Proc. of the 13th IEEE International Symposium on Wearable Computers (ISWC '09), pp. 9--16 (Sep. 2009).

A Motion Recognition Method for a Wearable Dancing Musical Instrument

Minoru Fujimoto*, Naotaka Fujita*, Yoshinari Takegawa†, Tsutomu Terada* and Masahiko Tsukamoto*
*Graduate School of Engineering, Kobe University, Japan
Email: {minoru, nfujita}@stu.kobe-u.ac.jp, tsutomu@eedept.kobe-u.ac.jp, tuka@kobe-u.ac.jp
†Organization of Advanced Science and Technology, Kobe University, Japan
Email: take@eedept.kobe-u.ac.jp

Abstract

In this paper, we constructed a system for realizing a new style of dance performance that dancers play music by dancing. From pilot study, we have found that the motion recognition for dance performance needed the synchronism to back ground music (BGM). Therefore, we propose a new motion recognition method specialized to dance performances. The key techniques of the proposed method are (1) adaptive decision of the size of recognition window to recognize a motion in sync with BGM, and (2) motion recognition in two-phase (rough and detailed) to fulfill the accuracy in high speed recognition. Data was recorded using a 3-axis wireless accelerometers mounted on both shoes. We evaluated the method on a dataset of 5 different dance steps (each repeated 100 times). The results show that this method is capable of improving recognition for all steps (in one case improving recognition from 62% to 99%) while retaining a feeling of seamless connection between movement and sound.

1. Introduction

Advances in computer technologies have enabled new means of musical expression with motion-enabled musical instruments. These instruments enable users to generate sound through physical movements, and they have created a new style of entertainment that enables physical and musical performances to be integrated. Although a lot of researchers and performing artists use such instruments, they do not fulfill the requirement that users want to keep the procedure for the whole performance under their own control.

In street dance, where dancers make steps according to a constant rhythm, allocating dance steps to sounds is effective approach. We created a system of dance performance that enables dancers to generate music while dancing, by using our sensor-equipped shoes. Our system not only generates sounds by dancing, but it also controls the scenario for the performance by changing the sets of sounds generated by motion commands. Dancers gain new dancing and musical expressiveness by using our system in dance battles. If one dancer is playing a drum-track by dancing, another is playing the bass line, and another is playing conga beats, i.e., they are playing music like a band.

We used an initial prototype on stage, and we realized that the relations between dance steps and music were important. However, since the movement was not synchronized with the sound exactly in the prototype, dancers did not gain the sense that they output the sound by themselves. Therefore, we propose a new motion recognition method specialized to dancers. The key techniques are the motion recognition considering the beat of back ground music(BGM) to decide the recognition window adaptively and the two-phase motion recognition to fulfill the accuracy in high speed recognition. The evaluation results confirmed that the system could recognize movements correctly by checking on the rest of the motion when the system made errors in recognition in the first part of motion, and it makes dancers feel a sense of satisfaction to be able to render music by their own physical movements.

The remainder of this paper is organized as follows. The next section explains related work. The subsequent sections describe our wearable dancing musical instrument system, sound sensitivity experiments, and improvements to the system to solve problems. We then present the results from an evaluation of our proposed recognition techniques. Finally, we present our conclusions and work we plan to do in the future.

2. Related Work

Several research projects and commercial products have appeared since the 1990s that have applied information on body motion and physiology to create music. These have employed various devices to extract information from the human body such as stripe potential sensors[1], bend sensors[2] between joints, and acceleration sensors[3]. There are also systems that use image processing[4].

A major commercial instrument, which senses information from the human body, is “Miburi”[5] by YAMAHA. The user wears sensors that are attached to a grip operated by both hands. The user plays sounds by using gestures while pushing a button. This system enables the user to play music by motions. However, players may not feel that they dance freely and emotionally but they are just forming their body to make intended sound. One of the reasons for this feeling is that these methods allocate the output sound directly to the sensor data while our system uses steps to make sounds.

An example of the use of acceleration sensors is the sensor-equipped shoes[6] by Joseph Paradiso's group at MIT. In this research, the output sound from the system is MIDI information obtained by acquiring 16 kinds of parameters from various sensors. However, this research does not consider the time limit to recognize motions while dancers usually make dance to BGM.

