The 2010 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 18-22, 2010, Taipei, Taiwan

Temporal Scaling of Leg Motion for Music Feedback System of a Dancing Humanoid Robot

Takahiro Okamoto† Takaaki Shiratori◇ Shunsuke Kudoh‡ Katsushi Ikeuchi†

Abstract— **In this paper, we propose a method to achieve temporal scaling of leg motions as a fundamental technique for a music feedback system of a dancing humanoid robot. We asked dancers to perform dance motion at normal musical tempo and faster musical tempos and observed how dancers modified performance for given musical tempos. The obtained insights from the observation are 1) a dancer needs to preserve leg postures that are important to emphasize dance expression, 2) there is a priority to determine what features of leg motion can be adjusted, and 3) stylistic leg motion resembles normal step motion if dancers cannot follow fast musical tempo completely. Based on these insights, we generate leg motion appropriately adjusted for changing musical tempo while maintaining balance. We validated our method via simulation experiments with a humanoid robot HRP-2.**

I. INTRODUCTION

Interest in entertainment robots has prompted researchers to achieve human-robot interaction [1] and to entertain people at theme parks and stage shows [2], [3]. Dance performance is one recognized form of this entertainment. Nakaoka *et al*. [4] have produced a humanoid robot that can maintain its balance while dancing. They employed a top-down approach called *task models*, which analyze the structure of human motion based on predesigned abstract models for imitation by a robot. However, the system still lacks a human's capability for synchronizing dance motion with musical tempo, a capability that we call a *music feedback system*. We believe that the music feedback system makes the dance performances more attractive to the audience.

In this paper, we focus on step motion of dance and propose a method to achieve temporal scaling of leg motions as a fundamental technique of the music feedback system. We basically extend the task models of leg motions proposed by Nakaoka *et al*. [4], and our technique allows for modifications of leg motion so that the robots can follow changing musical tempo, maintain balance, and preserve motion characteristics. Shiratori *et al*. [5] already proposed a similar technique for upper body motion based on various musical speeds. By combining upper body and leg motions, we can generate whole body motion that is adapted to new musical tempo.

Several researchers have applied temporal scaling, also called time warping, to human motion. Dynamic time warping is often used to adjust timings of motion-capture data [6], [7], [8]. McCann *et al*. [9] proposed a method of motion retiming based on physical properties. However, these techniques do not work for our case. Simple temporal scaling often causes movements that exceed joint limitations of robots or unstable motions which fail to maintain balance. Yoshii *et al*. and Murata *et al*. proposed a method that enables a biped humanoid robot to step with synchronization to musical tempo [10], [11]. However, their target motion is simple stepping and complicated motion like dance cannot be handled.

Our method is based on insights obtained through observation of how dancers modify performance when musical tempo becomes faster than the original one. According to the observation, we found that 1) a dancer needs to preserve postures that are important to emphasize dance expression, 2) there is a priority to determine what motion features can be attenuated, and 3) stylistic leg motion is adapted to resemble normal step motion if dancers cannot follow a fast musical tempo completely. Based on these insights, we generate appropriate leg motions that are adjusted for changing musical tempo while maintaining balance.

This paper is organized as follows. Section II describes our strategies for the dancing humanoid robot and clarifies our contributions. Section III describes details of the observation of dancer's motion, and Section IV explains our method of temporal scaling for leg motion based on insights obtained from the observation. Section V shows our simulation results. Section VI discusses our contributions and limitations, and Section VII concludes this paper.

II. DANCE IMITATION BASED ON TASK MODELS

As mentioned above, the proposed method in this paper is an extension of previous studies on imitating dance performance by a robot using task models. Before describing the proposed method, this section gives a summary of the previous studies. They are based on a paradigm called *Learning-From-Observation* (LFO), in which a robot observes a target dance performed by skillful dancers, recognizes what they do using task models, and maps the recognized motion to robot actions for mimicking them. Task models consist of *tasks* and *skill parameters*. Tasks are the knowledge of what to do, while skill parameters are the knowledge of how to do each task.

In general, when we say "mimicking an action," it does not mean mimicking the entire performed action. It is difficult, even if not impossible, to repeat the same trajectories to be mimicked, because each person has different dimensions in parts of his or her body. Instead, for this purpose, we extract

†Takahiro Okamoto and Katsushi Ikeuchi are with the University of Tokyo, Japan. {tokamoto, ki}@cvl.iis.u-tokyo.ac.jp

◇Takaaki Shiratori is with Carnegie Mellon University, USA. siratori@cs.cmu.edu

‡Shunsuke Kudoh is with the University of Electro-Communications, Japan. kudoh@is.uec.ac.jp

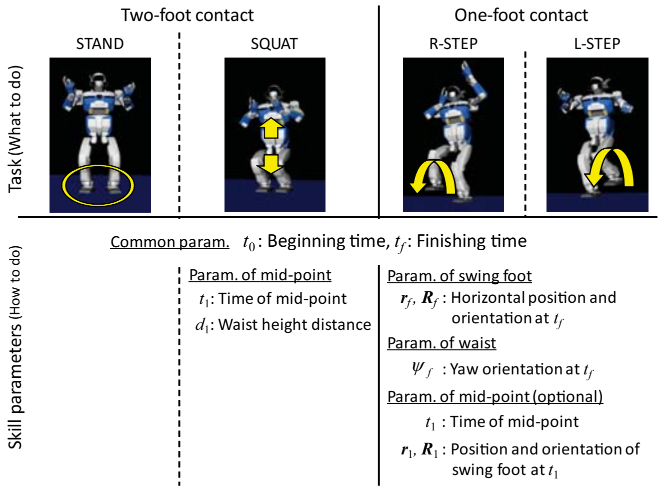

Fig. 1. Leg task models: the task model consists of *tasks* and *skill parameters*. The former explains what to do, and the latter explains how to do it.

