Construction and Evaluation of a Robot Dance System

Kuniya Shinozaki[1], Akitsugu Iwatani[2], and Ryohei Nakatsu[3]

[1]Kwansei Gakuin University, School of Science and Technology
2-1 Gakuen, Sanda Japan, 669-1337
scbc0052@kwansei.ac.jp

[2]Universal Studios Japan
Osaka, Japan
Iwatani@nirvana.ne.jp

[3]National University of Singapore, Interactive & Digital Media Institute
Block E3A, #02-04, 7 Engineering Drive 1, Singapore 11774
idmdir@nus.edu.sg

Abstract Dance is one form of entertainment where physical movement is the key factor. The main reason why robots are experiencing a kind of "boom" is that they have a physical body. We propose a robot dance system that combines these two elements. First, various factors concerning entertainment and dance are studied. Then we propose the dance system by robot using motion unit and the synthetic rule referring the speech synthesis. Also we describe the details of the system by focusing on its software functions. Finally we show the evaluation results of robot dance performances.

1 Introduction

The research and development of various kinds of robots is actively being carried out, especially in Japan [1][2][3][4][5]. Several reasons explain the current robot boom. One main reason is that robots have physical bodies, and so human-robot interaction extends beyond human-computer interaction.

Although in the future these robots are expected to support various aspects of our daily life, so far their capabilities are very limited. At present, installing such a task in robots remains very difficult. To break through such a situation, entertainment might be a good application area for robots.

Developing a dancing robot would be remarkable from various points of view. First, it might become a new form of entertainment, activates both the body and brain. Watching humans dance is already one established type of entertainment.

Please use the following format when citing this chapter:

Shinozaki, K., Iwatani, A. and Nakatsu, R., 2008, in IFIP International Federation for Information Processing, Volume 279; *New Frontiers for Entertainment Computing*; Paolo Ciancarini, Ryohei Nakatsu, Matthias Rauterberg, Marco Roccetti; (Boston: Springer), pp. 83–94.

Second, we might develop a new type of communication with computers, because dance can be considered one of the most sophisticated nonverbal communication methods.

Based on the above considerations we started to research dancing robots. In this paper we clarify the relationship among entertainment, humans, and robots and propose a robot dance system by robot using motion unit and the synthetic rule referring the speech synthesis. Also we will describe an evaluation experiment carried out to test this basic concept's feasibility.

2 Dance Entertainment and Robots

2.1 Entertainment

The role of entertainment in our daily life is very important. It offers relaxation and thus contributes to our mental health. Many aspects concerning entertainment must be considered and discussed [6]. One of the most important may be the existence of two sides: entertainer and audience. Although these two sides change positions depending on the case, the existence of performers and spectators is an absolute prerequisite for entertainment. Many entertainments have both entertainer and spectator characteristics. In the case of dance, people sometimes go to theaters to watch good dance performances, and they sometimes go to dance clubs or discos to dance themselves.

Furthermore, when viewed from a different aspect entertainment can be classified into two types. One is a real-time type that includes performers or entertainers performing live in front of an audience. Good examples include plays and/or concerts. Another is the non-real-time type; reading books and watching movies are good examples.

Following this classification, dance basically belongs to the real-time type of entertainment. For robot dancing, however, as described later, its position is somewhat special.

2.2 Dance Robot

One main reason why we choose dance as an entertainment for robots is that dance is quite sophisticated [7]. Based on the considerations described above, what is the role of robots in dance entertainment? Dance robots allow us to become both entertainers and spectators. When watching a robot dance, we are spectators. On the other side, many people will probably want to install dance motions on their robots and show these actions to others. In this case they are entertainers. For the classification between real-time and non-real-time entertainment, dance robots also have significant characteristics. If we want to show people the robot dance, we have to install the dance actions beforehand, meaning that the robot dance is non-real-time entertainment. At the same time, by developing interactive capabilities, the robot would show impromptu dancing behaviors. For example, it could change the dance depending on audience requests. Or it could sense the audience mood and could adopt its dancing behaviors to reflect the sensor results. A dance robot could provide flexible entertainment that ranges between real-time and non-real-time entertainment.

