Leg Motion Primitives for a Dancing Humanoid Robot

Shinichiro Nakaoka*, Atsushi Nakazawa†, Kazuhito Yokoi‡ and Katsushi Ikeuchi*
* Institute of Industrial Science, The University of Tokyo
4-6-1 Komaba, Meguro-ku, Tokyo 153-8505 JAPAN
Email: {nakaoka,ki}@cvl.iis.u-tokyo.ac.jp
†Cyber Media Center, Osaka University
1-32 Machikaneyama, Toyonaka, Osaka 560-0043 JAPAN
Email: nakazawa@ime.cmc.osaka-u.ac.jp
‡Intelligent Systems Institute, National Institute of Advanced Industrial Science and Technology
Tsukuba Central 2, 1-1-1 Umezono, Tsukuba, Ibaraki 304-8568 JAPAN
Email: Kazuhito.Yokoi@aist.go.jp

***Abstract*— The goal of the study described in this paper is to develop a total technology for archiving human dance motions. A key feature of this technology is a dance replay by a humanoid robot. Although human dance motions can be acquired by a motion capture system, a robot cannot exactly follow the captured data because of different body structure and physical properties between the human and the robot. In particular, leg motions are too constrained to be converted from the captured data because the legs must interact with the floor and keep dynamic balance within the mechanical constraints of current robots. To solve this problem, we have designed a symbolic description of leg motion primitives in a dance performance. Human dance actions are recognized as a sequence of primitives and the same actions of the robot can be regenerated from them. This framework is more reasonable than modifying the original motion to adapt the robot constraints. We have developed a system to generate feasible robot motions from a human performance, and realized a dance performance by the robot HRP-1S.**

I. Introduction

In our project, we have attempted to develop a total technology to preserve traditional folk dances which have been disappearing for the lack of successors[1]. Currently, our target dances are Japanese folk dances such as *Jongara-Bushi* shown as Fig. 1.

We believe that preserving dances requires replay of the dance performances. A typical solution for replay is representation by CG animation or virtual reality (VR). For example, Yukawa et al.[2] have proposed the 'BUYO-FU' system, which enables composition of choreography and reproduction by CG animation. However, these methods are insufficient because 'watching a dance' is not identical to'watching a CG animation' . Performances by an actual dancer, which bring a stronger impression than CG animation or VR are necessary. In this study, a humanoid robot is used to realize actual performances.

Pollard et al. [3] have developed a method to import human dance motion into a robot. Motion data acquired through a motion capturing system are converted to joint angle trajectories of a robot, and the trajectories are modified to be feasible ones within the constraints of the robot. However, their study does not deal with leg motion. Their robot's body is fixed to a stand. On the other hand, our study attempts to realize whole body performances by a robot, including leg actions. In generating leg motion, using motion data which are directly converted from captured motion is not easy, because the leg motion of a robot is too restricted for the application of such data. Leg structure of present robots easily causes self collision or overrun of movable ranges. In addition to avoiding those situations, leg motion must also consider dynamic balance keeping.

Our attempt is not to develop a robot just as a dance recorder and player, but rather to develop a robot which can recognize a dancer's intentions by observation and then imitate dances on the basis of recognition results. This approach is called 'Learning by Observation (LFO)', and learning frameworks based on LFO for another tasks have been proposed. For example, an assembly task can be modeled by several contact states between parts, and task behavior is expressed as transition of the states [4]. Or a knotting task can be modeled by knot states of a string based on the knot theory [5].

