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

Detecting Dance Motion Structure Using Body Components and Turning Motions

Björn Rennhak†, Takaaki Shiratori‡, Shunsuke Kudoh†
Phongtharin Vinayavekhin† and Katsushi Ikeuchi†

Abstract— **This paper presents a novel method for robust dance motion structure detection. In the japanese folk dance domain, teachers created illustrations of dance poses. These poses characterize the most important movements of a dance. So far there is no simple and reliable extraction method which can extract all poses as shown in these drawings. We use these poses for the Task Model (TM) in the context of Learning from Observation (LFO). LFO which is a well known technique for successful human to robot motion mapping, consists of tasks (what to do) and skills (how to do). We propose a novel approach, to extract special motions from a dance, called turning motions useful for skill mapping in the LFO paradigm. Furthermore, we use a modified version of this approach, to detect all poses as shown in the drawings, called turning poses. To achieve this we observe both forearms at the same time and analyze their movement in different 2-D coordinate planes. We evaluate the parameters with and without a weighting function where we minimize acceleration, velocity and power. We successfully demonstrate this novel method using two very different japanese folk dances and discuss further implications of this work in respect to the LFO paradigm and dances of other domains.**

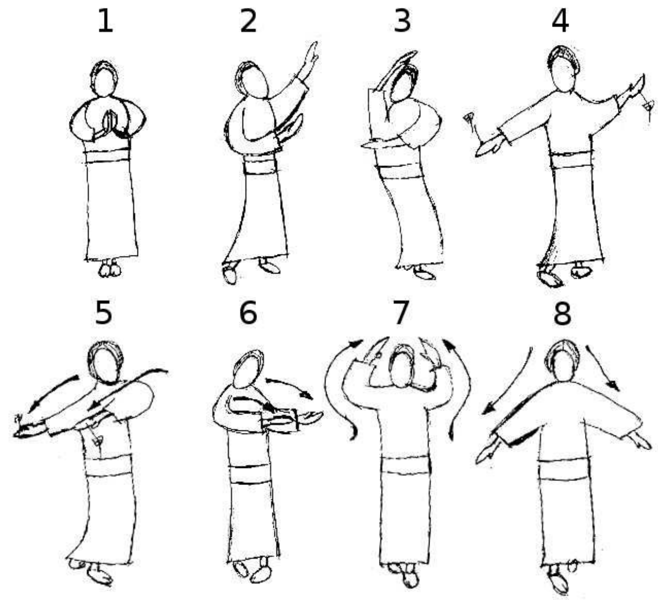


Fig. 1. Traditional japanese folk dance drawings showing one dance cycle by dance master Mr. Tokio Igarashi for *Aizu-bandaisan* created in 2004.

I. INTRODUCTION

In the recent years a steady decline of willing students to learn and to carry on traditional cultural heritage, particularly japanese folk dances, has been a significant problem in Japan. To prevent this loss of heritage, japanese dance teachers created elaborate drawings (*dance master illustrations*), as shown in Fig. 1, to illustrate the most significant poses and movements with annotations to guide the students in the same spirit as labanotation [1]. These important drawings together with the teacher are the basis to actually instruct and communicate students the order and configuration of each pose as well as timing and other information. Due to the importance of instructors in teaching, the fast overaging society has put a strong emphasis on saving cultural heritage by digitalisation and robotic reproduction.

Realizing human dance motions on robots is a challenging topic due to the dynamic and kinemeatic differences and can be overcome by using paradigms such as Learning from Observation (LFO) [2]. In this paradigm, robots observe human actions, recognize and map the detected actions to robot actions in order to mimic them. To recognize human motions the abstract Task Model (TM) is used to separate the motion stream into tasks *(what to do)* and skills *(how to do)*. In LFO terminology this means that tasks are state transitions and skills are the trajectory of the state transitions. This indirect mapping, which we base our work on, can overcome physical differences between humans and robots and allows a successful approximation of a given dance.