3 Dance Robot System

3.1 Basic Concept

Based on the above considerations we want to develop a system that can generate various dance motions. Since different dance genres exist, it is necessary to restrict dance genres to a specific one. Then the system would generate various dance motions by selecting several basic dance motions and by concatenating them. This basic idea resembles text-to-speech synthesis (TTS) [8], where by restricting the language to be synthesized and by selecting a basic speech unit, any kind of text described by the language can be generated. The following is the basic concept adopted in TTS:

(1) Speech consists of a concatenation of basic speech units.

(2) Selection of the speech unit is crucial.

(3) Connection of speech units is also crucial.

As basic speech units, various basic units such as phonemes, phoneme pairs, CV (consonant-vowel concatenation), CVC, VCV and so on have been studied [8]. Based on research of the last several decades, phonemes including variations that depend on previous and following phonemes are widely used as speech units. Taking these situations into consideration, the basic concept of dance generation is as follows:

(1) We restrict the generated dance to a specific genre.

(2) All dance motions consist of a concatenation of several basic dance motions.

(3) Deciding what to select dance units as basic dance motions is very important.

(4) Connecting dance units is crucial.

(5) Also it is crucial how to express a dance unit as robot motion.

In the following sections, we answer the above questions.

3.2 Dance Genre

For basic dance motions, there are several researches on classic ballet [9]. The classification of ballet motions is based on several leg positions and movements called steps. Although each leg position and step has its own name, basically no rules describe the details of whole body motions. We chose hip-hop as the dance genre because all of its dance steps and motions are classified into several categories, so it is easier to handle the whole body motions of hip-hop than ballet.

3.3 Dance Unit

Next we must decide the basic unit for dance motions. As described above, since each hip-hop step/body motion has its own name, it can be selected as a dance unit. However, it is difficult for an amateur to extract them from continuous dance motions. Therefore we collaborated with a professional dancer to simplify the extraction of basic motions from continuous dance motions. In addition, when constructing robot motions based on human motions, we must deform complicated human motions into rather simple robot motions. In this deformation process, a professional dancer's advice is also of great help.

3.4 Concatenation of Dance Units

The next question is how to connect each motion unit. One method interpolates the last posture of the previous motion and the first posture of the next motion. The difficulty in the case of a dancing robot is how to connect these two motions and prevent the robot from falling down. We introduced a method in which a neutral posture represented by a standing still pose is used as a transition posture between two dance units. In this case developing an algorithm is unnecessary to generate a transitional motion that connects two different motions.

3.5 Realization of Robot Dance Motion

The next issue is transforming human dance motions into the motions of robots. One common method adopts a motion capture system that is used to generate the motion of CG characters. For a robot, however, due to the limitations of the degree of freedom at each joint, directly transforming the motion captured by the system into robot motion does not work well. Research that transforms captured motions into robot motions is described in [10] that treats a Japanese traditional dance whose motions include legs moving slowly and smoothly front/back and left/right instead of dynamically. In this case it is relatively easy to maintain balance. However, hip-hop motions include dynamic body motions, and therefore it is difficult to maintain balance. Taking these situations into considerations, we chose a method where each motion unit extracted from continuous motion is transformed manually.

3.6 System Architecture

Based on the above considerations, we constructed the first prototype of a robot dance system, as shown in Fig. 1, that consists of dance unit sequence generation, a dance unit database, and dance unit concatenation.

(1) Dance unit database

A large amount of dance units are stored here; each one corresponds to a basic short dance motion and is expressed as robot motion data.

(2) Dance unit sequence generation

An input data that expresses a dance motion is analyzed and converted into a sequence of dance units by this part. At the present stage a sequence of dance units is directly used as input data and fed into the system.