In this study, leg actions of dance motion are modeled by

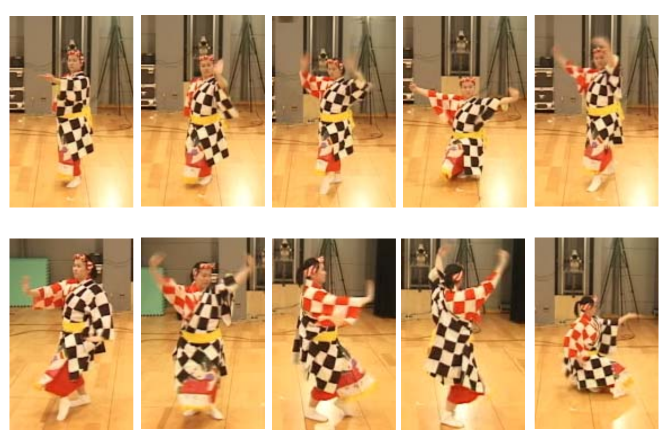

Fig. 1. Jongara-Bushi

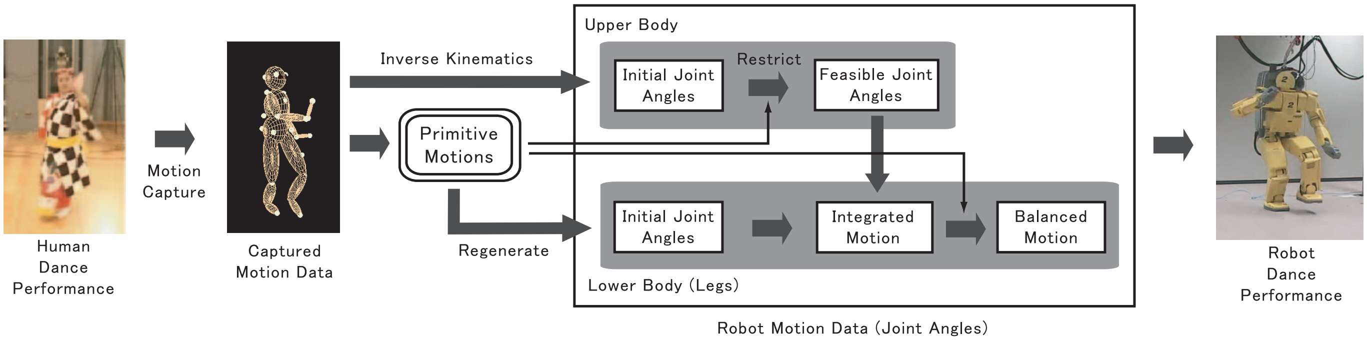

Fig. 2. Overview of the System

several motion primitives. Human dance motion is recognized as a sequence of the primitives and robot motion is generated from the recognized result. This approach enables a performance including leg actions under the severe limitations of robot legs. Also, this approach brings high flexibility to performances such as adaptation to various stage conditions, recomposed choreography sequences, or interaction with humans.

Inamura et al. [6] have proposed a framework of imitation, called *mimesis loop*. Their method can deal with general kinds of motions, but it seems that feasibility on actual robots is not considered sufficiently. On the other hand, our method specializes in dance performance and focuses attention on generating feasible motions of a robot.

In this paper, we propose a framework based on primitives of leg actions. The primitives are first defined and human dance performance is recognized as a sequence of primitives. Then robot motion is regenerated only from information of the primitive sequence. With this framework, feasible motions are simply generated rather than raw motion data being modified to adapt to a robot. In [7], we tested a motion generated by this framework on dynamics simulation. In addition to that, this paper shows the result of experiments on an actual robot hardware.

II. System Overview

A robot motion is converted from a human dance performance through the procedure illustrated in Fig. 2.

First, dance motion is acquired as digital data from a human performance by means of a motion capturing system. We use VICON, an optical type motion capturing system. Time series positions of 30 joint markers are acquired at the rate of 200 frames per second. We captured some Japanese folk dances such as Tsugaru Jongara-Bushi and Aizu Bandaisan.

Since captured motion data cannot be directly imported into a robot, the data must be converted to the joint angle trajectories of a robot. This process mainly consists of two parts: recognizing dance motion from captured motion data, and generating motion data for a robot. By the recognition process, symbolic representation of the dance is extracted. The representation is composed of *primitive motions*, which make up a minimal unit of dance motion.

