2010年10月25日(月)　社団法人映像情報メディア学会技術報告　ITE Technical Report Vol.34, No.44
ME2010-147(Oct.2010)

Dancing Sound

Vincent de Paul ATANGANA EKANI† Haoran XIE† Kazunori MIYATA†

† Japan Advanced Institute of Science and Technology 1-1 Asahidai, Nomi, Ishikawa, 923-1292 Japan

E-mail: † {Vincent-atangana, xiehr, miyata}@jaist.ac.jp

Abstract This paper describes a new way of playing music by using motion tracking with Wii MotionPlus^{TM} and sensor accelerator, by a person. The results of motion detection produce sound and make him/her be able to play a song by following the various movements like in a musical score. The gestures to be reproduced are projected in front of him/her, by using a video projector. The detection of Wii Remote accelerations and Euler angles will produce sound effects like volume variation.

Keyword Accelerometer, Gesture recognition, Human computer interaction, Motion capture, Wii Remote

1. Introduction

Wii remote development started in 2001. The same year, Nintendo Licensed from Gyration Inc, a company that produces wireless motion-sensing computer mice, a number of motion-sensing patents. Recent advances in motion detection and computation have created many opportunities of using computer vision and motion detection techniques to video game for entertainment[1], computer aided coaching for education. This mode of utility is an adaptation from its intended use as a game controller, and requires evaluation of its functions in music applications in order to understand its possibilities and limits. The goal of our project is to make a full body tracking system for low cost but effective. The player will be able to play a song with simple motions of dance as in a musical score. The system uses a low cost assembly formed by four Wii Remotes. Now, most dancing games use at most, two Wii Remotes and then only to coordinate the movements of hands. Our challenge is to use four Wii Remotes at the same time: two for hands, and two for feet. The main risk here is that certain movements cannot be detected in real time, for reasons of noise pollution, or because of poor execution movements. Summary of our research and founding is then presented to support the use of Wii Remote for full body tracking as a cheaper and reliable solution. It concludes with a brief look of our future research opportunities.

2. Related Work

Since the release of the Wii console, people have been exploring different ways in which to use the Wii Remote, and especially, how to make tracking with Wii Remote. Johnny Lee has posted video demos and sample codes in his website related to the use of the Wii Remote for finger tracking, and head tracking for desktop VR displays [2].

About the motion capture system using acceleration sensor, Tamura et al. made a motion capture system using biological motion data by Wii Remote[3], and proposed an approach of recognizing walk style with human emotion using single acceleration system[4], in four categories, anger, joy, sad an neutral. An AR system using a pair of accelerometers is conducted by Takada et al.[5], can recognize five types of body motion, sitting, standing, walking, running and biking.

One of the most popular in music games is “Just Dance”[6], developed by Ubisoft in 2009. The principle of this game is to hold the handle in player’s right hand and replicate the model improbable gestures on the screen. However, the recognition rate of a motion in “Just Dance” is imprecise, because only the movement of right hand has been taken into account. In our system, we will resolve all these problems. We use four Wii Remotes simultaneously, each Wii Remote is equipped by a Wii MotionPlus^{TM} ; It allows a player to make movement with hands and feet, in better recognition rate.

3. System Configuration

It is convenient to define the whole system as a new musical system interface using Wii Remotes. To implement our system(Fig.1), we need ① four Wii Remotes, four Wii MotionPlus^{TM}, ②a projector, a screen, ③a computer, ④speakers that will show players the various movements to execute. The Wii Remotes are

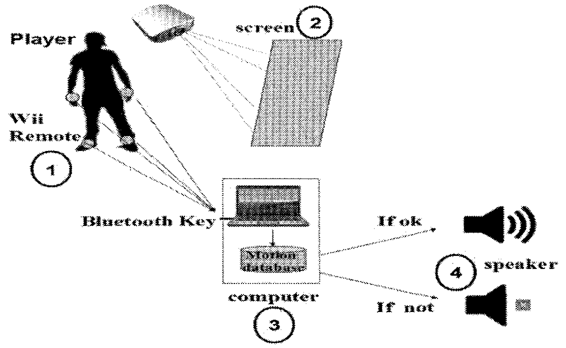

Fig.1 System Configuration

placed in the hands and feet of players, to allow them to use their wrists and ankles to better execute the movements. We use Wii MotionPlus™ that captures more accurate and complex motion. The acceleration data and Euler angle data are transmitted via Bluetooth. The user must follow every step of dance, with gestures of hands, and feet.

4. Experience

This chapter describes the system components in detail.

4.1. Data exploitation

An accelerometer based gesture recognition is known approach and was discussed by Hofmann et al.[7]. The gesture recognition on Wii Remote controller is also known task and it was discussed by Schlömer et al.[8]. Our approach tries to improve accuracy of gesture recognition using more features for classification and different design of recognition using analysis of data from accelerometers.