They can control music and compose dance through their movements. However, this technique is inadequate for musical performances because the beat of BGM or the timing of sound and motion have not been considered in its method of recognition.

3. Wearable Dancing Musical Instrument

We developed a system that played music according to dance steps in our previous work[7]. Street dancing includes various genres such as a break, lock, and hip-hop dancing. The dance movements in genres differ greatly, and combining them allows for an infinite varieties of expressiveness. Street dancing is highly innovative, and dancers combine short steps of one or two beats in synchronization to the music. Dancers feel a sense of satisfaction when the music and their movements are in synchronization.

3.1. Requirements

Dancing to music is done passively in conventional performances. Dancers or choreographers statically organize the choreography taking into consideration the flow and tempo of the music. Dancing configurations are static and it is difficult to change them in response to happening problems and the atmosphere. Up to now, dancers had only expressed themselves with bodily movements. However, if dancers were able to generate or control sound with their movements, a new style of performance would emerge. Our pilot study enabled us to clarify these system requirements:

- **No restrictions:** The equipment had to be compact not to restrict movements to maintain a high quality of dance.
- **Flexible movements:** To play/control music with movements, the system should allow dancers to flexibly configure the relationship between dance steps and phrasing.
- **Personalization:** Since movements vary among individuals, the mechanism for detecting motion had to be personalized to each dancer.
- **Flexible configurations:** A dance showcase is a kind of story. The system had to enable dancers to change sets of movements and allocated sounds flexibly.

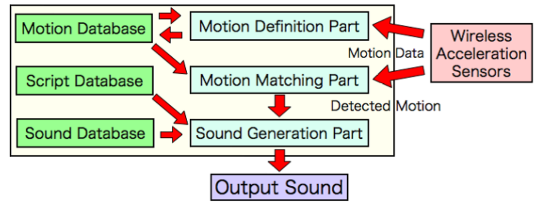

Figure 1. System structure

3.2. System Structure

Figure 1 outlines the flow of our system. It consists of 3-axis wireless acceleration sensors installed in both shoes and a PC application for motion recognition and generating sounds. Because dancers only have to wear small sensors on their shoes, the system does not interfere with their dancing (*No restrictions*). Our system uses feet movements to recognize the dancers' actions. In our system, dancers dance to the music of constant rhythm (we call it BGM). The system selects one music for BGM, and dancers can add short (one or two beats) sounds to BGM as sound effects by making a step.

When dancers first use our system, they register their movements in it. To register movements, a dancer carries out the actual motion while wearing the acceleration sensors on his/her shoes. This creates *flexible movements* and *personalization* because the user can freely register any motion. The dancer then defines the scripts. A script includes various definitions of mapping between movements and commands, such as generating sounds, changing music, and changing mapping sets (the details on functions in a script are described in the next section). It contributes to the *flexible configuration*.

When used on stage, the worn sensors continuously transmit movement data to the PC application via wireless communications. The system recognizes the movements by comparing acquired-motion and stored-motion data from the *motion database*. Our system uses a Dynamic Time Warping (DTW) based on dynamic programming[8] to measure similarity of two time-series data. DTW can be used to measure the similarity between two sequences, each of which may be differently stretched or compressed in time.

Brief algorithm is as follows. Comparing two time series data $X = (x_1, \cdots, x_n)$ and $Y = (y_1, \cdots, y_m)$, an $n \times m$ matrix d is defined by $d(x_i, y_j) = (x_i - y_j)^2$. Subsequently, it finds warping path $W = (w_1, \cdots, w_k)$, which is a path of pairs of indices of X and Y, meeting the following three conditions.

- $w_1 = (1, 1), w_k = (n, m)$
- $w_k = (a, b), w_{k-1} = (a', b') \Rightarrow a - a' \leq 1 \wedge b - b' \leq 1$
- $w_k = (a, b), w_{k-1} = (a', b') \Rightarrow a - a' \geq 0 \wedge b - b' \geq 0$

So as to find the path with the lowest cost with meeting the

Table 1. Commands can be used in script

Name	Explanation
play	specification of the sound to play
start-stop Loop	specification of the loop sound to start/stop
reset all Loop	reset of all loop sounds
change	changing a group
change BGM to	changing BGM to other sound
stop BGM	stop BGM
after	specification of the action of next steps
$\star$(multiply)	specification of the sequential number of steps

above conditions, the following equation is applied.

$$DTW(i,j) = d(x_i, y_j) + min \begin{cases} DTW(i-1, j-1) \\ DTW(i-1, j) \\ DTW(i, j-1) \end{cases}$$

The obtained lowest cost becomes a distance between X and Y. Each sensor outputs three-dimensional acceleration data, the system calculates the similarity between the present sensor data and stored motion data for each dimension. We simply add the costs for all dimensions' data. Each step has own threshold as described in Section 5.3. The system selects one step that has smallest cost. If all steps exceed their threshold, the system detects no step.