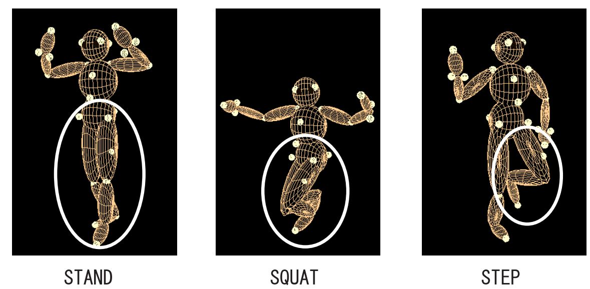

Fig. 3. Motion primitives of legs

In the present system, conversion processes are different between upper body motion and leg motion. Upper body motion is generated by using both raw motion data and extracted primitives [1]. First, joint angles are directly converted from raw motion data by inverse kinematics. Then, the data is modified to adapt to the constraints of a robot. The extracted primitives are used to acquire better expression in this time. On the other hand, leg motion is regenerated only from the information of the extracted primitives. Finally, upper body motion and leg motion are integrated and a waist trajectory is modified to keep the dynamic balance. Detailed process for leg motion is described in the following sections.

As the final step, a robot performs a dance according to the generated joint angle trajectories. We use the HRP robot platform, which has a common control interface between virtual robots for simulation and actual robots[8]. The validity of the generated motion data is tested by dynamics simulation on a virtual robot, then an actual robot performance is carried out. Currently, we use the humanoid robot HRP-1S[9]. The original hardware is developped by Honda [10] and the embedded control system is replaced by HRP Project.

III. Recognition of Leg Primitive Motions

A. Primitive Definition

In our target dances, three basic leg actions are observed as primitive motions: standing (STAND), stepping (STEP), and squatting (SQUAT). Figure 3 shows these actions. Each primitive has parameters which are required to recreate leg motion for a robot. Although there may be additional basic actions such as jumping or spinning in other dances, we

TABLE I

Parameters of Leg Primitives

Primitive	Parameters
STAND	- standing period - waist height
SQUAT	- period (lowest / final time) - the lowest waist height
STEP	- which is a swing / support foot ? - period (highest / final time) - position and orientation of the swing foot at each time - waist orientation at the final time

currently focus attention on the above three actions for present target dances.

STAND represents the motion whereby both legs support the body and maintain balance. This primitive has parameters of waist height and standing period.

STEP represents one stepping motion. To be precise, one foot is raised and then lowered to the floor while the other foot supports the body. The former foot is the *swing foot* and the latter foot is the *support foot*. To express various step actions, a period of primitive is expressed by medium time at which a swing foot takes the highest position, and final time at which the swing foot lands on the floor again. Primitives have states at these times: the position and orientation of a swing foot and the waist. Those properties are described as relative values from the support foot. The primitive does not require the initial state so that it can be adapted to any initial poses.

SQUAT represents one squatting motion, lowering the waist and raising it again. As well as STEP, SQUAT has medium time at which the waist takes the lowest position, and height of the waist at that time.

The definitions of three primitives are summarized in table I

Note that the primitive parameter values which are concerned with the length must be normalized into some standard human model. The normalization enables unified description and adaptation to various robots.

B. Extraction of Primitives

To extract STEP primitives, the speed graph of a foot (Fig. 4) is analyzed. This graph is basically a sequence of bell-shaped curves. During one unit of the curves, the foot moves and stops. This movement is regarded as stepping motion. Hence segments of STEP primitive are extracted as a unit of the curve. Medium time is extracted by finding the highest position of a swing foot in the segment. Note that small sliding movements of a support foot appear as small curves in the graph so those curves should be eliminated. Since a step does not always lift up a swing foot, the height of the foot is useless for judging whether a foot is supporting or swinging.

After each STEP segment is extracted, states at the key times are extracted from captured foot markers as relative values from a support foot.

