Kontrol: Hand Gesture Recognition for Music and Dance Interaction

Kameron Christopher   Jingyin He   Raakhi Sinha Kapur   Ajay Kapur

California Institute of the Arts
24700 McBean Parkway
Valencia, CA 91355, USA

ABSTRACT

This paper describes Kontrol, a new hand interface that extends the intuitive control of electronic music to traditional instrumentalist and dancers. The goal of the authors has been to provide users with a device that is capable of detecting the highly intricate and expressive gestures of the master performer, in order for that information to be interpreted and used for control of electronic music. This paper discusses related devices, the architecture of Kontrol, its potential as a gesture recognition device, and several performance applications.

Keywords

Hand controller, computational ethnomusicology, dance interface, conducting interface, Wekinator, wearable sensors

1. INTRODUCTION

In today's digital age, the traditional performing artist (be it instrumentalist or dancer) is faced with an ever-growing interest in extending their media. There is a long history of using micro controllers and sensors to create interfaces that interact in ways that appear more embodied.

The hands are an important form of non-verbal communication in which humans rely for interaction. They help relay emotions, emphasize concepts, and often help in explanation. The hands are the basis of sign language, which relies heavily on hand movements and gesture for everyday communication and interaction with the world. The hand movements of a skilled musician can provide very useful information on what the musician is playing, how they are playing it, and possibly what they intend to do next. Musicians develop complex muscle memories in order to execute various performative tasks with the greatest effect and efficiency. Many of these movements have become intuitive or have developed naturally and the musicians themselves would not be able to describe them. Building an interface that is capable of detecting and making use of this information in musical ways was the goal of the authors.

Many in the field of music have sought the creation of devices for musical control that function in a way that is more expressive and musical in a traditional sense. Tod Machover was an early innovator in these musical technologies with his idea of hyper-instruments [1].

In 1984, Michael Waisvisz brought popularity to the idea of hand based musical controllers with his introduction of *The Hands* [2]. The hands were constructed of numerous sensors attached to the user's arms, hands, and pants that allowed for the control of music. Since then, many gesture-based designs for musical control have been created [3].

Simultaneously, numerous hand-based devices were being developed for non-musical uses, such as communication, computing, and gaming [4]. Mattel's *Power Glove* was the first attempt to create a commercially available glove input gaming device and was being developed at the same time as the *DataGlove* [5], [6]. Stanford produced the *Thumbcode* as a digital sign language device in 1998 [7]. The development of hand-based devices has continued to grow in these fields[4]. Additionally, methods for gesture recognition in musical application continue to develop [3].

Kontrol aims to account for the traditional manner in which artists perform their art, and not interfere with their ability to do so. Kontrol is a hand interface built to extract the intricate expressivity of performers of various performative backgrounds, including dancers and instrumentalist. The goal of Kontrol is to provide a tool for the intuitive interpretation of gestures that interfaces with the performer in conjunction with the performative task. The use of triple-axis accelerometers on each finger and the hand provide high-resolution detail about a performer's actions and an ergonomically lightweight and non-intrusive design for the performer. Kontrol builds upon these findings as a tool for intuitive expression in digital music.

The design and architecture of Kontrol is described in section 2. Section 3 discusses the prospects of using gesture recognition with Kontrol. Several projects in which Kontrol has been used are described in section 4. Section 5 provides a conclusion.

2. Design and Architecture

Kontrol was designed to be lightweight and unobtrusive. Six triple-axis accelerometer are especially customized to sit on the left hand: 5 on the fingers and 1 on top of the hand (Figure 1). The sensors are carefully calibrated and designed to lay out ergonomically on the left-hand and are connect through a multiplexer to an Arduino Fio, as seen in Figure 2. The use of accelerometers allow for the recognition and parameterization of the subtlest of movements.

NIME'13, May 27-30, 2013, KAIST, Daejeon, Korea.