3.3. Script for Dancing

To achieve flexibility, we use a scripting mechanism to enable dancers to change the system settings according to their preferences. We implemented a GUI that enabled dancers to intuitively input scripts.

Our script consists of a set of commands, which are listed in Table 1. Figure 2 shows an example script description. The numbers in parentheses correspond to the following functions:

(1) The system outputs *Music1* for two beats when *Sample1* motion is recognized from the right foot.

(2) The system starts/stops the *Music2*-loop sound source when *Sample1* motion is recognized from the left foot.

(3) The system changes the group from *block1* to *block2* when *Sample2* motion is recognized from both feet.

(4) The system outputs *Music4* for one beat when *Sample2* motion is recognized from the right foot sequentially after *Sample2* has been recognized from the left foot.

A user can register multiple scripts as a group. Therefore, the system can change BGM and the combination of dance steps and sounds for every group. For example, dancers play the drum track by using all steps in group 1; in contrast, dancers play the bass line by using all steps in group 2. The system can change the group by performing steps corresponding to conditions to change the group like script description (3) in Fig. 2. We prepare the three types of description for configuring the group:

(a) The system sets BGM in the group to *Music7*.

(b) The system stops BGM.

(c) The system stops all loop sound except the BGM.

Using a script description, the system can not only output sound corresponding to movement but also control the BGM and loop sound sources. Therefore, dancers can control all the music, and define the relations between music and dance as if dancers were creating music with music-production software.

[Group 1]
BGM is *Music7*...(a)
Right *Sample1* play *Music1 2beats*...(1)
Left *Sample1* start-stop Loop *Music2* ...(2)
Both *Sample2* change *group1-2*...(3)
[Group 2]
BGM is *Music6*
Right *Sample1* change BGM to *Music3*
Right *Sample2* after Left *Sample2* play *Music4 1beat*...(4)
Both *Sample2* change *group2-3*
[Group 3]
stop BGM...(b)
reset all Loop...(c)
Both *Sample1* change *group3-1*

Figure 2. An example of script description

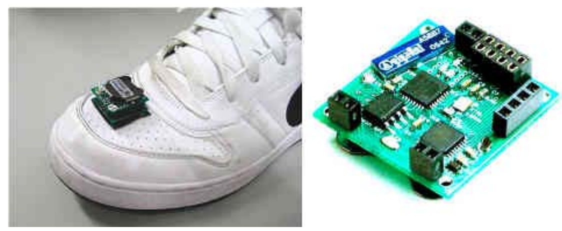

Figure 3. Acceleration sensor on a shoe

3.4. Initial Prototype

We implemented the prototype of our proposed system. Our research group developed a new wireless sensor module called the "Nao-RF Chip". This module has a 100-Hz 3-dimensional acceleration sensor, and its compact size and low weight allow it to be attached to shoes. Figure 3 shows a photograph of the sensor module and a shoe with the module attached. All software was implemented using Apple Xcode3.0 Objective-C2.0. We prepared 30 kinds of Audio Interchange File Format (AIFF) sound sources and also registered 30 movements.

We have used our prototype in several actual dance performances. Dancers were able to change musical patterns while dancing by using our system. Moreover, because we had added a function to change the projected background color in response to movements, the audience became more interested in the performance. However, there were two serious problems:

The difference in motion between that on the actual stage and that during practice

The accuracy of recognition was not as high as that in the practice environments because dancers tended to move intensely, roughly, and freely on the actual stage. Because the system did not often output correct sounds, dancers repeated the same step and strained themselves. As a result, the quality of the performance deteriorated.