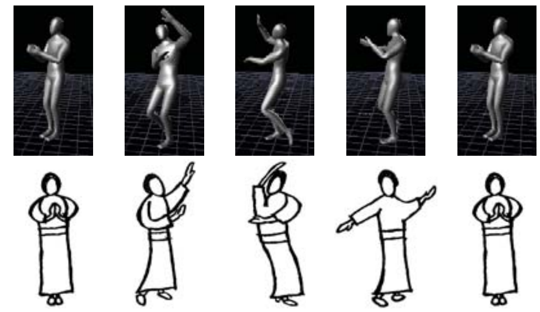

Fig. 2. Keyposes in Aizu-bandaisan dance: Top row: extracted keyposes by Shiratori *et al*.'s method, and bottom row: corresponding poses in a textbook written by a dance master.

important features of the actions, and perform only those important features. The LFO introduces abstract task models to represent those essential parts.

The LFO paradigm is not only for imitating dance, but for learning general human actions. In fact, the LFO paradigm has been applied to various hand-eye operations such as object assembly [12] and knot tying [13]. In each case, the appropriate task models are predesigned using the knowledge of action domains under the top-down approach.

For imitation of dance, we used different strategies to apply the LFO paradigm to leg and upper body motions, because leg and upper body motions of dancing robots have two different purposes. The leg motions need to stably support robot bodies, while upper body motions need to express dancing patterns. After generating leg and upper body actions separately, we then concatenated and adjusted those generated motions of leg and upper body.

Nakaoka *et al*. [4] proposed *leg task models* for generating leg motion of dance performance. Four tasks are defined by considering a contact state between the feet of a humanoid robot and the floor, through the top-down analytic approach (Fig. 1). Skill parameters describe task-specific timings and spatial characteristics of each task motion as shown in the bottom row of Fig. 1.

A task sequence is recognized from marker trajectories obtained by a motion-capture system. Temporal segments are extracted and then recognized as tasks. From each segment, the values of the skill parameters corresponding to the task are obtained.

Leg motion of a robot is reconstructed from an extracted task sequence and corresponding skill parameters. Robot have a standard trajectory for each task that produces stable motion, and leg motion is generated by modifying the trajectory based on the extracted skill parameters.

Shiratori *et al*. proposed a concept of *keyposes* as an essential component in dance motion [14]. A keypose can be defined as a fixed posture of a dancer for the purpose of providing the viewers with expression and meanings of dance. It is considered important in dance performance to properly represent those keyposes with appropriate timings. Theoretically, keyposes can be obtained by detecting brief pauses in a motion sequence, but for accurate detection musical information is utilized here to narrow keypose candidates. As an example, Fig. 2 shows the extracted keyposes from the Aizu-bandaisan dance (see also the accompanying video, which includes a video clip of the dance). Keyposes can be used for high-quality imitation of dance by a robot. Transitions between keyposes are considered as tasks for upper body motion. The timing and posture of keyposes are represented as precisely as possible in generating robot motion, while transition between keyposes can be modified if it is difficult for a robot to make the same motion as the original.

This method was extended to enable a robot to dance at a different musical tempo from the captured one [5]. The temporal scaling of motion is necessary, for example, when we have a robot dance to a live performance of music when the musical tempo is not always constant. In this method, the upper body motion is modeled using hierarchical B-spline. When the musical tempo gets so fast that motion to the music by simple temporal scaling is beyond the robot's capability, the upper body motion is modified by their method based on observation of dance performance by humans. However, this method can cope with only upper body motion. Temporal scaling of leg motion is difficult by this method because constraints in leg motion, such as balance and a closed loop by legs and the floor, are different from upper body constraints. Therefore, leg motion was modified manually in the study.

This paper focuses on step motion and proposes a method for temporal scaling of leg motion. Adding the method to the previous studies, we can generate whole body motion to arbitrary tempos in music. First, human performance is observed to determine how to adjust motion to the musical tempo. Next, a model for temporal scaling is designed based on the observation.

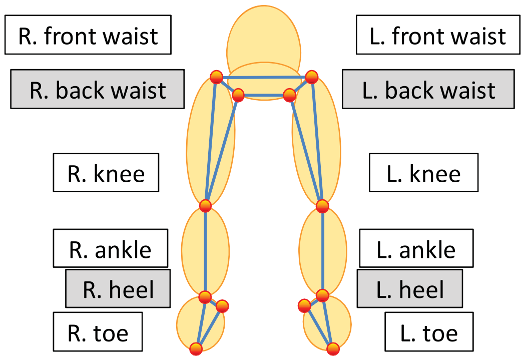

Fig. 3. Optical markers. Each box represents an optical marker with a marker name, and gray-colored boxes represent markers behind the body from this viewpoint. COM of each foot is computed to get foot trajectories.

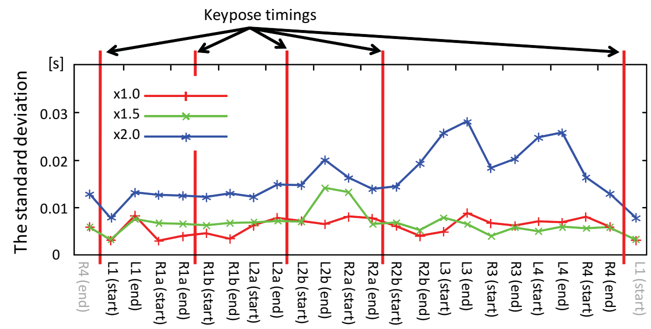

Fig. 5. Standard deviation of the start/end timing of STEP tasks (Dancer B).