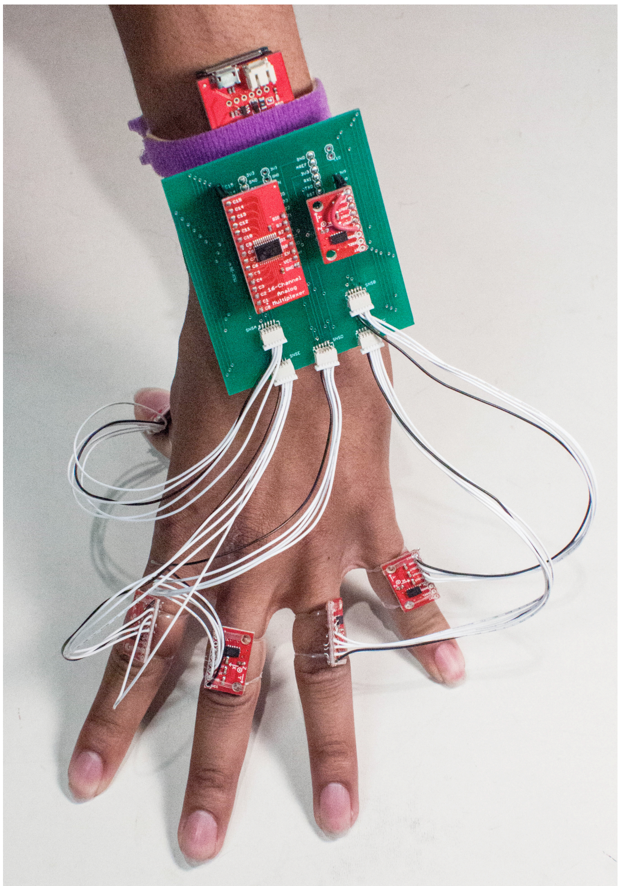

Figure 1: Kontrol

The number of accelerometers necessitated the use of a multiplexer in order to input all of the accelerometers' data into the Fio. A lithium rechargeable battery powers the system, which provides reliable mobility for Kontrol. The use of Xbees allows for wireless communication, sending the data from the Fio directly to the computer.

An overview of data flow from the microcontroller to the computer through wireless communications can be seen in Figure 3.

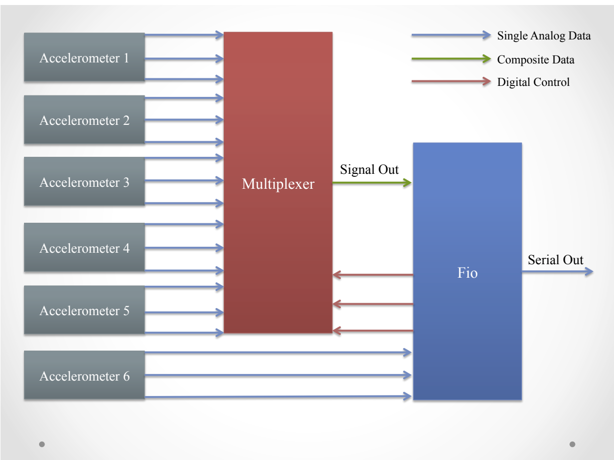

Figure 2: Overview of Data-Handling

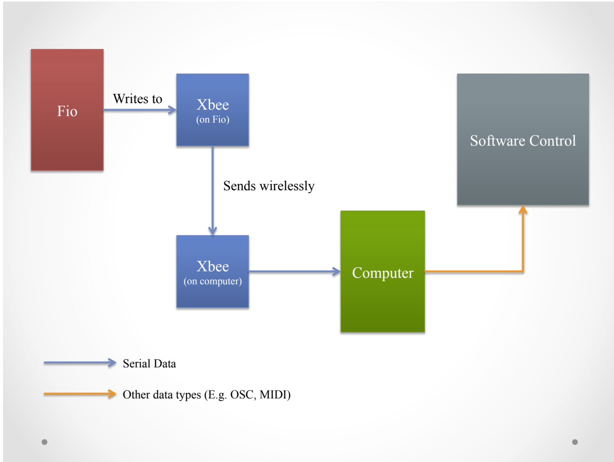

Figure 3: Overview of System Communication

To achieve a compact setup, the authors designed a custom PCB to connect all of the components for Kontrol to the Arduino Fio. The PCB also provides for a stable structure for Kontrol to sit on one's hand.

3. Gesture Recognition

Kontrol provides very high-resolution detail about the performer's hand and finger movements, which can be extremely useful in determining a performer's intentions and what a performer's action should trigger. As such, we need a way to extract the meaning of these gestures that is useful for various types of performances. Individually addressing the data from the accelerometers into meaningful and useful parameters is time consuming, and requires the logging and processing of large data sets for each performer. This inconvenience necessitates the use of machine learning to classify gestures. The classification of the data falls into two modes: Static and Dynamic.