The difficulty of outputting rhythmic sound

The system output sound when a movement had been recognized. However, audiences found it difficult to understand that dancers controlled all sounds. Moreover, because the start of output sound was not synchronized with the motion by the difference of recognition delay for each motion, dancers did not feel the sense of togetherness they should have been. We should improve the mechanism responsible for motion recognition to specialize it for dancing.

4. Experiments on Sensibility to Motion and Sound

As described in Section 3, it is a serious problem if sound is output with some delay after motion. However, the more accurate we make recognition techniques, the more possible it becomes for sound to lag behind motion. To find the permissible range of delay, we did three pilot experiments (Experiments 1–3).

The subjects in each experiment were 20 college students (10 dancers and 10 audience members). The dancers had an average of five years of dance experience. They danced being aware that sound was being output by their movements, and the audience subjectively watched the performance. We investigated the sense of incongruity they felt between the dancers' movements and the sound that was output.

These experiments were conducted in a room with a volume like showcase using speakers. We used Apple's GarageBand to make the sounds, and created a BGM *Club Dance Beat*, i.e., an AIFF-loop sound source. The tempo of all sounds was 120 beats per minute (bpm). Subjects answered questions about their performances with 1(worst) to 5(best). The order of the experiments for all subjects was random.

4.1. Experiment 1: Permissible Range of Delay in the Output Sound

In making sound by recognizing dance steps, each step matches the beat for BGM and the output sound is also

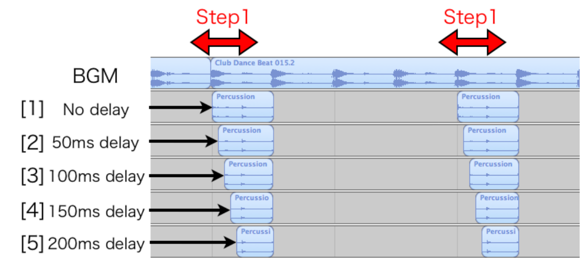

Figure 4. Changing delay of sound output for Experiment 1

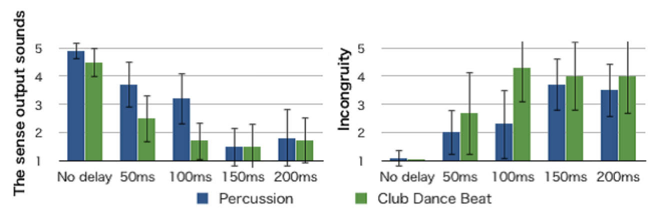

Figure 5. Feeling of synchronization vs. delay of sound output for dancers(left) and Incongruity vs. delay of sound output for and audience(right)

adjusted to the timing of them. However, this makes it difficult for the system to recognize movements because there is too little time to recognize the steps. Therefore, we evaluated how long the permissible range was when output lagged behind the beat of BGM when sounds were output by motion. This was carried out to find how long the system could take time to recognize the steps.

The sound sources for the output sound were *Percussion* and *Club Dance Beat*. As shown in Fig. 4, sounds were created under five conditions: (1) No delay (2) 50ms delay (3) 100ms delay (4) 150ms delay (5) 200ms delay.
First, we demonstrated steps four times for all sound sources according to BGM after we had told dancers the correct timing for the output sounds. The dancers then repeated the same steps that we had demonstrated. The audience watched the dancers' performance without giving them any explanation about the timing of the output sounds. We assessed whether dancers could feel that they had produced the sounds themselves, and whether the audience gained any sense of incongruity between the movements and the sounds.

Figure 5 has barcharts that shows the average results for this experiment. These results mean that the delay time has strong effect on the feeling on both dancers and audience, especially over 100ms delay both feel incongruity between dance motions and output sounds. The difference between two types of sounds is not so much effect on feeling but Club Dance Beat is more sensitive of the delay time. Therefore, we need to adjust the output sounds to the timing of the beat for the BGM, and it means that the system needs to

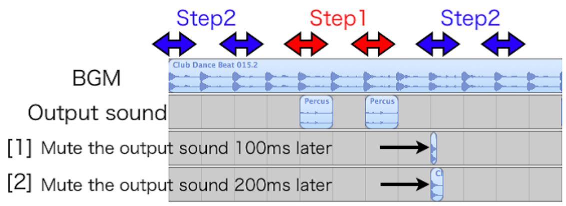

Figure 6. Adding misrecognition for Experiment 2

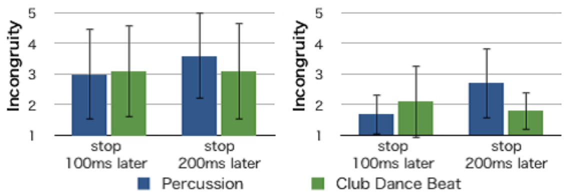

Figure 7. Incongruity vs. length of incorrect sound-output for dancers(left) and for audience(right)

recognize motions before the beat timing.