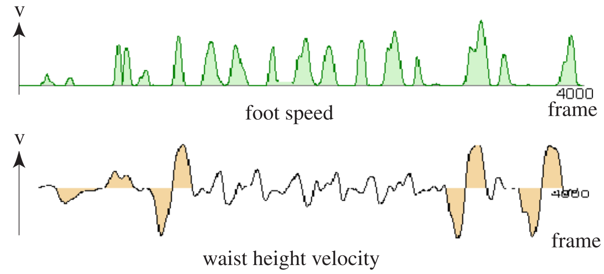

Fig. 4. Graph of foot speed and waist height velocity

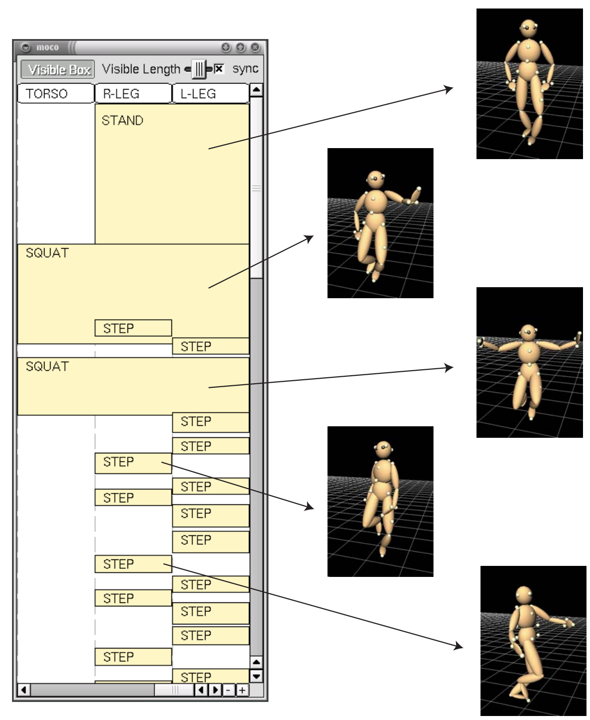

Fig. 5. Extracted primitive sequence and poses in some primitives

To extract SQUAT primitives, a velocity graph of waist height (Fig. 4) is analyzed. In this graph, the squatting action appears as a set of a concave curve and a convex curve, that is, the movement to lower the waist and raise it again. The extraction process has to find this set of curves. As well as STEP, a small area which is nothing but a small swinging action should be eliminated. Waist height at the lowest position is extracted as a medium state.

The segment of the STAND primitive corresponds to the frames where the speed remains approximately zero for a certain period in both foot speed and waist height velocity.

Figure 5 shows an extracted primitive sequence in Jongara-Bushi.

IV. Generation of Leg Motion for a Robot

Leg motion for a robot is generated from a primitive sequence, which is extracted from an original dance motion through the recognition process described in the above section.

First, position and orientation trajectories of feet are generated. Then, initial joint angle trajectories are calculated from the foot trajectories by inverse kinematics of the legs. After leg motion and arm motion are integrated, the waist trajectory is modified to satisfy the dynamic balance.

A. Generating Initial Motion

For each primitive in the acquired sequence, the foot trajectories which represent an action of the primitive are generated. To be precise, the values of the foot trajectory are the position of the ankle joint and the orientation of the sole. These values are expressed on the waist coordinate for using inverse kinematics easily.

For a STEP primitive, foot trajectories are basically created by interpolating initial, medium, and final states of position and orientation. The initial state is the state just before entering the new primitive motion. The medium state and the final state are determined from the parameters of the primitive.

There is the case that the medium state and the final state are modified according to the robot constraints. The modification is caused by contact condition, the constraint of movable range, and self collision.

Although humans can contact their soles freely with the floor, existing robots must contact a sole flat against the floor. Otherwise, the foot cannot support the body stably. Thus the final foot orientation in acquired STEP primitives must be constrained to be horizontal, and the final position must be level with the floor.