Static classification allows us to classify single feature vectors. Theses classifiers are used when the performer wants instantaneous feedback on a given hand position. An example would be a flautist using a particular fingering as an indicator or trigger.

Dynamic classification allows us to classify a sequence of vectors. These algorithms take into account temporality, and allow us to classify gestures unfolding over time - A dancer performing a movement, which is to be accompanied by a triggered musical gesture.

Wekinator is the software used to realize the classification process. Developed by Rebecca Fiebrink, Wekinator is a tool for real-time feature extraction and on-the-fly machine learning [8]. It includes several machine learning algorithms and allows for the input of the data from Kontrol. The system allows us to easily classify data from Kontrol simultaneously in multiple modes. Using our custom-built software in conjunction with Wekinator, the incoming x, y and z values from each of the accelerometers are packed into a single Open Sound Control (OSC) [9] message. Once Wekinator classifies the incoming data, it sends the results of the classification back to our software, which then allows it to be used for various purposes. This process is demonstrated in Figure 4.

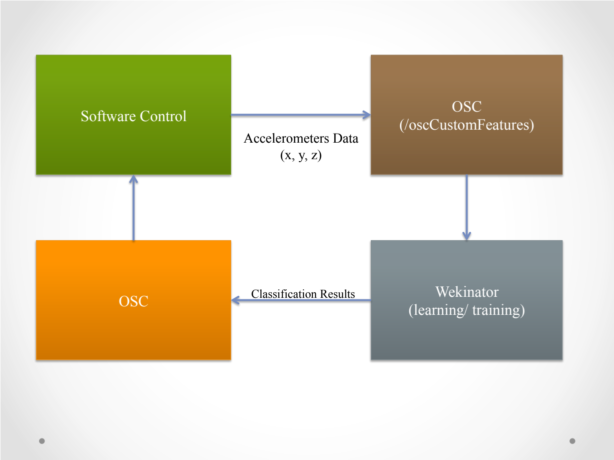

Figure 4: Overview of Feature Extraction to Wekinator

4. Applications

This section describes several experiments in which Kontrol is used in a variety of performance contexts, including: The Guqin, The Saxophone, South Asian Dance, and Conducting. Through these contexts, we were able to fine to our software and design.

4.1 Extending the Guqin

Kontrol is extendable to various performance contexts. One such application of Kontrol is with the Guqin, a Chinese Han instrument. The Guqin is a seven-stringed, fretless plucked instrument with a history of about 5000 years. In Qin music tradition, the timbre of the notes is of utmost importance [10] It is said that the fluid onset of tone(s) with varying timbres is the main focus of its musical aesthetics. The Qin's beauty lies more in each separate note as compared to those in succession, where each individual tone is unique to the listener. One can only fully understand the aesthetical pleasures of the Guqin when they are able to distinguish the subtle differences in "tone colors" produced on the same pitch. The essential quality of traditional Qin playing is said to be "Painting with sound" [10]. The finger technique used can be compared to the "brush strokes", where as which string the note is played on can be compared to the physical "color" of the sound. One can play the same pitch several ways: on different strings, or with different fingers.

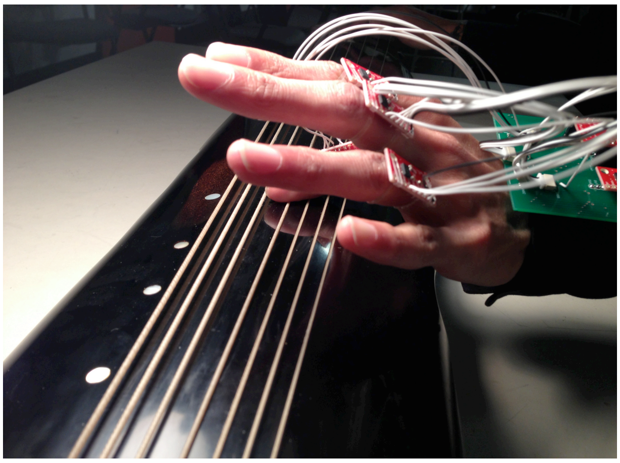

Figure 5: Kontrol in use with Guqin

A distinctive feature of Qin playing is the execution of finger techniques that bring about sonorous tones [11]. In performing Qin music, playing techniques are often subjective to one's interpretations and teachings, resulting in the formation of different playing styles [11]. Kontrol gives the performers the freedom of movement and the alibi of extended limbs by consolidating micro-controllers and sensors in a small package (Figure 5), while using wireless technology for data communication between the micro-controller and the computer. The objective is to explore and further the instrument's potential in the modern digital age, by extending performance practices beyond the techniques of traditional repertoire using micro-controllers.

