarded to be an effective and engaging mapping.

7. CONCLUSIONS

This paper set out a toolbox of control mechanisms for gestural music control in scenarios where finger flexion and hand orientation data is available. A selection of continuous, discrete and combined control mechanisms have been organised and presented which were developed in anticipation of a live performance to enable the rapid development of a gestural mapping system. The implementation details for the control mechanisms have been provided followed by an example application in the form of a live musical performance at the TEDGlobal2011 conference in Edinburgh. With easy access to a diversity of control mechanisms it was possible to quickly develop a usable and robust gesture to sound mapping in collaboration with the performer, facilitating a constructive and iterative development process resulting in mappings which emerged from the development sessions and rehearsals.

8. ACKNOWLEDGMENTS

The authors would like to thank Kelly Snook, Professor Tony Pipe, Professor Chris Melhuish and all colleagues at the Bristol Robotics Laboratory.

9. REFERENCES

[1] Oblong industries.

[2] Tedglobal, conferences.ted.com/TEDGlobal2011/.

[3] Fifth dimension technologies, www.5dt.com, 2012.

[4] x-io technologies, www.x-io.co.uk, 2012.

[5] D. Arfib, J. M. Couturier, L. Kessous, and V. Verfaille. Strategies of mapping between gesture data and synthesis model parameters using perceptual spaces. *Organised Sound*, 7(2), 2002.

[6] R. Bencina. oscpack, www.rossbencina.com/code/oscpack, 2012.

[7] F. Bevilacqua, B. Zamborlin, A. Sypniewski, N. Schnell, F. Guédy, and N. H. Rasamimanana. Continuous realtime gesture following and recognition. In *Gesture Workshop*, 2009.

[8] A. Blum. *Neural networks in C++: an object-oriented framework for building connectionist systems.* John Wiley & Sons, Inc., New York, NY, USA, 1992.

[9] M. A. Bokowiec. V'oct (ritual): An interactive vocal work for bodycoder system and 8 channel spatialization. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, 2011.

[10] S. Chantasuban and S. Thiemjarus. Ubiband: A framework for music composition with BSNs. In *Sixth International Workshop on Wearable and Implantable Body Sensor Networks*, 2009.

[11] S. Fels, A. Gadd, and A. Mulder. Mapping transparency through metaphor: towards more expressive musical instruments. *Organised Sound*, 7(2):109–126, 2002.

[12] N. Gillian, R. B. Knapp, and S. O'Modhrain. Recognition of multivariate temporal musical gestures using n-dimensional dynamic time warping. In *Proceedings of the International Coference on New Interfaces for Musical Expression*, 2011.

[13] R. I. Godøy, E. Haga, and A. R. Jensenius. Exploring music-related gestures by sound-tracing - a preliminary study. In *Proceedings of the COST287-ConGAS 2nd International Symposium on Gesture Interfaces for Multimedia Systems*, 2006.

[14] K. Hayafuchi and K. Suzuki. Musicglove: A wearable musical controller for massive media library. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, 2008.

[15] H. Ip, K. Law, and B. Kwong. Cyber composer: Hand gesture-driven intelligent music composition and generation. In *Proceedings of the 11th International Multimedia Modelling Conference*, 2005.

[16] I. Jang and W. Park. Signal processing of the accelerometer for gesture awareness on handheld devices. In *IEEE lnternational Workshopon Robot and Human Interactive Communication*, 2003.

[17] E. Jessop. The vocal augmentation and manipulation prosthesis (vamp): A conducting-based gestural controller for vocal performance. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, 2009.

[18] J.-H. Kim, N. D. Thang, and T.-S. Kim. 3-d hand motion tracking and gesture recognition using a data glove. In *IEEE International Symposium on Industrial Electronics*, 2009.

[19] M. Lee and D. Wessel. Soft computing for real-time control of musical processes. In *International Conference on Systems, Man and Cybernetics, Intelligent Systems for the 21st Century.*, 1995.