Movable ranges of robot legs are usually narrower than those of humans. If a state of a primitive is beyond the ranges, it must be restricted to lie within the ranges.

Since a robot like HRP-1S has fat legs, self collision of legs easily occurs. Compared with humans, self collision is significant problem for robots. Collisions must be detected and eliminated in the generation process.

For a SQUAT primitive and a STAND primitive, foot trajectories are generated by a similar process to that of STEP.

After foot trajectories are generated, joint angle trajectories are calculated by inverse kinematics.

B. Dynamic Balance

A target robot has the ability to move joints according to the motion generated through the above process. However, if the actual robot tries to perform the motion on a floor, the robot is unable to maintain balance and therefore, it falls down.

Leg motion is generated under the assumption that all the area of a foot sole is in contact with the floor when it is supporting the body. In other words, a sole does not rotate during that time. In terms of dynamics, this assumption is satisfied when the point at which the moment to the robot body is zero is in the area of the sole surface. During this time, the sole does not rotate. The point is called 'zero moment point (ZMP)' [11] and the area is called *supporting area*. If a robot is supported by both feet, the supporting area corresponds to the convex area which consists of both soles.

Given the physical model of a robot, a trajectory of ZMP can be calculated from motion data for the robot, under the assumption that the supporting area is infinite. If ZMP moves out of an actual supporting area, the motion is impossible to perform because the actual motion must imply rotation of the supporting sole at that time, so that the sole moves away from the ground and the robot falls down. Therefore, ZMP must always be inside the supproting area [12]

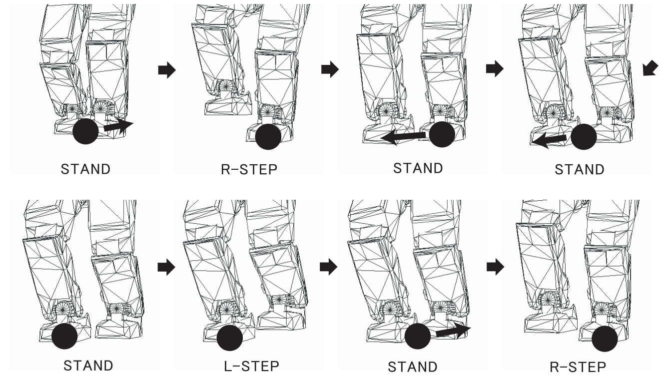

Fig. 6. ZMP transition with support state transition (Each round marker shows ZMP)

In this study, a desired ZMP trajectory which is always in the supporting area is prepared first. Then the motion is modified to realize the trajectory.

C. Desired ZMP Trajectory

The condition that ZMP is inside a supporting area must be satisfied, but it is not a sufficient condition for actual stability. If a supporting area remains in one state, ZMP should remain a stable point in the supporting area, but supporting state changes with steps. Stability of motion depends on a ZMP trajectory with state transitions. In this study, we applied the following criteria.

- In STEP period, ZMP must be located at the center of a supporting sole.
- In STAND period, ZMP moves from a previous position to the next supporting position by third order polynomial equation. Initial velocity and accelerations and final ones are kept zero.
- If the STAND period is long, transition is separated into three steps: (1) from a previous position to the center of the supporting area, (2) stay there and (3) move to the next supporting position.
- If the STAND period is short, ZMP movement becomes rapid and robot motion becomes unstable. Adequate transition time is required for stability. In this case, ZMP movement is expanded so that it starts in the previous state and extends into the next state. Acceleration and deceleration of ZMP is done in those states.

Figure 6 shows a sequence of foot primitives and ZMP movements.

D. How to realize a desired ZMP

Given a desired ZMP trajectory, a motion must be modified to realize it. Nishiwaki et al. [13] proposed a method to solve this problem. On a discrete system, supposing all the segments are restricted to be translated horizontally in the same distance, the following equation is acquired:

$$x_{zmp}^e(t_i) = \frac{-hx^e(t_{i+1}) + (2h + g\Delta t^2)x^e(t_i) - hx^e(t_{i-1})}{g\Delta t^2}$$

where x_{zmp}^e is a difference between a calculated ZMP and a desired ZMP, x^e is a translation distance to realize the desired ZMP, t_i is time at frame i, h is height of center of mass, Δt is time per one frame. This equation is about x-axis, and similar equation applies to y-axis.