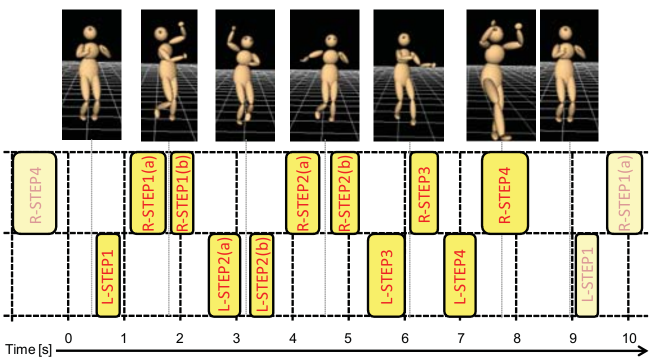

Fig. 4. STEP tasks in a cycle of Aizu-bandaisan dance.

III. Observation of Human Dance Motion

Human performance of dance was captured using the VICON's optical motion capture system. Configuration of optical markers for legs is illustrated in Fig. 3. We computed center of mass (COM) of each foot using the ankle, heel and toe markers to obtain foot trajectories. We asked three dancers to demonstrate Aizu-bandaisan dance, a Japanese folk dance. One is a dance master (Dancer A) and the others are her students, a veteran (Dancer B) and a beginner (Dancer C). As their experience in dance varies widely, the common features among the three can be considered essential in this dance. Aizu-bandaisan dance is a cyclic dance, and one cycle takes about 10 seconds. Dancers keep repeating it during the dance. In a cycle, there are eleven STEP tasks as shown in Fig. 4.

The musical speeds we captured are the original speed, 1.2 times faster, 1.5 times faster, 1.8 times faster, and 2.0 times faster. Music at double speed is actually not practical for the dance, but, because humans can move more quickly than robots, we need to investigate human performance to the music when tempo is too fast for humans to move perfectly. For each musical speed, we captured dance performance of 10–15 cycles.

The observation of captured dance provides four insights about leg motion:

1) Timing of STEP tasks near keyposes tends to be maintained,
2a) Strides of each STEP task tend to be maintained,
2b) Speed of a swing foot does not accelerate as much as musical tempo, and
3) Kicking actions, which appear at the end of each cycle, converge to a normal step as musical tempo increases.

The first insight is the relationship of the timing of STEP tasks to keyposes. When the start or end timing of a STEP task exists near a keypose, the timing tends to be maintained even though the musical tempo gets faster. Fig. 5 shows this phenomenon. The vertical axis represents the standard deviation of timing for STEP tasks in the horizontal axis. For comparing timing among different musical tempo, we introduce the concept of *relative time*, which is a scaled time as 0.0 at the beginning of a cycle and 1.0 at the end of it. We can compare timing of actions captured in several musical tempos by the relative times.

For the normal musical tempo, the values are roughly same for all STEP tasks. For 2.0 times faster speed, the values increase for all tasks, but the values for tasks near keyposes are larger than those for the other tasks. This indicates that dancers made more effort to maintain the original timings for tasks around keyposes, when keeping up with faster musical tempo becomes difficult.

The second insight is that the stride for each STEP task tends to be maintained as shown in Fig. 6. The horizontal axis represents musical tempo, and the vertical axis represents the average stride length. The length is almost constant despite the musical tempo becoming faster. It is natural that dancers tend to maintain strides because strides are considered one of the dominant factors in dance expression. In some strides, the length of stride decreases, but these are tasks that do not adjoin keypose timings, such as L-STEP3 and R-STEP3. From this observation, we can consider that the dancer tried to maintain stride length as close to original one as possible around keypose timings.

The third insight is about the speed of a swing foot. Fig. 7 shows the relationship between the maximum speed in each STEP task and the musical tempo. The horizontal axis represents musical tempo, and the vertical axis represents the maximum speed. The value increases following the musical tempo. However, the amount of the increase is slight and not

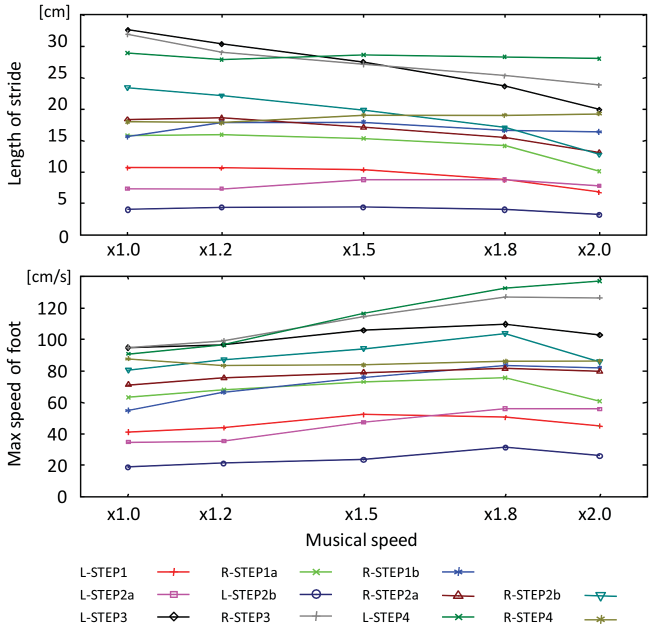

Fig. 6. Stride (top) and maximum speed (bottom) of each STEP task (Dancer B).