In order to detect natural boundaries of human motion (LFO states), Shiratori et al. [3], [4] successfully analyzed motion capture data as well as music information. They coined a concept called *keyposes* which are important stop motions of hand and foot occurring together with the music beat. This method successfully extracts a subset of the given *dance master illustrations* which they call *keyposes*. A problem of this extraction is that it fails to detect and explain all of the illustrations. They furthermore use hierarchical B-splines [5], [6] to preserve characteristic trajectory information (LFO skills) under robotic constraints with different temporal scaling. This method, despite its merits, is non-optimal (data size) as the entire input information is kept and error prone to missing states or larger gaps between states (*keyposes*) when their detection fails. Our recent experiments, as shown in Fig. 2, also indicate the importance of poses beside the *keyposes* during slow dancing and when reproducing a more real looking dance performance.

‡ T. Shiratori is with the The Robotics Institute, Carnegie Mellon University, 4615 Forbes Avenue, Suite 420, Pittsburgh, 15213 USA, E-Mail: siratori@cs.cmu.edu.
† The remaining authors are with the Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo, 153-8505 JAPAN, E-Mail: {br, kudoh, pv-milk, ki}@cvl.iis.u-tokyo.ac.jp

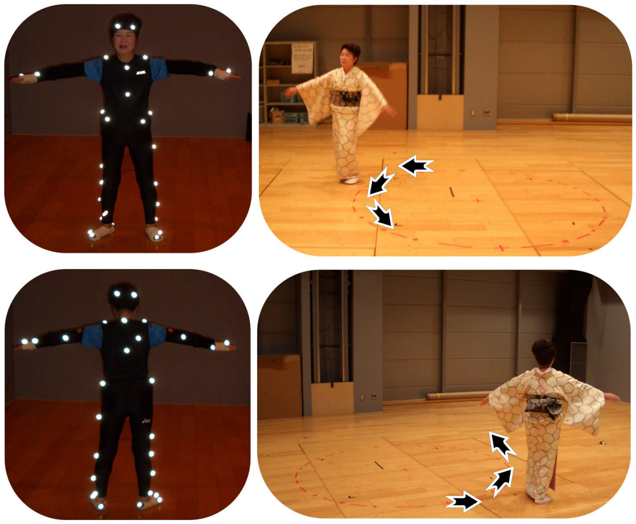


Fig. 2. Experiments, using motion capture equipment Vicon iQ, in Aizu-wakamatsu with *Dance Master* Mrs. Hisako Yamada and her students from 26^{th} to 28^{th} October 2009.

To solve all these problems, without the use of musical information, we introduce in this paper a novel approach based on *turning motions*. This method allows us to successfully detect characteristic poses for trajectory interpolation from state to state (LFO skills). Furthermore, we use a modified version of this approach, to detect all poses as shown in the *dance master illustrations*, which we refer to as *turning poses* (LFO tasks).

II. Related work

Previous works focused on motion analysis [7], [8], [9] and others used musical information [10]. In general there is a difference in the quality of achieving the poses as shown in the *dance master illustrations* (LFO tasks) from dancer to dancer. Shiratori et al. [3] postulates that keyposes help to recognize the skill-level of a dancer. In these works trajectories (LFO skills) are represented by hierarchical B-splines [5], [6]. Although a proper pose extraction for trajectory representation was very difficult, good results could be achieved by preserving the motion data shape. A popular paradigm to map these human motions to robot motions is LFO widely used by Ikeuchi et al. [11] and others [12], [13], [14]. Kovar et al. [15], [16] followed another approach which enabled them to efficiently identify and remove redundant motions for a very large data sets. Other researchers also tried to capture and transfer style [17] while preserving its original content using a translation model. Pollard et al. adapted human motion data to robots by using a PD Data filter [18]. This paper is organized as follows: Section III describes our observations during experiments in Aizu-wakamatsu. Section IV describes our detection method based on *turning motions* and *turning poses*. Section V shows experimental results, and Section VI discusses the implications of this technique. Section VII concludes this paper by mentioning possible future work directions.

III. Observation of Human Dance Motion

A closer study of the poses shown in the *dance master illustrations* which could not be extracted by previous works reveals that they seem to be of a more dynamic, not stopping but slowing down, nature. The dancers slow down at these points a little as they need to perform a more exact motion. This dynamic behavior was indicated by the dance teachers by the arrows in the drawings four to eight as shown in Fig. 1. Further study of dance video and motion capture material of many traditional japanese dances, reveals that these motions excerpt complex elliptical-like arm motions which indicates a turning or critical point in their maximal arm displacement, not occurring at a musical beat. This lack of pose relation to musical information allows us in this paper to omit all use of music data. Furthermore, japanese dances often show a general symmetry in their movement, particularly in the arm movements which makes this finding so interesting.