(3) Dance unit concatenation

As is described in 3.4, a neutral posture is introduced as an intermediate posture between two dance units, and therefore, they can be easily connected.

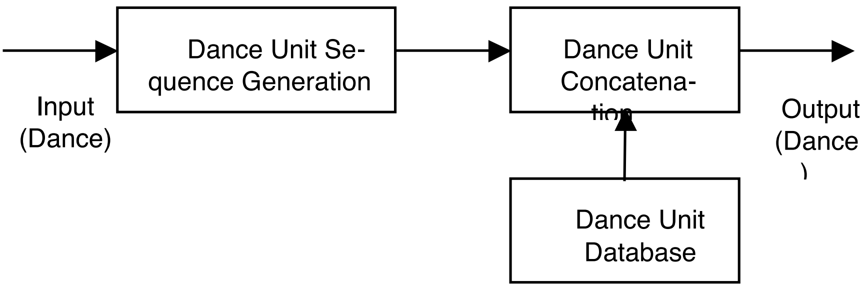


Fig. 1 Structure of dance robot system

4 System Development and Evaluation

4.1 Humanoid Robot

From the several humanoid robots already available on the market, we selected a humanoid robot developed by Nirvana Technology [11] and installed dance motions on it. Figure 2 shows its appearance, and Table 1 shows its basic specifications. Various robot motions can be designed and produced on PC using a "motion editor" realized by motion making and editing software.

Fig. 2 Humanoid robot

Table 1 Specifications of humanoid robot

Size/Weight	34 cm / 1.7 kg
Degree of flexibility	22 (12 legs, 8 arms, 1 waist, 1 head)
CPU	SH2/7047F
Motor	KO PDS-2144, FUTABA S3003, FUTABA S3102, FUTABA S3103
Battery	DC6V

4.2 Development of Dance Unit Database

As described above, we collaborated with a dancer to develop a dance unit database and conducted the following database generation:

(1) First, a typical hip-hop motion of several minutes long was recorded.

(2) Then we observed and discussed the dance sequence and selected about 60 motions as dance units that included almost all the representative hip-hop motions.

(3) We asked the dancer to separately perform each motion corresponding to each dance unit and recorded it. At the same time we asked him to start each dance motion from a “natural standing posture” and to finish in the same posture.

(4) By watching each dance motion being performed, we tried to create a robot dance motion that corresponds to human dance motion using motion editor.

4.3 Evaluation of Robot Dancing

Using the system described above we carried out simple evaluation experiments.

4.3.1 Comparison of the two types of robot dance units

We evaluated the two types of dance units; one was generated by the professional dancer (type 1) and the other by non-experts (type 2). First we classified all the dance motions into three categories according to the complications of the motions; primary, intermediate, and advanced. And we selected one representative motion for each category. These dance motions are "Lock"(primary), "Rolling Arm" (intermediate), and "Club"(advanced). Then we generated two types of robot dance motions for each of these motions.

Ten subjects were asked to compare these two types of robot dance motions by giving a score ranging from 1 to 5 to each dance motion (1 is the worst and 5 is the best). Figure 3 shows the comparison between the two types of dance motions; robot dance motions developed by the dancer himself (type 1) and those developed by non-experts (type 2) for three kinds of motions; (a) Lock, (b) Rolling arm, and (c) Crab. Also the live dance motions performed by the dancer is shown as references. Figure 4 shows the evaluation results for each of the three kinds of motions. The evaluation result and the consideration for each motion are described below.