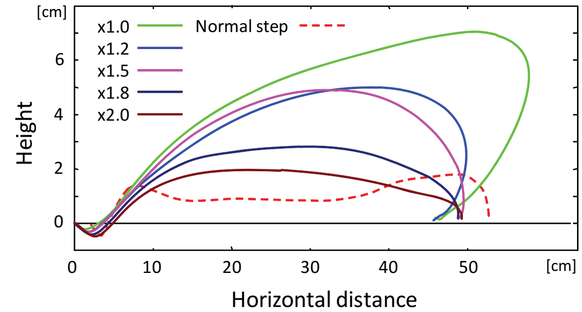

Fig. 7. Trajectory of a swing foot in kicking actions, R-STEP4 (Dancer B).

as great as that of the musical tempo, probably due to the physical limitation of the dancer. In order to maintain strides, the dancer increases the duration of STEP tasks as well as the speed of a swing foot.

The last insight is about stylistic motion such as kicking actions that appear at the end of a cycle, which appear as R-STEP4 in Fig. 4. All the other tasks are *normal steps*, for which only modifications of speed and duration described above are sufficient (leg task models distinguish two types of STEP tasks: normal steps and other *expressive* actions. Details of these are described in Section IV). With regard to kicking actions, the trajectories of the foot change as the musical tempo increases, as shown in Fig. 7. The horizontal and vertical axes represent horizontal and vertical distances of a swing foot in a kicking action for each musical tempo, respectively. We can see that the foot trajectory of a kicking action becomes closer to that of a normal step.

Although we just show the graphs for one dancer (Dancer B) in the above discussion due to the limitation of space, observation of the other two dancers (Dancers A and C) also supports the above insights. An algorithm for temporal scaling of the leg motion is designed based on these insights.

All of these insights can be expressed by skill parameters of the leg task models in Fig. 1, such as the beginning time (t_0), the finishing time (t_f), and the middle and landing positions of a swing foot ($\boldsymbol{r}_f$, $\boldsymbol{r}_1$). This makes it easy to extend the leg task models for temporal scaling to music. The details of these parameters and generation of leg motion with them are described in the following section.

IV. Temporal Scaling of Leg Motion by Musical Speed

This section provides details on the leg motion generation from the skill parameters and the concept of the extension that we designed.

A. Leg Motion Generation based on Task Model

In Nakaoka *et al.*'s task models, a trajectory of a swing foot in a STEP task is reconstructed using a smooth interpolation function based on a cubic polynomial that passes along three points computed from skill parameters: the starting point $(t_0, \boldsymbol{r}_0)$, the middle point $(t_1, \boldsymbol{r}_1)$, and the landing point$(t_f, \boldsymbol{r}_f)$. This function is expressed as follows:

$$\boldsymbol{f}_n\langle(t_0, \boldsymbol{r}_0), (t_1, \boldsymbol{r}_1), (t_f, \boldsymbol{r}_f)\rangle(t). \tag{1}$$

There are two ways to determine the middle point. One is the case of a *normal step*, in which a captured trajectory of the swing foot is similar to a step in usual walking. For a normal step, the middle point calculated by the following equations is used:

$$t_1 = \frac{t_0 + t_f}{2}, \tag{2}$$

$$\boldsymbol{r}_1 = \left(\frac{r_x^0 + r_x^f}{2}, \frac{r_y^0 + r_y^f}{2}, h\right)^T, \tag{3}$$

where h is a predefined value as the normal step height. The other case is when the captured trajectory differs largely from the trajectory of a normal step. For such a *stylistic* action, the middle point is determined directly from the foot position of human motion data at the timings of the middle point. In our implementation, these two cases are distinguished by a degree of the difference between the middle point of the captured trajectory and the point calculated by Equation (3).

A time series of leg tasks in performance is called a *task sequence*. An example of a task sequence is shown in Fig. 4. In a task sequence, R-STEP and L-STEP tasks must not overlap. Although an overlap of R and L STEP tasks means jumping or flight phase of running, the current leg task models do not treat these motions. If there is no interval between neighboring STEP tasks or an interval is shorter than a certain length, an interval is made in extracting the task sequence so that Zero Moment Point (ZMP) can move

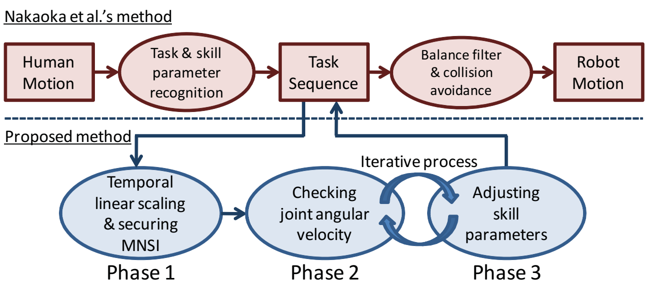

Fig. 8. Overview of the temporal scaling algorithm for leg motion.

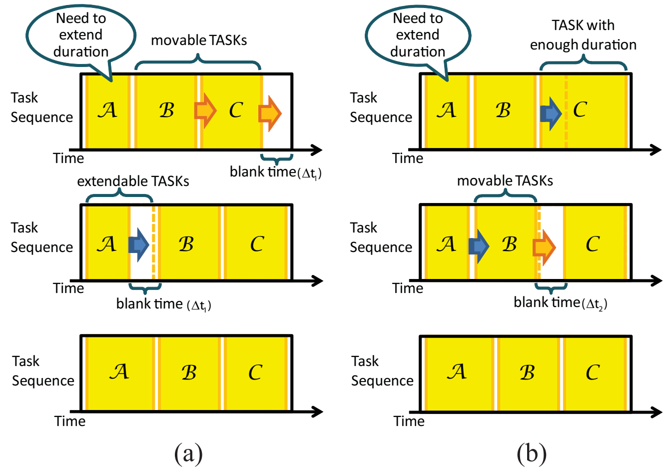

Fig. 9. Extension of duration by (a) shifting and (b) contraction. Orange and blue arrows represent shifting task timing and extending/shortening of task duration, respectively. Yellow boxes indicate STEP tasks.

stably in the support polygon for balance maintenance. The minimal length of the interval is called *Minimum Non-Step Interval* (MNSI), and we use empirical value of 0.07 sec. MNSI is a time interval required in order for the ZMP move to from one foot to another.