Utilizing Kontrol's gesture recognition system, it allows the performers to parameterize performance gesture for further applications. One example is to map different performance gestures as event triggers while parameterizing derivative data of hand movements to parameters of audio signal processes. Other applications with Kontrol and the Guqin include its pedagogical use in training finger techniques for both hands.

4.2 Extending the Saxophone

Kontrol offers the saxophonist the ability to directly control software parameters while playing his instrument. The saxophonist often requires both hands while performing and when interfacing with computer music often has to rely on strictly predefined scripts, others controlling parameters, or stepping away from the instrument itself to perform a computer operation.

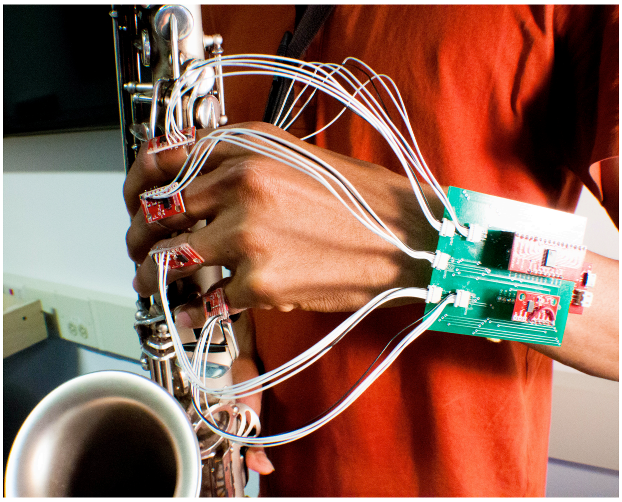

Figure 6: Kontrol with Saxophone

Kontrol provides a solution that is more intuitive to the saxophonist and has removed the need to step away from the instrument in order to gain fine control over the computer.

The gesture recognition system with Kontrol has allowed for the same fingering and hand positions that the saxophonist is accustomed to using while playing his instrument to be assigned to customizable parameters and the direct control of software synthesizers. Future works with Kontrol will include intelligent computer systems that can respond musically to what the saxophonist is playing.

4.3 Digitizing South Asian Dance Forms

Intricate body isolations are a defining feature of South Asian dance. From the eyes, eyebrows, and angles of the body, each story is conveyed through an understanding of symbolic hand poses. The wrists, palms and fingers of the hand express different meanings through 'Mudras' the stretching of the fingers in different positions. Single and double hand gestures are a sign language expressing masculine and feminine ideology within dance. For example, the splayed fingers can symbolize a lotus flower and blossoming of the mind, where the flat palm represents self, direction and limitations. The use of hands alone can lead the audience through a full composition. Figure 7 shows a variety of different "Mudras".

Figure 7: Variety of Different "Mudras"

The third and fourth authors have designed wearable sensor technology with the use of accelerometers attached to wrists of performers, inspired by Tomie Hahn's *PikaPika* [12]. Though this has been a very fruitful endeavor, the motivation to make fully implemented hand is driven by the need for more control of different aspects of multimedia interaction during choreography. The ability to use gesture recognition to trigger sections of composition is very powerful. Before now, the dancers had to rely on either fixed timing of a composition, or a second operator manning a computer controlling which section beginning and endings. Being able to tie traditional indian dance forms from "Mudras" described above, allows for extended control using the Kontrol, but in a developed and historical context.

4.4 Conducting

The use of Kontrol in the context of a conductor leading a virtual or mechanical ensemble is also very powerful. Using musical techniques inspired by Max Mathews score following [13], the forth author has done work conducting the robots in the Machine Orchestra [14]. Tiny finger gestures can be used to control quiet by very fast repetitive movements in mechanical movement, coined by the forth author as the "butterfly effect" [14]. Gesture recognition can also be used recall entire sections of music without the use of a button driven interface. In its present form, Kontrol is a one-hand interface, but experiments in dance and conducting inspire making sensors for the second hand.