[20] M. Leman, F. DESMET, F. Styns, L. Van Noorden, and D. Moelants. Embodied listening performances reveal relationships between movements of player and listeners. In *Proceedings of ENACTIVE*, 2007.

[21] D. J. Levitin. Control parameters for musical instruments: a foundation for new mappings of gesture to sound. *Organised Sound*, 7(2), 2002.

[22] R. Lugo and D. Jack. Beat boxing: Expressive control for electronic music performance and musical applications. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, 2005.

[23] P.-J. Maes, M. Leman, K. Kochman, M. Lesaffre, and M. Demey. The "one-person choir": A multidisciplinary approach to the development of an embodied human-computer interface. *Computer Music Journal*, 35(2):22–35, June 2011.

[24] P.-J. Maes, M. Leman, M. Lesaffre, M. Demey, and D. Moelants. From expressive gesture to sound. *Journal on Multimodal User Interfaces*, 3(1), 2010.

[25] E. R. Miranda and M. M. Wanderley. *New Digital Musical Instruments:Control and Interaction Beyond the Keyboard.* A-R Editions, 2006.

[26] T. Mitchell and I. Heap. Soundgrasp: A gestural interface for the performance of live music. In *International Conference on New Interfaces for Musical Expression*, 2011.

[27] J. Moen. From hand-held to body-worn: embodied experiences of the design and use of a wearable movement-based interaction concept. In *Proceedings of the International Conference on Tangible and Embedded Interaction*, 2007.

[28] G. Saggio, F. Giannini, M. Todisco, and G. Costantini. A data glove based sensor interface to expressively control musical processes. In *4th IEEE International Workshop on Advances in Sensors and Interfaces*, 2011.

[29] M. Sama, V. Pacella, E. Farella, L. Benini, and B. Riccó. 3dID: a low-power, low-cost hand motion capture device. In *Proceedings of the conference on Design, automation and test in Europe: Designers' forum*, 2006.

[30] H. Sawada and S. Hashimoto. Gesture recognition using an acceleration sensor and its application to musical performance control. *Electronics and Communications in Japan (Part III: Fundamental Electronic Science)*, 80(5), 1997.

[31] J. Storer. Juce, www.rawmaterialsoftware.com, 2011.

[32] D. Sturman and D. Zeltzer. A survey of glove-based input. *Computer Graphics and Applications, IEEE*, 14(1):30 –39, jan 1994.

[33] T. Todoroff. Wireless digital/analog sensors for music and dance performances. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, 2011.

[34] G. Torre, J. Torres, and M. Fernstrom. The development of motion tracking algorithms for low cost inertial measurement units. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, 2008.

[35] M. Wanderley and P. Depalle. Gestural control of sound synthesis. *Proceedings of the IEEE*, 92(4):632 – 644, April 2004.

[36] M. Wang, Y. Yang, R. Hatch, and Y. Zhang. Adaptive filter for a miniature mems based attitude and heading reference system. In *Position Location and Navigation Symposium*, 2004.

[37] R. Y. Wang and J. Popović. Real-time hand-tracking with a color glove. *ACM Transaction on Graphics (SIGGRAPH 2009)*, 28(3), 2009.

[38] G. Welch and E. Foxlin. Motion tracking: No silver bullet, but a respectable arsenal. *IEEE Computer Graphics and Applications*, 22(6):24–38, 2002.

[39] D. Wessel and M. Wright. Problems and prospects for intimate musical control of computers. *Computer Music Journal*, 26(3), 2002.

[40] M. Wright. Open sound control: an enabling technology for musical networking. *Organised Sound*, 10(3), 2005.

[41] X. Yun, E. R. Bachmann, H. Moore, and J. Calusdian. Self-contained position tracking of human movement using small inertial/magnetic sensor modules. In *IEEE International Conference on Robotics and Automation*, 2007.