Finally, leg motion of a robot is generated by solving inverse kinematics from foot trajectories obtained from a task sequence, and skill parameters are refined for creating reasonable robot motion with balance maintenance and collision avoidance. The top row of Fig. 8 illustrates these procedures.

B. Temporal Scaling of Task Sequence

As we discussed in Section III, all the insights from the observation can be expressed by the skill parameters of the leg task models. Therefore, we can easily achieve temporal scaling by extending the Nakaoka *et al.*'s leg task models [4]. Fig. 8 illustrates the relationship between our temporal scaling method and the leg motion generation of the Nakaoka *et al.*'s algorithm. After a task sequence is extracted from human motion, we adjust it to new musical tempo by our temporal scaling method, and then it is passed to the skill refinement stage for balance maintenance and collision avoidance.

Our temporal scaling algorithm consists of the following three phases:

- **Phase 1:** the whole task sequence is temporally scaled linearly, and the intervals between STEP tasks are adjusted if they are shorter than MNSI.
- **Phase 2:** STEP tasks in which joint angular velocity exceeds the limit are detected by calculating inverse kinematics.
- **Phase 3:** skill parameters of the STEP tasks are modified by changing parameters of duration and stride.

Phases 2 and 3 are performed iteratively until joint angular velocities do not exceed limits for all STEP tasks.

Phase 1: First, linear temporal scaling is applied to a task sequence. The start and end timings t_0 and t_f for all tasks are divided by the musical tempo. This process generates a new task sequence approximately adjusted to the new musical tempo, but may shorten intervals between STEP tasks too much. For stable dance performance, all the STEP intervals need to be longer than MNSI. If an interval is shorter than MNSI, it is lengthened by moving forward the finish time t_f of the task right before the interval. Although modification may change the end timing of tasks around keyposes, the modification is not noticeable because MNSI is significantly small.

Phase 2: Foot trajectories are generated from the adjusted task sequence, and entire joint angle trajectories are calculated from the foot trajectories by inverse kinematics. Then, whether or not joint angular velocity resulting from skill parameters of STEP tasks exceeds the limit is checked. If there exist joints that exceed the limits of joint angular velocities, the excess rate and index of the task are passed to Phase 3.

Phase 3: Based on the information acquired by Phase 2, skill parameters are modified to adjust the joint angular velocities that exceed the limit. The insights 2a and 2b indicate that duration of a task can be changed more flexibly than stride. Therefore, our algorithm gives higher priority to modifying duration than to modifying stride, and it modifies stride only if the modification of duration cannot generate a motion satisfying the limitation. If leg motions exceed joint limitation, insight 2b indicates that duration of the task can be extended to reduce joint angular velocity. However, all STEP task intervals must be longer than MNSI for stable motion, and this makes simple extension of duration difficult. This phase is divided into two parts, *a* and *b*, depending on whether or not an extensible STEP interval exists.

Phase 3a: This is a case in which at least one extensible STEP interval exists somewhere in a task sequence, though robot motion still exceeds joint limitation. Fig. 9 (a) illustrates this case. Let a task sequence consist of three tasks, $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$. $\mathcal{A}$ is a task in which robot motion exceeds the limitation of joint angular velocity, and $\mathcal{B}$ and $\mathcal{C}$ are not. If an extensible task interval adjoins $\mathcal{A}$, the duration of $\mathcal{A}$ is simply extended. If the interval adjoining $\mathcal{A}$ is not sufficiently long and $\mathcal{C}$ adjoins a task interval longer than MNSI, $\mathcal{B}$ and $\mathcal{C}$ are shifted into the interval next to $\mathcal{C}$ so that the interval adjoining $\mathcal{A}$ becomes long enough to extend the duration of $\mathcal{A}$. The start time t_0 and finish time t_f of $\mathcal{B}$ and $\mathcal{C}$ are

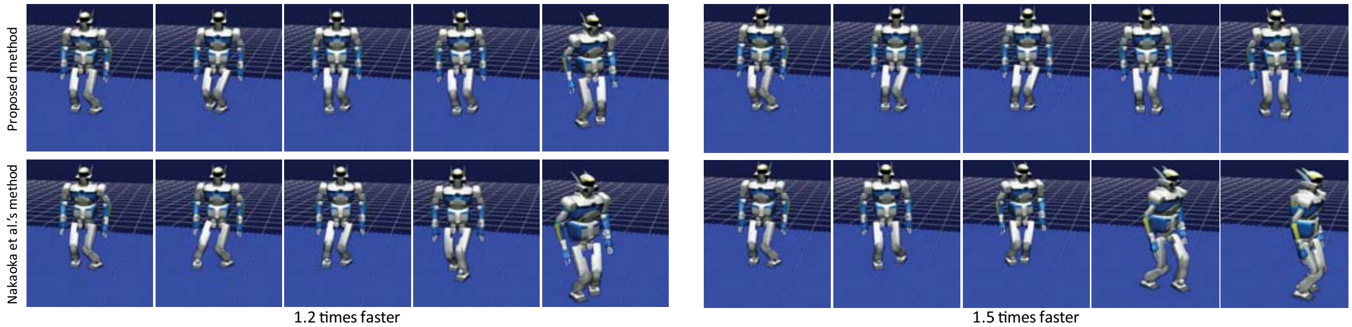

Fig. 10. Simulation results of Aizu-bandaisan dance at music 1.2 times faster than the original tempo. Top: result by the proposed method, and bottom: result by Nakaoka *et al*.'s method for comparison.

changed as:

$$t'_0 = t_0 + \Delta t_1, \tag{4}$$
$$t'_f = t_f + \Delta t_1, \tag{5}$$

and $\mathcal{A}$ is extended using the extended step interval as:

$$t'_0 = t_0, \tag{6}$$
$$t'_f = t_f + \Delta t_1, \tag{7}$$

where Δt_1 depends on length of the long task intervals and timing of keypose, which must be preserved, according to the insight 1. If robot motion after Phase 3a still exceeds the limitation, Phase 3b is applied.