This equation is expressed by information from 3 consecutive frames. These kinds of equations are solved as tridiagonal simultaneous linear equations.

This method cannot figure out a result which completely follows the desired ZMP trajectory in one calculation because the constraint that all the segments translate parallel in the same distance is actually impossible. However, by iterating the calculation, a converged result is acquired.

V. Results

In recognition, primitive sequences were successfully extracted as Fig.5. In this section, we examine the validity of generated robot motion.

A. Appearance of Generated Motion

Figure 7 shows a captured dance motion and a generated robot motion for HRP-1S in Jongara-Bushi. Animation here is not dynamics simulation, but just kinematic replay of motion data.

The result can be regarded as a good imitation of the human performance. The most remarkable difference is the trajectories of a whole body. Since movable range of the robot legs is restricted, the robot cannot always follow a rapid turn or a wide step in the original motion.

Although a method to evaluate similarity or skill of performances is necessary, we currently have no proper criteria. Evaluation such as distance comparison of trajectories of some body parts makes no sense because body type is different between that of a robot and a human; this applies to skilled human dancers.

B. Feasibility on Actual Robots

Compared with appearance, validity of dynamics is clearly evaluated by whether a robot accomplishes a performance or falls down. First, a motion is tested on the dynamics simulator. If the robot can stably perform a dance from beginning to end on simulation, a performance by an actual robot is experimented.

Dynamics simulation was tested with the controller developped by HRP Project [9]. Although the controller has the ability to maintain balance against small disturbances, the initial motion could not accomplish a performance. After balance modification was applied, the robot successfully performed a whole dance without falling down. However, a support foot slid and motion became unstable when the robot widely rotated the waist to turn. Balance modification does not consider horizontal (yaw) factor of moment, which is not concerned with ZMP. Balance modification should be improved to solve this behavior. Finally, we tested the motion on the actual robot HRP-1S.

Because of the above problem, an experiment was done under the condition that rotations in turn steps were eliminated. Additionally, speed of the robot motion was restricted as half of the original one because of the capacity of the current actuators. Under these conditions, the robot successfully performed the whole dance without falling down (Fig. 8). Although some restrictions were applied, it is a remarkable result that this kind of dynamic motion with the whole body was performed by the actual robot.

VI. Conclusion

In this paper, we have proposed primitives of leg actions for dance imitation for a robot. On the basis of the primitives, human dance performance is recognized as a sequence of primitives, and feasible robot motion is generated from the recognition result. Our method realized a dance performance with leg actions by an actual humanoid robot.

However, the current method requires some restrictions to actual performances. For future work, the generation method should be improved to acquire a more stable performance. Additonally, evaluation criteria of performance skill should be defined and a robot should achieve high-skilled performances with the criteria.

Acknowledgment

This research is supported in part by Japan Science and Technology Corporation under the Ikeuchi CREST project.

References

[1] A. Nakazawa, S. Nakaoka, and K. Ikeuchi, “Imitating human dance motions through motion structure analysis,” in *Proceedings of International Conference on Intelligent Robots and Systems*, EPFL, Lausanne, Swizerland, September 2002.

[2] T. Yukawa, T. Kaiga, K. Nagase, and H. Tamamoto, “Human motion description system using buyo-fu (in japanese),” *Information Processing Society of Japan*, October 2000.

[3] N. S. Pollard, J. K.Hodgins, M. J. Riley, and C. G. Atkeson, “Adapting human motion for the control of a humanoid robot,” in *Proceedings of International Conference on Robotics and Automation*, Washington DC, April 2002.