5. Conclusion

In this paper, we have presented the design and architecture of a newly customized performance interface, Kontrol. Through its applications in the various forms of performing arts, it has demonstrated its functionality as an intuitive device for the realization of extending performance aesthetics without compromising traditional techniques. The experiments in the various contexts include instrumental performances on the Guqin and Saxophone, South Asian Dance and conducting.

Many of the design considerations for Kontrol proved to be successful. The utilization of Xbee allows seamless and reliable wireless communication between the Fio and the computer. This provides mobility for the performers and the absence of a computer on stage, which may otherwise affect the performance aesthetics. More so, the accelerometers allow high precision in gesture recognition and provide accurate data to even the subtlest movement of the fingers.

However, using readily available hardware has prevented us from further reducing the overall dimensions of Kontrol. To overcome this, the newer version will look into custom-design circuits that integrates the multiplexer and Fio into a consolidated package. Furthermore, the surface mount JST-like connectors require high level soldering skills and are very fragile.

Future designs may include an ergonomic calibration for the right-hand, and perhaps the incorporation of nanotechnology. The authors hope to see Kontrol in many other performing arts context, as well as in pedagogy.

6. REFERENCES

[1] T. Machover, “Hyperinstruments: A composer’s approach to the evolution of intelligent musical instruments,” *Organized Sound*, pp. 67–76, 1991.

[2] M. Waisvisz, “The hands, a set of remote midi-controllers,” in *Proceedings of the International Computer Music Conference, San Francisco, CA: International Computer Music Association*, 1985, pp. 86–89.

[3] M. M. Wanderley, “Gestural control of music,” in *International Workshop Human Supervision and Control in Engineering and Music*, 2001, pp. 632–644.

[4] L. Dipietro, A. M. Sabatini, and P. Dario, “A survey of glove-based systems and their applications,” *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 38, no. 4, pp. 461–482, 2008.

[5] D. J. Sturman and D. Zeltzer, “A survey of glove-based input,” *Computer Graphics and Applications, IEEE*, vol. 14, no. 1, pp. 30–39, 1994.

[6] T. G. Zimmerman, J. Lanier, C. Blanchard, S. Bryson, and Y. Harvill, “A hand gesture interface device,” in *ACM SIGCHI Bulletin*, 1987, vol. 18, pp. 189–192.

[7] V. R. Pratt, “Thumbcode: A Device-Independent Digital Sign Language,” in *Proceedings of the 13th Annual IEEE Symposium on Logic in Computer Science, Brunswick, NJ*, 1998.

[8] R. Fiebrink and P. R. Cook, “The Wekinator: A System for Real-time, Interactive Machine Learning in Music,” 2010.

[9] M. Wright and A. Freed, “Open sound control: A new protocol for communicating with sound synthesizers,” in *Proceedings of the 1997 International Computer Music Conference*, 1997, pp. 101–104.

[10] R. H. Van Gulik, “The Lore of the Chinese Lute. An Essay in Ch’in Ideology,” *Monumenta Nipponica*, pp. 386–438, 1938.

[11] “John Thompson on the Guqin Silk String Zither.” [Online]. Available: http://www.silkqin.com/. [Accessed: 11-Feb-2013].

[12] T. Hahn and C. Bahn, “Pikapika–the collaborative composition of an interactive sonic character,” *Organised sound*, vol. 7, no. 03, pp. 229–238, 2002.

[13] R. Boulanger, M. Mathews, B. Vercoe, and R. Dannenberg, “Conducting the MIDI Orchestra, Part 1: Interviews with Max Mathews, Barry Vercoe, and Roger Dannenberg,” *Computer Music Journal*, vol. 14, no. 2, pp. 34–46, 1990.

[14] A. Kapur, M. Darling, D. Diakopoulos, J. W. Murphy, J. Hochenbaum, O. Vallis, and C. Bahn, “The machine orchestra: An ensemble of human laptop performers and robotic musical instruments,” *Computer Music Journal*, vol. 35, no. 4, pp. 49–63, 2011.