4.2. Experiment 2: Sense of Incongruity in Muting Output Sound During Performance

Where extra time is needed to recognize motion beyond the permissible range, the system only has to recognize movements temporally in time to the beat. In this case, incorrect recognition can occur because there is not enough time for accurate recognition. We assessed whether people felt any sense of incongruity when the system recognized movements incorrectly and muted the wrong output sounds after it had output the sounds.

We provided an explanation about two kinds of steps to dancers. Step 1 outputs *Percussion* sounds and Step 2 outputs no sound. The dancers performed these two steps one after the other. In this assumption, we adjusted the system to output short sound at random in Step 2 for simulating recognition errors as shown in Fig. 6. The dancers made steps 12 times including two recognition errors. We assessed if they felt any sense of incongruity in their performance. We used two kinds of output sounds of Experiment 1, and the system output sound adjusting it to the timing of the beats of the BGM, i.e., (1) the output sound was muted 100 ms later (2) it was muted 200 ms later.

The audience watched the dance performance without any explanation and we assessed whether they felt any sense of incongruity with the dancers' performance.

Figure 7 has barcharts that shows the average results for this experiment. Regardless of the difference in the sounds, dancers who played music felt a constant sense of incongruity. However, audience members felt practically no sense of incongruity to short sound outputs. Where the sound was a successive loop like *Percussion*, they did not feel any sense of incongruity if the sound was short. However, where the volume of the first sound of the beat was loud as in the *Club Dance Beat*, there was no relationship with the length of the sound if they heard the first part of the sound.

These results indicate that dancers felt a sense of incongruity when they heard unexpected sound while audience members who did not know which movements had created the sound did not feel any sense of incongruity. From the results of actual use of initial prototype on stage, it is important for dancers to feel the sense of togetherness between motions and sounds by performance. Therefore, a method of muting incorrect sound is not effective for dancers to play performance as they want.

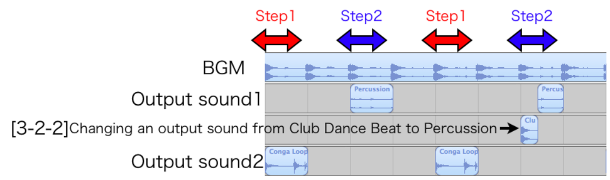

Figure 8. Changing incorrect output to correct sound for Experiment 3

Table 2. Patterns of changing sounds for Experiment 3

Number	Explanation
3-1-1	Change an output sound from *Percussion2* to *Percussion* 100ms later
3-1-2	Change an output sound from *Percussion2* to *Percussion* 200ms later
3-2-1	Change an output sound from *Club Dance Beat* to *Percussion* 100ms later
3-2-2	Change an output sound from *Club Dance Beat* to *Percussion* 200ms later
3-3-1	Change an output sound from *Percussion* to *Club Dance Beat* 100ms later
3-3-2	Change an output sound from *Percussion* to *Club Dance Beat* 200ms later
3-4-1	Change an output sound from *Club Dance Beat2* to *Club Dance Beat* 100ms later
3-4-2	Change an output sound from *Club Dance Beat2* to *Club Dance Beat* 200ms later

4.3. Experiment 3: Sense of Incongruity in Changing Sound Output During Performance

In the same situation as Experiment 2, the system changed the sound from incorrect one to correct one. We told the dancers different sounds could be made by using two kinds of motions and dancers did two kinds of steps one after the other. Actually, as Fig. 8 shows, we adjusted the system to change the output sound only once at a time at random while dancers repeated the steps eight times, and assessed if they felt any sense of incongruity during their performance.

