The SARC EyesWeb Catalog: A Pattern Recognition Toolbox for Musician-Computer Interaction

Nicholas Gillian
Sonic Arts Research Centre
Queen's University Belfast
Northern Ireland
ngillian01@qub.ac.uk

R. Benjamin Knapp
Sonic Arts Research Centre
Queen's University Belfast
Northern Ireland
b.knapp@qub.ac.uk

Sile O'Modhrain
Sonic Arts Research Centre
Queen's University Belfast
Northern Ireland
sile@qub.ac.uk

Abstract

This paper presents the SARC EyesWeb Catalog (SEC), a group of blocks designed for real-time gesture recognition that have been developed for the open source program EyesWeb. We describe how the recognition of real-time body movements can be used for musician-computer-interaction.

Keywords: SARC EyesWeb Catalog, gesture recognition

1. Introduction

In human-human interaction, gestures afford the transmission of communicative information that can directly convey emotion, intention, assertion or indirectly augment and elaborate spoken language. It has long been the goal for researchers in the field of human-computer interaction (HCI) to create interactive systems that can afford the user the ability to interact using gestures, such as body movement, in a manner that reflects the simplicity, robustness, and effortlessness of human-human interaction. One area of HCI that would greatly benefit from the advance in gestural recognition is the field of musician-computer interaction (MCI).

Achieving this level of specific control is difficult due to a number of factors. Complex pattern recognition algorithms are normally required to achieve such a dynamic state of control and expensive or specific equipment is required to facilitate the use of pre-designed MCI systems. It may also be impractical for musicians to use working systems built by prior artists due to the availability of matching sensor hardware or inflexible software design.

The term gesture itself is very difficult to define, as gestures only make sense when made in their correct context and when both performer (the individual performing the gesture) and perceiver (the individual perceiving the gesture), are using the same 'gestural vocabulary' [1]. This 'gestural vocabulary' is acquired through many different methods, such as being naturally acquired over time through interacting with other individuals or learnt as a specific skill. Gestures are a natural and informative method of communication; however they are context dependent and are confined by the performer and perceivers culture, skills and prior knowledge. For a machine therefore to act as a perceiver it must be context aware and understand the 'gestural vocabulary' being used by the performer.

The work presented here tries to address these problems, not by giving a specific definition of what a gesture must be or by giving a one-off fix that will only work for one type of performer or hardware setup, but by providing a specific set of software tools that can be quickly and easily shaped by the user to create dynamic user specific gesture recognition interfaces and that allow the user to specify what type of gesture they want to use and how they want to use it.

2. Background

There has been much research carried out in the field of gestural recognition for MCI, especially over the last decade, with advances in computer processing power and the decreasing price of pre-built wireless sensor units such as the Wii^{TM}, allowing users even with a limited budget to create their own unique MCI systems. The work here is certainly inspired by previous work by the likes of Mulder et. al. [2] and Pritchard et. al. [3].

3. SEC

The SARC EyesWeb Catalog (SEC) is a group of blocks that can be used in the open source platform EyesWeb. EyesWeb is a graphical development environment (GDE) similar to programs such as Max/MSP or Labview, in which the user can build a working system graphically by dragging and dropping blocks onto an area that resembles a signal-flow diagram. The SEC expands the current wide ranging EyesWeb algortihms with a number of blocks designed for real-time feature extraction, classification and gesture analysis. Version 1.0 of the SEC, which is available to download and use free of charge from [4], contains over 40 blocks such as:

- Blocks to perform mathematical algorithms such as

Matrix Inverse, Invert, Transpose, Vector Cosine Angle, Euclidean Distance and Mahalanobis Distance

- Blocks for signal processing such as FIR Filter, Savitzky-Golay Filter, Dead Zone and Envelope Extraction
- Blocks for gesture recognition such as Principal Component Analysis, Fuzzy C-Means clustering and K-Means clustering
- Blocks for interfacing with external hardware sensors such as the Wi-mirco Dig or the SHAKE SK6 inertial measurement unit