Phase 3b: This is a case in which no extensible task interval exists in the whole task sequence. Fig. 9 (b) illustrates this case. Let $\mathcal{A}$ be a task in which resulting robot motion exceeds the limitation of an actuator and $\mathcal{C}$ be a task whose joint angular velocity does not exceed the limitation. First, the duration of $\mathcal{C}$ is shortened by Δt_2, which is a small constant value (0.01 sec in our implementation) as:

$$t_0^{\mathcal{C}\prime} = t_0^{\mathcal{C}} + \Delta t_2, \tag{8}$$
$$t_f^{\mathcal{C}\prime} = t_f^{\mathcal{C}}, \tag{9}$$

and $\mathcal{B}$ shifts by Δt_2 as

$$t_0^{\mathcal{B}\prime} = t_0^{\mathcal{B}} + \Delta t_2, \tag{10}$$
$$t_f^{\mathcal{B}\prime} = t_f^{\mathcal{B}} + \Delta t_2, \tag{11}$$

and, the duration of $\mathcal{A}$ is extended by Δt_2 as:

$$t_0^{\mathcal{A}\prime} = t_0^{\mathcal{A}}, \tag{12}$$
$$t_f^{\mathcal{A}\prime} = t_f^{\mathcal{A}} + \Delta t_2. \tag{13}$$

The above procedures and Phase 2 are repeated until robot motion of all tasks satisfies the joint limitations. If these duration extensions are still not sufficient, we apply Phase 3c finally.

Phase 3c: If the excess of joint angular velocity cannot be solved by extension of duration, the stride is shortened based on insight 2a as follows:

$$\boldsymbol{r}_f - \boldsymbol{r}_0 \rightarrow \alpha(\boldsymbol{r}_f - \boldsymbol{r}_0), \tag{14}$$

where α is calculated with the excess ratio provided by Phase 2 and the ratio of the duration extension via Phases 3a and 3b, and $\boldsymbol{r}_0$ represents a starting position of the swing foot while stepping respectively.

If a STEP task contains skill parameters related to a middle point, we modify them based on insight 3. Our method modifies the skill parameter of a middle point position $\boldsymbol{r}_1$ to realize this phenomenon as:

$$\boldsymbol{r'}_1 = (1-\alpha)\boldsymbol{r}_1 + \alpha \boldsymbol{r}_{1s}, \tag{15}$$

where $\boldsymbol{r'}_1$ represents a modified middle point position, and $\boldsymbol{r}_{1s}$ represents the standard middle point defined by the right side of Equation (3).

V. Simulation Results

As an experiment, the motion of the Aisu-bandaisan dance was generated to music 1.2 and 1.5 times faster than the original. The algorithm was integrated into our software for handling whole body motion [15]. The methods by Shiratori *et al*. and Nakaoka *et al*. described above were also incorporated into it. Model files of robots and motion data in the system are compatible to OpenHRP [16]. We used functions of OpenHRP for dynamic simulation and control. As a robot platform, we used HRP-2 [17].

The first experiment is generation of leg motion. Joints in the upper body are fixed here. Musical tempos we tested are 1.2 and 1.5 times faster than the original. The results are shown in Fig. 10. The top figures are results by the proposed method. The bottom figures are results by the Nakaoka *et al*.'s method with simple temporal scaling, which can be done by dividing only the start and end timings of all STEP tasks by input musical speed. The left side is resulting motion for 1.2 times fast music, and the right side is that of 1.5 times fast music. Since the joint limitations cannot be considered in the current simulator, the robot can perform motion which violates the limitations in the simulation, though this is an infeasible motion due to limitation of the physical actuators' power. The proposed method shrinks the stride appropriately to satisfy the limitation of the actuators.

The next experiment is generation of whole body motion. The upper body motion is generated by Shiratori *et al*.'s method [5]. The results are shown in Fig. 11. The top figures are results for music 1.2 times faster than the original tempo. The bottom figures are results for music 1.5 times faster

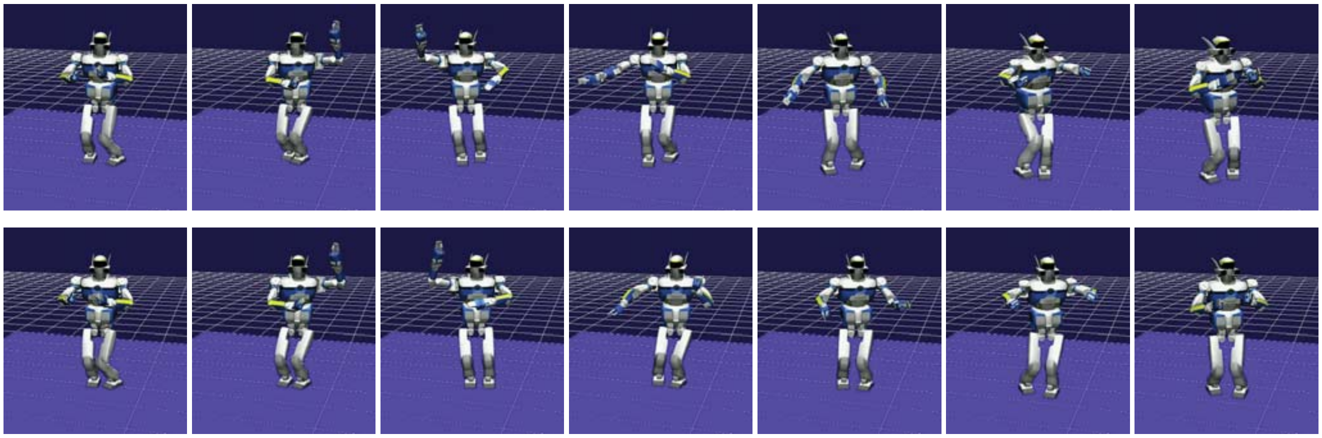

Fig. 11. Simulation results of whole body motion for Aizu-bandaisan dance. Top row: result for music 1.2 times faster than the original speed, and bottom row: result for music 1.5 times faster than the original speed. The upper body motion is created by Shiratori *et al*.'s method.

than the original tempo. The robot could perform dance performances with balance maintenance.