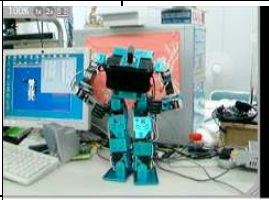
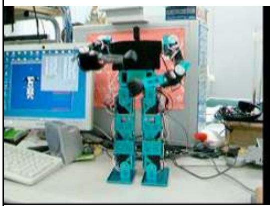
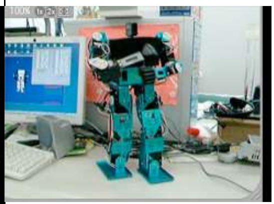
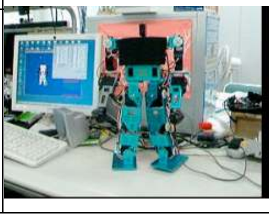
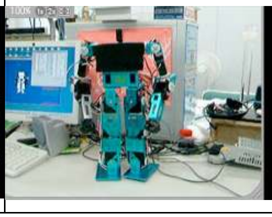
	Robot motion (type 1)	Robot motion (type 2)	Dancer Performance
(a)			
(b)			
(c)			

Fig. 3 Comparison of three dance motions

(1) Lock

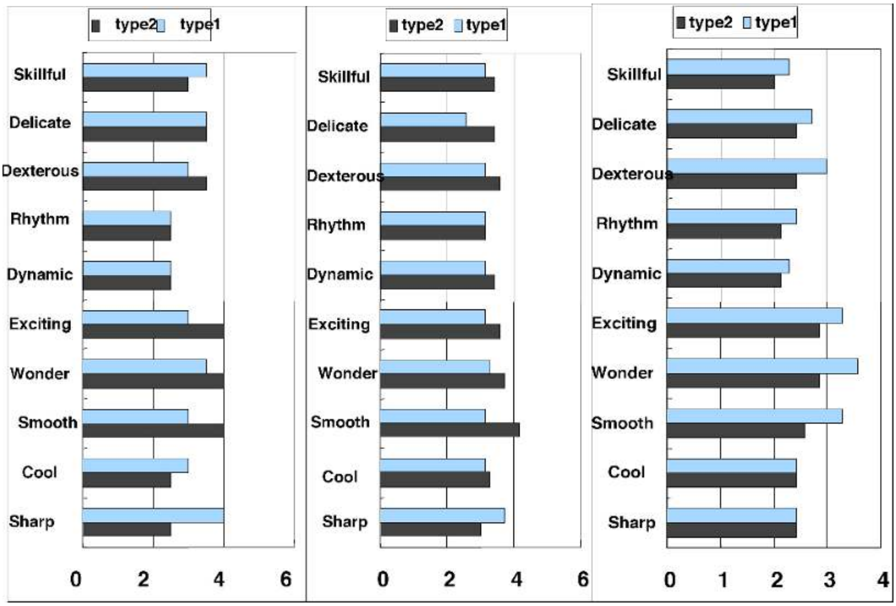
This is a repeating motion of moving and stopping like being locked. In this move the sharpness of stopping motion is an important factor as a dance. For “sharpness,” type 1 motion (motion designed by a professional dancer) obtained the higher score than type 2 (motions designed by non-experts) as expected. On the other hand, for such evaluation items as “exciting,” “wonder,” and “smooth,” the type 2 motion got higher scores than the type 1 motion. It seems that the stop-and-go motion designed by the dancer was judged awkward by the subjects.

(2) Rolling arm

This is a motion of moving body while turning arms smoothly. For the sharpness, the type 1 motion obtained higher score than the type 2. But for other evaluation items, the type 2 motions generally got slightly higher scores. Especially for “smooth” type 2 received much higher scores against type 1. Originally this motion contains a step raising legs, and the type 1 motion precisely simulates this process and in the case of sharpness it worked well and obtained the high score. On the other hand, the type 2 motion achieves this move by sliding legs without raising legs. As a result, it was judged that the type 2 motion looked smoother than the type 1, and this gave a influence to the result of smoothness evaluation and others.