We used the output sounds of Experiment 1 and the system output sound adjusting the timing to the beat of the

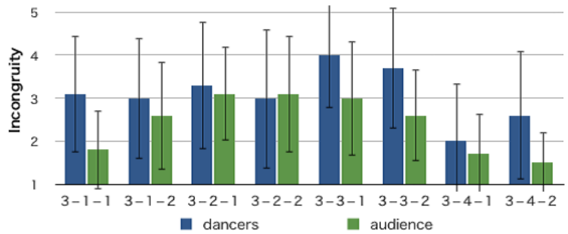

Figure 9. Incongruity vs. patterns of changing sounds

BGM, and did eight kinds of patterns that are listed in Table 2. The audience watched the dance performance without any explanation and we assessed whether they felt any sense of incongruity in the dancers' performance.

Figure 9 shows the average results for this experiment. In changing the sound from *Percussion* to *Club Dance Beat* as in (3-3-1) or (3-3-2), both dancers and audience members felt a sense of incongruity because the connection of two kinds of sounds made the strange sound. In changing the sound from *Club Dance Beat* to similar *Club Dance Beat* as in (3-4-1) or (3-4-2), (the volume or tone were similar,) neither dancers nor audience members seemed to feel any sense of incongruity. This means it is more effective to change sound than mute it during a performance. Therefore, the method of changing sound when the incorrect sound are output is effective if we define optimal conditions for the sound connections.

5. System Design

The importance of the output timing of sound was clarified in previous sections. The steps of the dancer and the beat of BGM should match to combine movements and sound to enable music to be played by dancing. Furthermore, from the results of Sound Experiment 1, the output sounds should start on the beat of BGM.

To satisfy this requirement, we propose a motion-recognition technique that takes into consideration the beat of BGM and a new two-phase recognition method. The system matches the output sound with movement by taking into consideration the beat of BGM. However, the sensing data for recognition is limited by only using this technique, and the accuracy of recognition is not so high. Therefore, the system applies the former technique but also the two-phase motion recognition. In addition, we propose a function for adjusting the threshold of motion recognition to resolve the problem of the difference in movements between that on the actual stage and practice.

5.1. Motion Recognition Considering the Beat

From the result of Experiment 1, the timing of sound generation by dancer's step should be adjusted to the beat of

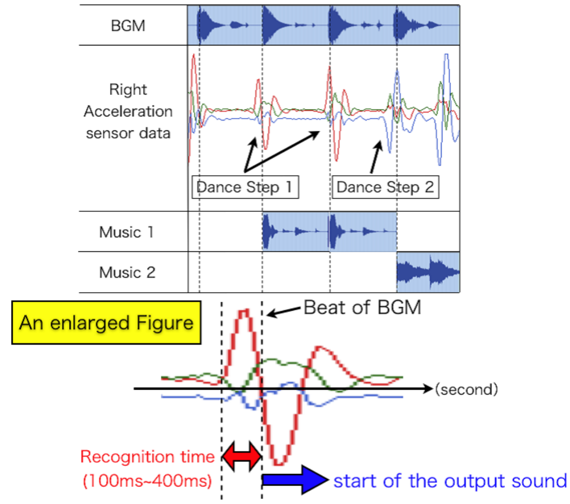

Figure 10. Motion-recognition technique considering the beat of BGM

BGM. This is because dancers step to just timing of the beat of BGM, and they feel good when a dance step synchronizes with it. Generally, since one motion (a wave form sample of acceleration sensor data) does not end at the just timing to the beat, we have to know that a part of wave form to be used for recognition as shown in Figure 10. Therefore, we implemented a function that automatically calculates the range of recognition by stepping a sample to BGM several times actually. The procedure of defining the recognition time for each sample is as follows:

(1) The dancer makes a step to BGM.
(2) The system defines the recognition time for the step as the time from the start of movement to the just timing of the beat.
(3) The dancer repeats the above procedure several times and the system employs the average of the recognition time for these steps.

As we can see in Fig. 10, the system determines an effective range of recognition by applying this mode to individual steps. Using this range of recognition for each step, the system can output the sound matched to the beat of BGM.

5.2. Two-phase motion recognition

The system solved the problem with the output timing of sound by using the motion-recognition technique that took into consideration the beat of BGM. However, it makes the recognition time only 100 – 400 ms. As described in Section II, recognizing the motion of a dance step is not very accurate, and recognition errors increase.