[4] J. Takamatsu, H. Kimura, and K. Ikeuchi, “Calculating optimal trajectories from contact transitions,” in *Proceedings of International Conference on Intelligent Robots and Systems*, EPFL, Lausanne, Swizerland, October 2003.

[5] T. Morita, J. Takamatsu, K. Ogawara, H. Kimura, and K. Ikeuchi, “Knot planning from observation,” in *Proceedings of International Conference on Robotics and Automation*, Taipei, Taiwan, September 2003.

[6] T. Inamura, I. Toshima, and Y. Nakamura, “Acquisition and embodiment of motion elements in closed mimesis loop,” in *Proceedings of International Conference on Robotics and Automation*, Washington DC, April 2002.

[7] S. Nakaoka, A. Nakazawa, K. Yokoi, and K. Ikeuchi, “Recognition and generation of leg primitive motions for dance imitation by a humanoid robot,” in *Proceedings of 2nd International Symposium on Adaptive Motion of Animals and Machines*, Kyoto, Japan, March 2003.

[8] F. KANEHIRO, K. FUJIWARA, S. KAJITA, K. YOKOI, K. KANEKO, H. HIRUKAWA, Y. NAKAMURA, and K. YAMANE, “Open architecture humanoid robotics platform,” in *Proceedings of International Conference on Robotics and Automation*, Washington DC, April 2002.