Data retrieved from accelerometer sensor in a Wii Remote controller are defined by five components: ax, ay, az, which represents device acceleration in X, Y, Z axes respectively, and Pitch, Roll for Euler angles. The simple dance motion data are captured by using Wiiyourself![9], native C++ library for Windows, as shown in Fig.2. The angle range of pitch is [-180,180], and for roll [-90, 90]. Because the range of the value of acceleration data is less than Euler angles, so in order to show all the five curves in one figure, the values of Euler angles are represented by the values of (angles/20) in Fig.2. The computer receives the data at approximately 89Hz sampling rate. We adopted the motion in 30 seconds in Fig.2. The ordinate is the data of acceleration (G), and abscissa is time in seconds. In this example, because of a more drastic motion by hand than foot, so the curves are more zigzag.

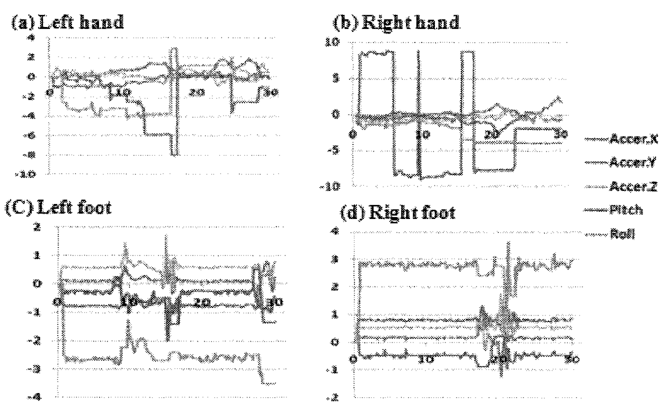

Fig.2 Five captured attributes

4.2. Determination of Feature Data

The value $P = \sqrt{x^2 + y^2 + z^2}$ is the length of the composite vector of the acceleration data in three directions in orthogonal coordinate. It is easier to analyze a characteristic of motion data by the curve using P than three curves in different directions (Fig.3). The calculated result is in ordinate, and time in second for abscissa. This feature data P are used in Discrete Fourier Transform (DFT) by FFTW[10] library. The curves of four parts motion data is shown in Fig.3.

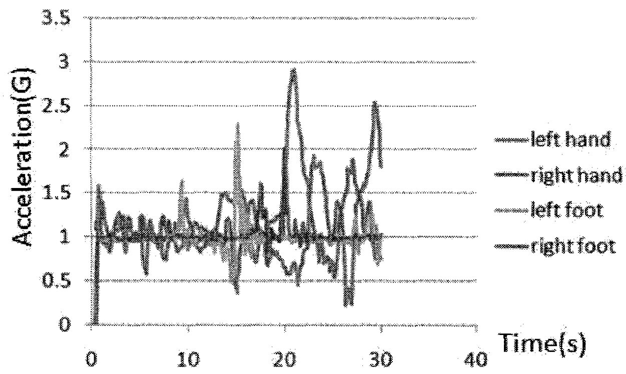

Fig.3 Feature data using P

And then about the DFT calculation, the window size is set to 256 data while using FFTW library, and the data can be acquired by the order of frequency. Then principal components analysis (PCA) is applied to compress the data, contribution rate is set to 95% above.

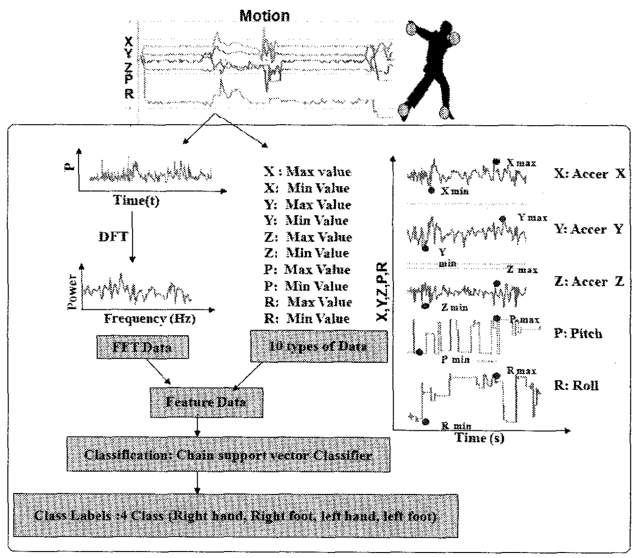

Fig.4 Dataflow of motion recognition system

In Fig.4, we also calculate the max and min values of Wii Remote signals. Then 10 types of results can be given from this process, which are considered with FFT data as feature data. These data are set into Chain Support Vector Classifier(C-SVC)[11]. The C-SVC is a four class's classifier based on right hand, right foot, left hand and left foot. This implement use RBF kernels for the experiment.