Therefore, from the results of Experiment 3, we used the technique to change the sound when it was not correct but it had the good connection to the correct sound. We propose

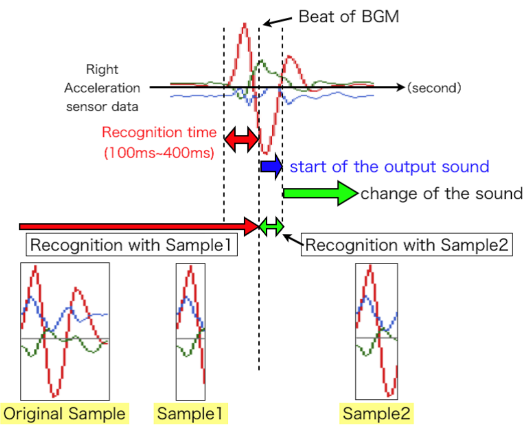

Figure 11. Motion-recognition technique by two-phase

a two-phase motion-recognition method, which recognizes movement in 2 phases and changes the output sound according to the results at each step. This technique maintains a high recognition rate and the output timing to the beat of BGM. Figure 11 shows the proposed method.

The technique that takes into consideration the beat of BGM uses the former part of the sample for recognition. In the same way, the two-phase technique uses the former part of the sample (Sample 1) for recognition, and outputs a sound. After the sound is output, the system recognizes it by using longer sample (Sample 2) than Sample 1 during 50 – 200 ms. If the recognition results for Sample 2 differ from those of Sample 1, the system changes the output sound to the correct (Sample 2) sound.

5.3. Threshold Adjustment

Our system recognizes motion by DTW. We set the different threshold for each sample, and if the similarity between the sample and acquired sensor data is smaller than the threshold, the system detects the motion happens. Generally, if the system makes the threshold high, the false positive decreases but the false negative increases. Moreover, the change of threshold for a motion has effects on the false negative and false positive for other motions. Therefore, the system calculates the false negative and false positive for all motions using actual data of motions each time the user changes one parameter. The user decides the optimal setting that fulfills the user requirements as the following steps:

1) Dancers actually step all samples several times in turn.
2) The system calculates the recognition rate, false-negative, and false-positive for all samples.
3) The system recalculates them by changing the threshold of each sample until fulfilling the user requirements.

The user sets the requirements for each sample with an appropriate value such as the intended recognition rate. The

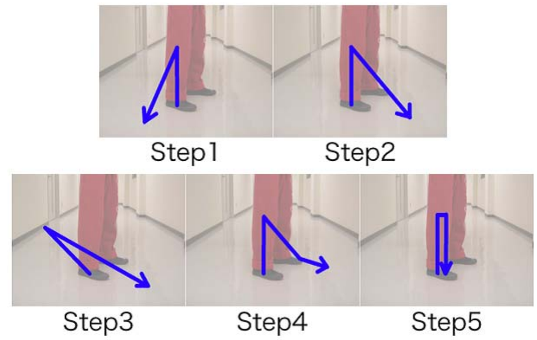

Figure 12. Five dance steps for experiment

system uses an appropriate threshold with this method in the evaluation discussed in the following section.

6. Evaluation

To evaluate our proposed method, we prepared five kinds of steps as outlined in Fig. 12. A total of 100 steps were recorded for each of the 5 dance steps (comprising 10 sets of 10 consecutive repetitions). The steps were similar at the beginning of motion. We assessed the recognition rate in three ways as follows:

1) Real-time method: Recognition only from the first part of the sample.
2) Accurate method: Recognition only from the first part plus 100 ms of the sample.
3) Proposed method

Because there is a need to change sound in our method, we used the sound that was easy to change. The sound in this experiment was changed within 100 ms after the first part of the sample was played. The first author whose had seven years of dance experience participated in this experiment.

Table 3 lists the results of the experiment. In this table, the numbers means the number of times, 'Error recognition' means the system recognized an incorrect step, and 'Non-recognition' means the system recognized no step. In proposed method, 'Correct substitute' means the system substituted the correct sound due to recognition of a correct step at the second phase, and ' False substitute' means the system substituted an incorrect sound by misrecognition at the second phase.

In Real-time method, there are many errors in recognition because the time to recognize was too short. However, as sound could be output along with motion, we felt as though we had created the music ourselves. In Accurate method, as Table 3 shows, fewer errors are recognized than in Real-time method. However, we can easily see that sound was delayed by motion.