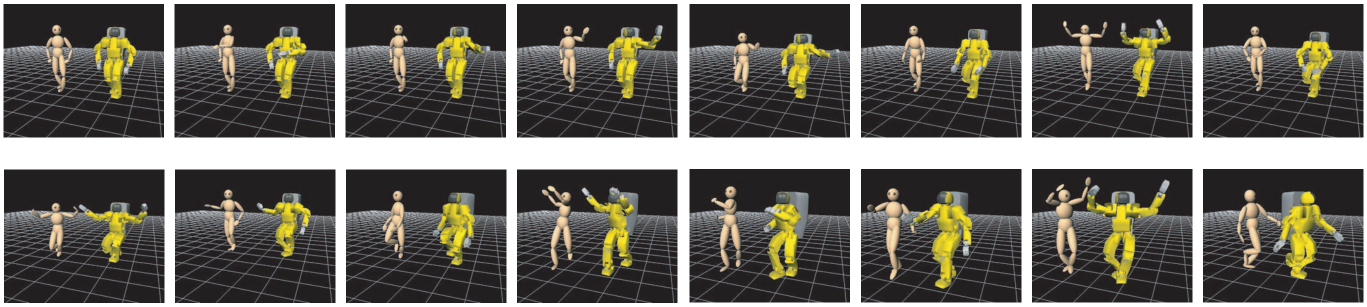

Fig. 7. Original motion and generated robot motion

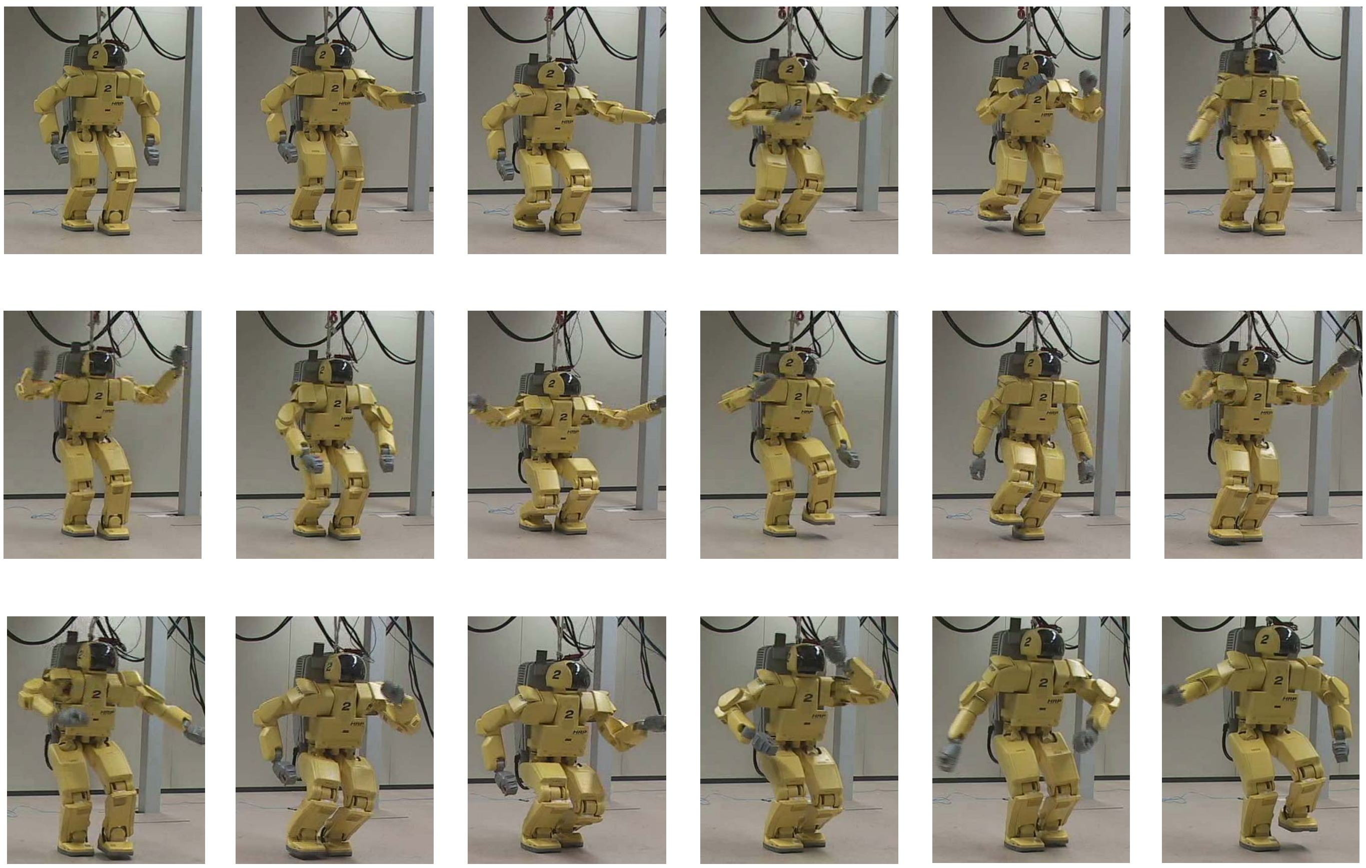

Fig. 8. Actual Dance Performance by HRP-1S

[9] K. Yokoi, F. Kanehiro, K. Kaneko, K. Fujiwara, S. Kajita, and H. Hirukawa, "A honda humanoid robot controlled by aist software," in *Proceedings of the IEEE-RAS International Conference on Humanoid Robots*, Tokyo, November 2001.

[10] K. Hirai, M. Hirose, Y. Haikawa, and T. Takenaka, "The development of honda humanoid robot," in *Proceedings of International Conference on Robotics and Automation*, Leuven, May 1998.

[11] M. Vukobratovic, B.Borovac, D.Surla, and D.Stokic, "Biped locomotion: Dynamics, stability, control and application," *volume 7 of Scientific Fundamentals of Robotics*, 1990.

[12] S. Tak, O. young Song, and H.-S. Ko, "Motion balance filtering," in *Proceedings of EUROGRAPHICS*, vol. 19, no. 3, INTERLAKEN, August 2000.

[13] K. NISHIWAKI, S. KAGAMI, Y. Kuniyoshi, M. INABA, and H. INOUE, "Online generation of humanoid walking motion based on a fast generation method of motion pattern that follows desired zmp," in *Proceedings of International Conference on Intelligent Robots and Systems*, EPFL, Lausanne, Swizerland, October 2002.