Fig. 12 shows the angular velocity of an actuator at the right knee during these simulations. As quick motions appear at a knee in kicking actions, we chose it as a typical example for comparison. The top graphs show results with fixed upper body posture and the bottom graphs show results with upper body motion generated by Shiratori *et al*.'s method. Velocities generated by our method (blue) are always within the limits in the simulation, while velocities generated by Nakaoka *et al*.'s method (green) often exceed the limits. As for other actuators, it is observed that velocity limits are kept by the proposed method and whole body balance of the robot is maintained successfully.

VI. DISCUSSION

The experiments showed that our method could generate appropriate leg motions for the Aizu-bandaisan dance at a musical speed faster than the original. By combining the Shiratori *et al*.'s [5] method that could generate upper body motion based on musical tempo, we could generate whole body motion that synchronized motion at various musical speeds.

One may question the general validity of our method because we showed results for only one dance performance. We believe that our method is as general as Nakaoka *et al*.'s [4] method. They demonstrated several dance performances with different robots. Our method is an extension of their method and, therefore, can be applied to motions that can be handled by their method. We will conduct further experiments with different dance performances for justification.

In the data we used for experiment, there is only one *stylistic* STEP task, which contains skill parameters for a mid-point. Our investigation led the insight that the trajectory of the stylistic task becomes closer to that of a normal step as the musical tempo. It worked well in our experiment, the case of Aizu-bandaisan dance, but we have to admit that the current results are not sufficient to validate that the insight is always appropriate. We need to collect more data that contain leg motions with special expression. However, for example in Aizu-bandaisan dance, because such stylish tasks appear less than 10 percent of all tasks, this does not critically undermine the generality of the proposed method.

Additionally, we discussed only STEP tasks so far though leg task models include SQUAT tasks as well as STEP tasks (STAND tasks are also included, but they are just considered as the rest of STEP and SQUAT tasks). This is again because we investigated only Aizu-bandaisan dance here. Only one SQUAT task appears in a cycle of the dance and, moreover, the action is very slight. Therefore, from Aizu-bandaisan dance, we cannot extract any general knowledge about SQUAT tasks. Investigation of other dances to improve our system is an indispensable future work.

VII. CONCLUSION

In this paper, we focused on STEP tasks of Nakaoka *et al*.'s leg task models, and proposed a method to achieve temporal scaling of leg motions based on musical speed. Our method is based on the insights obtained in the observation of dancer's motion. The insights are 1) a dancer needs to preserve leg postures that are important to emphasize dance expression, 2) there is a priority to determine what leg motion features can be adjusted, and 3) stylistic leg motion resembles normal step motion if dancers cannot follow fast musical tempo completely. We validated our method via simulation experiments.

For future work, we will design temporal scaling models for SQUAT tasks, and perform more experiments with physical robots. MNSI is obtained when we calculate the longest time interval during performance using a target ZMP trajectory. It is one of our future work to use MNSI obtained by calculation. We also aim at completing the music feedback system that enables robots to listen and dance to varying musical tempo performed by humans in real time.