5. The musical System

To control music, we have used a "stop" and "go" player program With DirectX.

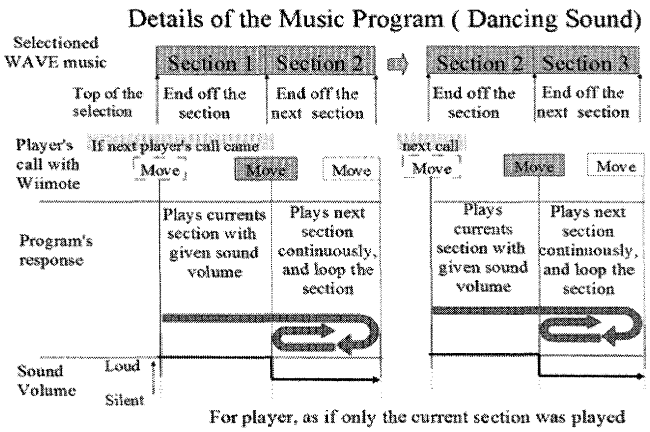

Fig.5 Musical System

The program is based on dividing the soundtrack into several sections equivalents (Fig.5). We have decided to allocate four seconds at each section. When the program receives the signal from the player through the Wii Remote, the first part of the soundtrack is played. The music stops every four seconds and restarts at the next section when the program receives a new signal from the player, and so on until the end of the soundtrack

6. Experiment and Results

To assess the performance of the proposed motion recognition and music controller system--Dancing Sound, an evaluation of the designated music-based dancing assistant system[12][13] has been carried out. The system is written in a C++ language by using wiiyouself! [9] Library and DirectX objects for sound control. To capture some specific motion such as moon walk, video of the "Thriller (Michael Jackson)" is selected for the experiment.

6.1. Pre-process

Before motion recognition and sound controller is started, motion databases are necessary as training set in SVM classifier. About the selected song of *Thriller*, the time length is 4 minutes and 10 seconds, all streams of sensory data include five features: Euler angles Pitch and Roll and accelerations for three dimensions. Four CSV files are outputted based on the body parts (left or right, hand or foot). In the experiment, database sizes are approximately 220KB, the size of hands motion database are apparently bigger than foot because of the choreograph. Also the video of a master dance motion is taken by a professional dancer (or good at dancing). The database format is designed as shown in Fig.6.

A	B	C	D	E	F	G	H	I
Time	Timelength	Accel.X	Accel.Y	Accel.Z	Pitch	Roll	Update	Class
190916519	24	0	0	0	0	0	1	1
190917519	1024	0	0	1	0	0	0	1
190918519	2024	0	0.042	1	−2.386	0	0	1
190919519	3024	0	0	1	0	0	0	1
190920519	4024	−0.167	−0.958	0.44	63.851	−20.746	0	1

Fig.6 Motion Database

In this table, "Timelength" is calculated in ms, and using four seconds as a designated time section. "Update" label is the value of *UpdateAge*, which is a variation of acceleration data structure in wiiyouself! library, if the value of "Update" is high, the motion is out-of-date and probably shouldn't be used, so the "Update" value is used for sampling point selection. "Class" is the value to discriminate each body part which is used in LIBSVM[11] library.; "1" for left hand, "2" for right hand, "3" for left foot, and "4" for right foot. Fig.7 shows the final curves of acceleration data.

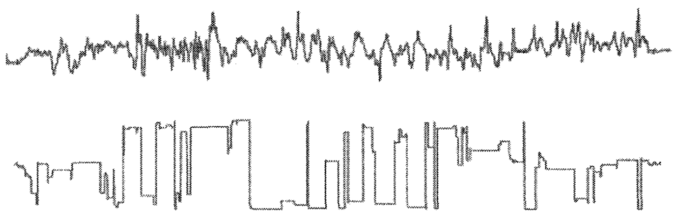

Fig.7 Curve performance of motion database; top: acceleration data, bottom: Euler angle

6.2. Recognition and Training

In our experiment, the hands and feet of two participators (Player-A and Player-B) are equipped with four Wii Remotes with Wii MotionPlus^{TM}. As comparison, Player-B has three years above experiences in dancing, while Player-A had never danced before. The participators were dancing while watching the video made in the pre-process step. In our system, the volume of song is controlled by DirectX to help player to realize the result of motion in each time steps(250/4 ≒ 63 time sections). And the recognition rate is calculated by S/N*100%, where S is the amount of right motions, and N is the total amount of motions (63 here).