(3) Crab

This motion is a move peculiar to the Hip-hop dance. It includes a move of sliding legs sideways without raising them and fixing their backside on floor and thus moving the body sideways. The motion designed by the professional dancer (type 1) receives higher scores than the motion designed by non-expert (type 2) for almost all evaluation items. Especially, important evaluation items for this move such as “exciting,” “wonder,” and “smooth,” the type 1 obtains fairly higher evaluation scores than the type 2.



(a) (b) (c)

Fig. 4 Evaluation results for three kinds of motions

These result shows that as the robot dance motions become more complex, they can get higher scores. The reason for this would be that the professional dancer understands so well the characteristics of each dance motion and his knowledge and now-how is reflected on the robot dance motion. Even though it does not appear so well in the case of simple motions, this characteristic reveals itself in the case of complicated motions. On the other hand, the motion designed by non-expert (type 2) obtained higher evaluation scores than the type 1 for simple motions. The explanation for this would be that the subjects got good impressions for the over-actions and the unstableness that the type 2 motions generally contain and express themselves. Contrarily, the type 1 motions designed by a professional dancer are sophisticated without containing such over-action nor unstableness. This characteristic sometimes leads to rather low evaluation scores as the subjects are non-expert of dances and thus could not understand the details of the dance motions where the knowledge and now-how of the professional are stored.

4.3.2 Evaluation of the continuous dance motion

Then we carried out the experiment to evaluate the feasibility of the dance generation system. We compared two types of continuous dance motions. One is a continuous dance motion which is automatically generated by this system and has the length of about one minute (type 3). Another is the same dance motion where instead of automatic generation the professional dancer designed the whole continuous dance motion from scratch (type 4).

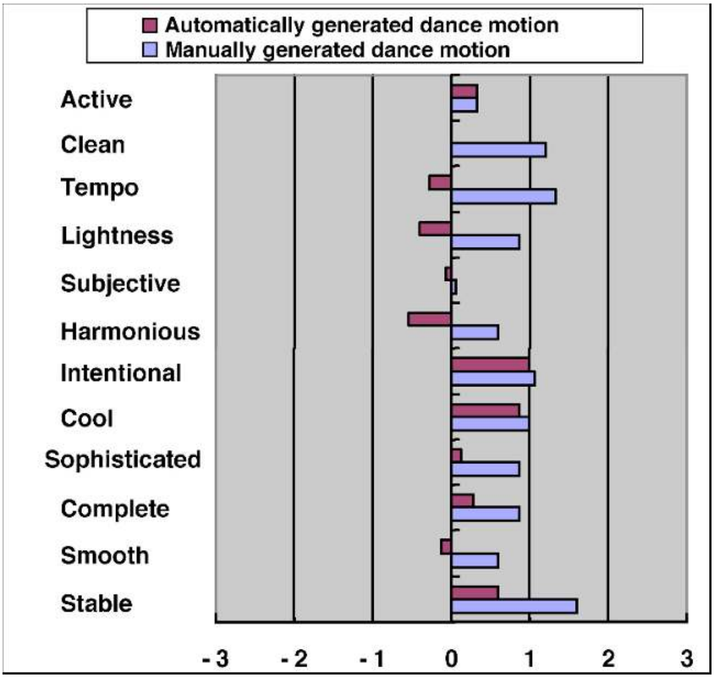


Fig. 5 Comparison between automatically generated motions and manually generated motions

For evaluation twelve items generally used for the sensibility evaluation such as "stable," "soft", “smooth,” and so on were selected. Each evaluation item has a seven level score ranging from -3 to 3. For example, for the evaluation item "stable" the 0 means neutral, 3 means very stable, and -3 is very unstable. Figure 5 shows the evaluation result. The type 4 obtained fairly good results for most of the evaluation items. This means that the evaluation items were fairly well selected. Generally the dance motion generated by this dance generation system (type 3) obtained lower evaluation scores than the type 4 motion. Especially, for such evaluation items as "harmony," "lightness," and "tempo, " the type 3 motion obtained minus evaluation scores. This is because the subject felt unnaturalness due to the neutral posture effect used to connect the two dance units. This means that the system still needs further improvement to generate continuous dance motion, especially for the connection of two dance units. At the same time, however, the type 3 motion got plus scores for "stability", "cool", and "intentional." Especially for "cool" and "intentional" the evaluation results are almost as high as the results of the type 4 motion. This shows that the continuous dance motion generated by this system would be effective as far as it is used as a performance even at the present stage.