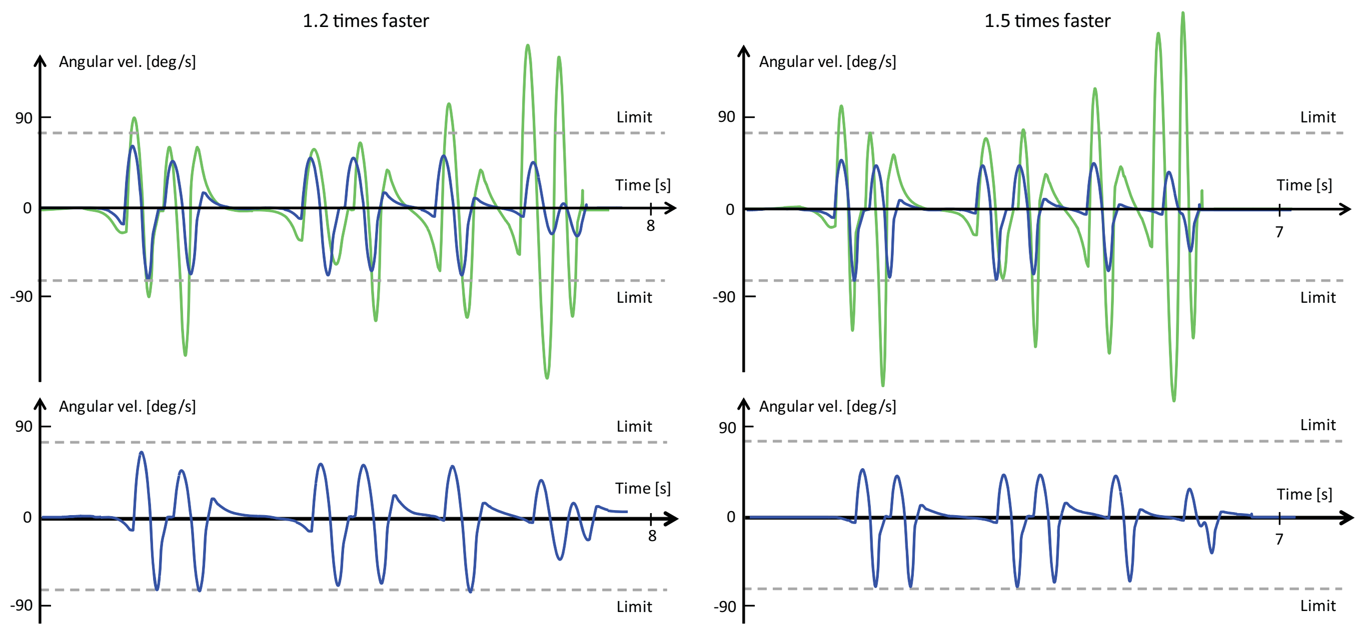

Fig. 12. Velocity sequences of the right knee generated for 1.2 and 1.5 times faster tempos than the original one. Top row: results with fixed upper body posture, and bottom row: results with upper body motion generated by Shiratori *et al.*'s method, and left: results generated for 1.2 times faster tempos, and right: results generated for 1.5 times faster tempos. Blue and green curves represent sequences of joint angular velocity of right knee generated by the proposed method and Nakaoka *et al.*'s method, respectively.

ACKNOWLEDGEMENT

This research is sponsored, in part, by Ministry of Education, Culture, Sports, Science and Technology under Digital Museum Project.

REFERENCES

[1] H. Kozima, M. P. Michalowski, and C. Nakagawa, "Keepon: A playful robot for research, therapy, and entertainment," *International Journal of Social Robotics*, vol. 1, no. 1, pp. 3–18, 2009.

[2] K. Kosuge, T. Hayashi, Y. Hirata, and R. Tobiyama, "Dance partner robot -Ms DancerR-," in *Proc. IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, 2003, pp. 3459–3464.

[3] K. Yamane and J. K. Hodgins, "Simultaneous tracking and balancing of humanoid robots for imitating human motion capture data," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2009, pp. 2510–2517.

[4] S. Nakaoka, A. Nakazawa, F. Kanehiro, K. Kaneko, M. Morisawa, H. Hirukawa, and K. Ikeuchi, "Learning from observation paradigm: Leg task models for enabling a biped humanoid robot to imitate human dances," *Int'l Journal of Robotics Research*, vol. 26, no. 8, pp. 829–844, August 2007.

[5] T. Shiratori, S. Kudoh, S. Nakaoka, and K. Ikeuchi, "Temporal scaling of upper body motion for sound feedback system of a dancing humanoid robot," in *Proc. IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, 2007.

[6] A. Bruderlin and L. Williams, "Motion signal processing," in *Proc. ACM SIGGRAPH 95*, 1995, pp. 97–104.

[7] E. Hsu, K. Pulli, and J. Popović, "Style translation for human motion," *ACM Trans. on Graphics (Proc. ACM SIGGRAPH 2005)*, vol. 24, no. 3, pp. 1082–1089, 2005.

[8] R. Heck, L. Kovar, and M. Gleicher, "Splicing upper-body actions with locomotion," *Computer Graphics Forum (Proc. Eurographics 2006)*, vol. 25, no. 3, pp. 219–227, 2006.

[9] J. McCann, N. S. Pollard, and S. Srinivasa, "Physics-based motion retiming," in *Proc. ACM SIGGRAPH/Eurographics Symp. on Computer Animation*, 2006, pp. 205–214.

[10] K. Yoshii, K. Nakadai, T. Torii, Y. Hasegawa, H. Tsujino, K. Komatani, T. Ogata, and H. G. Okuno, "A biped robot that keeps steps in time with musical beats while listening to music with its own ears," in *Proc. IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, 2007, pp. 1743–1750.

[11] K. Murata, K. Nakadai, K. Yoshii, R. Takeda, T. Torii, H. G. Okuno, Y. Hasegawa, and H. Tsujino, "A robot uses its own microphone to synchronize its steps to musical beats while scatting and singing," in *Proc. IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, 2008, pp. 2459–2464.

[12] K. Ikeuchi and T. Suehiro, "Toward an assembly plan from observation, part I: Task recognition with polyhedral objects," *IEEE Transactions on Robotics and Automation*, vol. 10, no. 3, pp. 368–385, 1994.

[13] J. Takamatsu, T. Morita, K. Ogawara, H. Kimura, and K. Ikeuchi, "Representation for knot-tying tasks," *IEEE Transactions on Robotics*, vol. 22, no. 1, pp. 65–78, 2006.

[14] T. Shiratori, A. Nakazawa, and K. Ikeuchi, "Detecting dance motion structure through music analysis," in *Proceedings IEEE International Conference on Automatic Face and Gesture Recognition (FG2004)*, 2004, pp. 857–862.

[15] S. Nakaoka, A. Nakazawa, and K. Ikeuchi, "An efficient method for composing whole body motions of a humanoid robot," in *Proceedings of the Tenth International Conference on Virtual Systems and Multimedia (VSMM)*, 2004, pp. 1142–1151.

[16] F. Kanehiro, H. Hirukawa, and S. Kajita, "OpenHRP: open architecture humanoid robotics platform," *The International Journal of Robotics Research*, vol. 23, no. 2, pp. 155–165, 2004.

[17] K. Kaneko, F. Kanehiro, S. Kajita, H. Hirukawa, T. Kawasaki, M. Hirata, K. Akachi, and T. Isozumi, "Humanoid robot HRP-2," in *Proceedings of the 2004 IEEE International Conference on Robotics and Automation (ICRA)*, 2004, pp. 1083–1090.