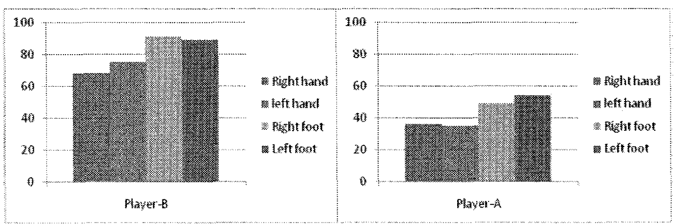

Fig.8 Recognition rates before training

In Fig.8, the recognition rates of hand motions are lower than foot motions, because of the more complex motions designed in hand parts. After all, the dance of *Thriller* is very difficult to study. For training, the info of sound control can help player to make a better performance, in our experiment, both player had tried the system five times. Fig.9 shows the result after training. The recognition rates in both players are improved after training, and there is a more apparent improvement in foot motions, because the foot motions are so simple that easier to study than hand motions.

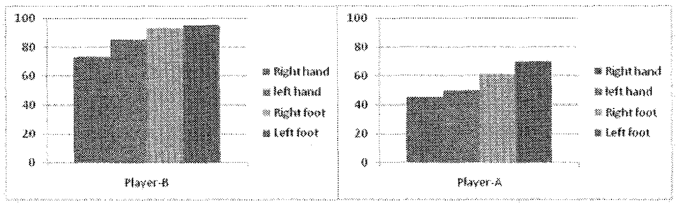

Fig.9 Recognition rate after training

6.3. Limitations and futures Works

Although our objective is to have a good system of gesture recognition, we haven't reached that point yet. However, we have a very good working program which records and saves key Wii Remote data, accelerometer data, infrared position values, pitch, roll, and time. Thus, the player should always stay within that specific region for the infrared position recording accurately and correctly. Furthermore, a better non-flickering strong infrared light source can also help out a lot in our experimental set-up.

7. Conclusion

We've examined the potential musical applications of the Wii Remote, and evaluated its usability in what we believe to be some of the core contexts of its use. This evaluation has shed light on some of the problems that might occur when employing the device, and on situations, such as expressive control, where the controller may yield more creative potential. The Wii Remote is relatively cheap and easily connectable to home computers, making it widely accessible to musicians. The results show it can add interesting and novel dimensions to musical control, provided that some limitations are accounted for. This study doesn't claim to be exhaustive, although we hope that we've covered the core functions of the Wii Remote. In terms of applications for the accelerometer data, more evaluation could be carried out in the area of FFT analysis contexts.

References

[1] Johnny Chung Lee, "Hacking the Nintendo Wii Remote", IEEE Pervasive Computing, 7, 3, pp.39-45, (July.2008)

[2] http://johnnylee.net/projects/wii/

[3] 村 宏樹, 笹島 大輔, 淡野 公一, 外山 貴子, "wiiリモコンを用いた人物の動作識別システムの構築", 宮崎大学工学部紀要, 37, pp.141-146, (Aug.2008)

[4] Hiroki TAMURA, Koichi TANNO, Masahiro ISHII and Zheng TANG, "The Discrimination Experiments of the Walking Style with Human Emotion using Single Acceleration Sensor", Journal of Japan Society for Fuzzy Theory and Intelligent Informatics, 22, 1, pp.65-72, (2010)

[5] 高田 大輔, 小川 剛史, 清川 清, 竹村 治雄, "身体動作に基づき提示情報を切り替えるコンテキストアウェアなウェアラブル AR システム", ヒューマンインタフェース学会論文誌, 12, 1, pp. 47-56, (Feb.2010)

[6] http://www.ubi.com/UK/Games/Info.aspx?pId=851

[7] Frank G. Hofmann, Peter Heyer, and Günter Hommel. "Velocity profile based recognition of dynamic gestures with discrete hidden markov models", Proceedings of the International Gesture Workshop on Gesture and Sign Language in Human-Computer Interaction, pp.81–95, (Sep.1997)

[8] Thomas Schlömer, Benjamin Poppinga, Niels Henze, and Susanne Boll. "Gesture recognition with a wii controller". In TEI '08: Proceedings of the 2nd international conference on Tangible and embedded interaction, pp. 11–14, (Feb.2008)

[9] Wiiyourself! library. http://wiiyourself.gl.tter.org/, 2008

[10] http://www.fftw.org/

[11] http://www.csie.ntu.edu.tw/~cjlin/libsvm/

[12] A. Camurri, G. Castellano, R. Cowie, D. Glowinski, B. Knapp, C. L. Krumhansl, O. Villon, and G. Volpe. "The premio paganini project: a multimodal gesture-based approach for explaining emotional processes in music performance", Proceedings of the 7th International Workshop on Gesture in Human-Computer Interaction and Simulation. Lisbon, (May.2007)

[13] E. Maestre, J. Janer, M. Blaauw, A. P´erez, and E. Guaus. "Acquisition of violin instrumental gestures using a commercial EMF tracking device". In Proceedings of the 2007 International Computer Music Conference, Copenhagen, 2007