The difference between type 3 and type 4 motions are that in the case of type 3 motion it goes back to a neutral position at the point of the dance unit connection. It is necessary to improve this point by introducing better neutral posture or introducing multiple neutral postures.

5 Conclusion

In this paper we proposed a dance robot system as a new application area for humanoid robots. We clarified several distinctive entertainment characteristics and investigated the role of robots in entertainment.

Based on these basic considerations we proposed a dance robot system in which a humanoid robot performs various dance motions. We hypothesized that any dance motion consists of a concatenation of short dance motions called dance units. This basic idea was imported from TTS, where any text can be converted into speech by concatenating short basic speech called speech units. Based on this basic idea, we collaborated with a professional dancer. After recording and analyzing his hip-hop dancing, we extracted about sixty dance units and converted them into the motions of a humanoid robot. By concatenating these dance units we found that a huge amount of dance variations for the hip-hop genre could be achieved.

Then we carried out two types of evaluation experiments. First we compared dance motions designed by the professional dancer and the ones by non-experts of dancing. We found that as the dance motions become more complicated and sophisticated, the dance motions by the dancer got higher evaluation results. Then we compared a continuous dance motion automatically generated by this system and one fully manually designed. Although the automatically generated dance got lower evaluation results, for some evaluation items it got almost the same scores.

This means that this system is promising from a point of automatic dance generation. Further studies must address the following issues. First we have to investigate how many dance units are enough to generate any type of hip-hop dance. Also we have to investigate the feasibility of a neutral posture that connects two dance units. As only one type of neutral posture was used so far, still there is some unnaturalness for the automatically generated continuous dance motion. We expect that by introducing several other neutral postures, continuous dance motions achieved by the robot would become more natural.

References

1. Golubovic, D., Li, B., and Hu, H. A Hybrid Software Platform for Sony AIBO Robots. (2003). *RoboCup 2003: Robot Soccer World Cup* VII, 478-486.
2. Ishimura, T., Kato, T., Oda, T., and Ohashi, T. An Open Robot Simulator Environment. (2003). *RoboCup 2003: Robot Soccer World Cup* VII, 621-627
3. Kerepesi, A., Kubinyi, E., Jonsson, G. K., Magnusson, M. S., and Kiklosi, A. Behavioural Comparison of Human-Animal (Dog) and Human-Robot (AIBO) Interactions. (2006). *Behavioural Processes*, Vol. 73, No.1, 92-99.
4. Wama, T., Higuchi, M., Sakamoto, H., and Nakatsu, R. Realization of Tai-chi Motion Using a Humanoid Robot. (2004). *Entertainment Computing*, Springer LNCS, 14-19.
5. http://www.expo2005.or.jp/en/index.html
6. Callois, R. *Les Jeux et les Hommes*. (1958). Paris: Callimard.
7. Laban, R. *The Mastery of Movement*. (1980). Macdonald and Evans; 4th ed., revised and enlarged edition.
8. Kleijn, W. B. and Paliwal, K. K. (ed.), *Speech Coding and Synthesis*. (1995). Elsevier.
9. Lee, C. *Ballet in Western Culture: A History of Its Origins and Evolution*. (2002). Routledge, London.
10. Nakaoka, S., Nakazawa, A., Yokoi, K., Hirukawa, H., and Ikeuchi, K. Generating Whole Body Motions for a Biped Humanoid Robot from Captured Human Dances. (2003). *IEEE 2003 International Conference on Robotics and Automation*.
11. http://www.nirvana.ne.jp/