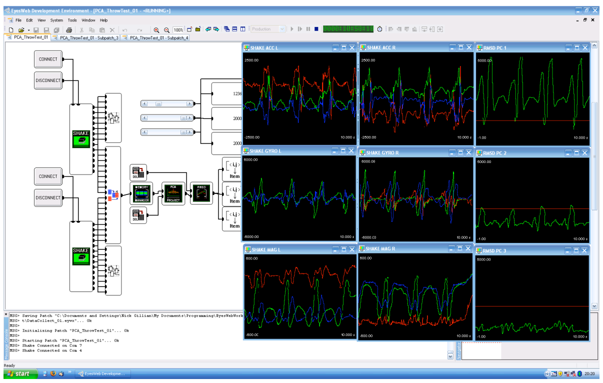

Figure 1. A recognition patch built using some SEC blocks.

Encapsulating the gesture recognition algorithms such as PCA within a GDE such as EyesWeb gives the benefits of creating fast prototypes for testing different gestures using various sensor configurations quickly and efficiently. Users can modify patches used to recognise a set of gestures, creating a brand new patch that can recognise a set of completely different gestures in a matter of minutes. The blocks have been developed so that they are as flexible and adaptable as possible; allowing, for example, different users to use different hardware units as the sensor inputs to the same recognition patch.

4. Gestural control for MCI

In an MCI scenario, two types of gestural control are desirable: continuous control and discrete state control. Both types of control are useful in a live performance scenario, for example continuous control could be used to control the level of a particular parameter or effect, whilst discrete state control could be used to indicate to the computer that a specific section of the piece has been reached and it should therefore perform task *x*.

A system using pattern recognition algorithms can be used as a discrete state controller that, once a specific gesture is detected, can trigger any number of sonic events. This creates the capability for the performer to sequence actions without the need for mediation by a classical physical interface (what Wanderley calls “remote interaction” [1]). The disadvantage of this technique is that only after a gesture is made, can recognition occur.

To mitigate this disadvantage, various distance measures can be used to provide continuous control. As the performer creates a gesture that begins to approach the gesture being classified, the distance measure will begin to drop, reaching its minima when the performer most approximates the target gesture. The disadvantage of this technique is the complex interaction between the gesture and the derived distance measure that makes it difficult to have a high degree of precise control. Combining both the discrete and continuous aspects of gestural control creates a more natural interface that fuses the advantages of each method while overcoming some of the disadvantages. The continuous control provides *a-priori* information and the discrete control provides a precise end-point of the gesture.

This technique is currently being used for a piece, The Biomuse Trio, for Violinist, Laptop, and BioMuse by Eric Lyon. In this piece, the Biomuse performer uses kinematic sensors combined with physiological sensors to measure overt physical gestures as well as changes in emotional state. Smooth arm gestures by the Biomuse performer, measured and processed using the SEC, are used to control continuous parameters within MAX/MSP on the laptop. The discrete end-points are then recognized and used as triggers. The movements and physiological changes associated with “anger” and “serenity” are also used for event control.

5. Future Work

The current version of the SEC does not feature any algorithms for accounting for the temporal variability inherent within any biological system, therefore future versions of the SEC will feature blocks for dynamic time warping to overcome this problem. Blocks are also currently being developed to extend the pattern recognition toolbox, including algorithms for Fuzzy Logic, Hidden Markov Models and Artificial Neural Networks.

6. Conclusion

We have presented the SARC EyesWeb Catalog, a new group of blocks for the open source program EyesWeb that can be used to recognise human movement gestures.

References

[1] E. R. Miranda and M. M. Wanderley, *“New Digital Musical Instruments: Control and Interaction beyond the Keyboard”*, A-R Editions, Spring 2006.

[2] A. G. Mulder, *“Design of Virtual Three-dimensional Instruments for Sound Control”*, PhD thesis, School of Kinesiology, Simon Fraser Universtiy, 1998.

[3] B. Pritchard and S. Fels, *“GRASSP: Gesturally-Realized Audio, Speech and Song Performance”*, Proceedings of the 2006 International Coference on New Interfaces for Musical Expression (NIME06), Paris, France, June 4-8, 2006.

[4] SEC [Online] http://www.sarc.qub.ac.uk/~ngillian/SEC.html