In Proposed method, fewer errors were recognized than in Real-time and Accurate method. When movements were

Table 3. Recognition accuracy in three methods

	Real-time method					Accurate method					Proposed method				
Dance Step	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Recognition	99	62	100	96	100	100	98	100	98	99	99	62	100	96	99
Error recognition	0	38	0	3	0	0	2	0	0	1	0	0	0	1	0
Non-recognition	1	0	0	1	0	0	0	0	2	0	1	0	0	1	0
Correct Substitute											0	37	0	2	0
False substitute											0	1	0	0	1

similar, if the system made errors in recognition in the first part of motion, the system could recognize movements correctly by checking on the rest of the motion. Although error recognition made it difficult for dancers to play music as they wanted, the technique we propose is effective when sounds integrate seamlessly. The current system always substituted incorrect sound 100 ms after sound was output. However, good timing to substitute sounds should be varied depending on when sounds are synchronized. Therefore, the system became more effective by setting the time in second phase for all combinations of sound sources and incorporating these as metadata. In the result of step 2, for example, 'recognition' is 62 times in Real-time method, and 'recognition' is 62 times and 'correct substitute' is 37 times in Proposed method. This means the accuracy of this step improved from 62% to 99% by our method. This result confirmed that the proposed method has almost same accuracy as the accurate method in high speed as the real-time method.

There were also a few incorrect modifications. The system could hardly recognize errors in the last part after it had accurately recognized errors in the first part, and one cause of this was the difficulty of setting a proper threshold. As the beginning of the steps were exactly the same, the system could not distinguish them.

We used our system on a dance performance at the Kobe Luminarie Live Stage on December 13th and 14th of 2008. By applying the motion-recognition technique that took into consideration the beat of BGM, sound was output that could be adjusted to the timing of movements. A performance where various sounds were output by various steps created excitement in the audience. Not only skillful dancing but also sensibility to sound was important for dancers because they played music by using this system. We intend to assess the degree of comfort generated by this combination of motion and sound. We did not implement the two-phase motion-recognition technique in at that stage. We intend to evaluate our system in practical use in the future.

7. Conclusion

We described a new system that enables dancers to make music by dancing. Our system uses a scripting mechanism to flexibly control performances. Dancers express their emotions intuitively and create a new type of entertainment using our system. We focused attention on the sense of music being played by dancing. We propose new techniques of recognition specialized to enabling adjustments to be made in the recognition timing of movements. Using our proposed techniques, all dance steps correspond to sounds, resulting in greater plasticity during the performance.

Acknowledgment

This research was supported in part by a Grant-in-Aid for Scientific Research(A)(20240009) and Priority Areas(21013034) from the Japanese Ministry of Education, Culture, Sports, Science and Technology and by Hayao Nakayama Foundation for Science and Technology and Culture.

References

[1] A. Tanaka: Musical Technical Issues in Using Interactive Instrument Technology with Application to the BioMuse, Proceedings of the International Computer Music Conference (ICMC93), pp. 124–126(1993).

[2] W. Siegel, and J. Jacobsen: "Composing for the Digital Dance Interface," Proceedings of the International Computer Music Conference (ICMC99), pp. 276–277(1999).

[3] D. Merrill and J. Paradiso: Personalization, Expressivity, and Learnability of an Implicit Mapping Strategy for Physical Interfaces, In the Extended Abstracts of the Conference on Human Factors in Computing Systems (CHI05), pp. 2152–2161(2005)

[4] F. Sidney and K. Mase: Iamascope: a graphical musical instrument, Computer & Graphics, Vol. 23, pp. 277–286(1999).

[5] YAMAHA[MIBURI], http://www.yamaha.co.jp/design/products/1990/miburi/.

[6] J. Paradiso, K. Hsiao, A. Benbasat, and Z. Teegarden: Design and Implementation of Expressive Footwear, IBM Systems Journal, Vol. 39, No. 3&4. pp. 511–529(2000).

[7] M. Fujimoto, N. Fujita, Y. Takegawa, T. Terada, and M. Tsukamoto: Musical B-boying: a Wearable Musical Instrument by Dancing," Proceedings of the International Conference on Entertainment Computing (ICEC08), pp. 155–160(2008).

[8] R.E. Bellman: Dynamic Programming, Princeton University Press(1957).