To further analyze the symmetry and importance of these poses we conducted music and dancing speed varying experiments in Aizu-wakamatsu as shown in Fig. 2. During this experiment we used the 3D movement analysis system (38 markers per dancer) to capture the dance information with varying speeds (0.5x, 0.8x, 1.0x, 1.2x, 1.5x, 2.0x) and four female dancers (two master and two senior level). Resulting from these experiments, we found that the poses of dynamic nature were particularly valuable for slower dancing while on the other hand during faster speeds especially these poses were omitted. From these observations, we obtained the following two insights:

a) Insight: These poses are characteristic for **dance structure** though less important than *keyposes*. When time is short, these poses are sacrificed first to have more time to execute the *keyposes*. When more time is available these poses become equally important.

b) Insight: When these poses are executed a overall higher completeness of a dance reproduction was the result. They convey **style and details** of a given dance.

Based on these insights, we propose a method to detect the dynamic poses of human dance motion.

IV. Turning Motions and Turning Poses

In this section, we propose a method to detect *turning motions* (elliptical motions), based on the acquired insights, when the configuration of the dancers forearms change. This special motion occurs when both arms cross, change its movement direction or orientation. This concept is deceivingly simple, but enables us particularly in the japanese dance domain to capture the important characteristic relationship between body and arm configuration.

We consider all poses as seen in the *dance master illustrations* to be special *turning motions* which we call *turning poses*. In order to successfully extract the poses shown by the *dance master illustrations* we minimize the speed, velocity and the power of the forearms to extract the critical points in their elliptical motion. By extracting these poses, however, we cannot determine which of them essentially **are** *keyposes* or **not** due to the lack of musical information in this work.

Only by using musical beat we can determine if a given pose by a *dance master* is a *keypose* or *turning pose*.

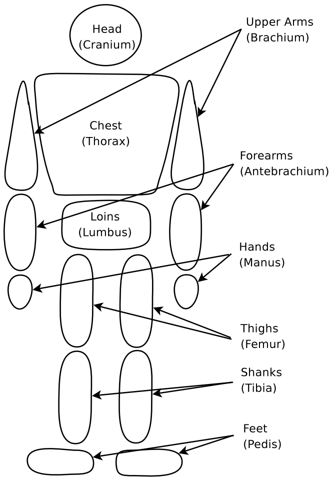


Fig. 3. Different body components, divided into fifteen parts

In order to extract the important body/arm relationship called *turning motions*, we place our local coordinate system in the *umbilicus* (navel, at the level between L3/L4 vertebrae) of the performer as shown in Fig. 4. The main purpose is to utilize the natural body geometry of the person which is commonly known as φ. Furthermore, we use a fixed coordinate system which changes all motions measured from our global to a local system. We are specifically interested in the *X-Y* (*Coronal*), *Y-Z* (*Sagittal*) and *X-Z* (*Transverse*) planes from the navel point.

Instead of looking only at certain points (e.g. hand and foot point) over a time series we based our method on the naturally rigid body components as shown in Fig. 3. This means that certain parts of the human body are fixed in relation to each other by a bone so there is no flexing movement in the components. So span for instance the elbow and the shoulder one component, the upper arm, the elbow and wrist another one, the forearm.

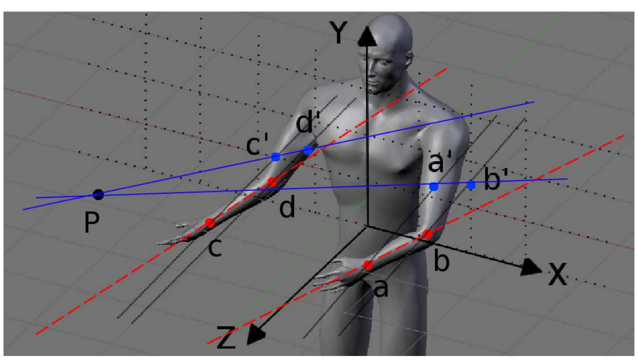
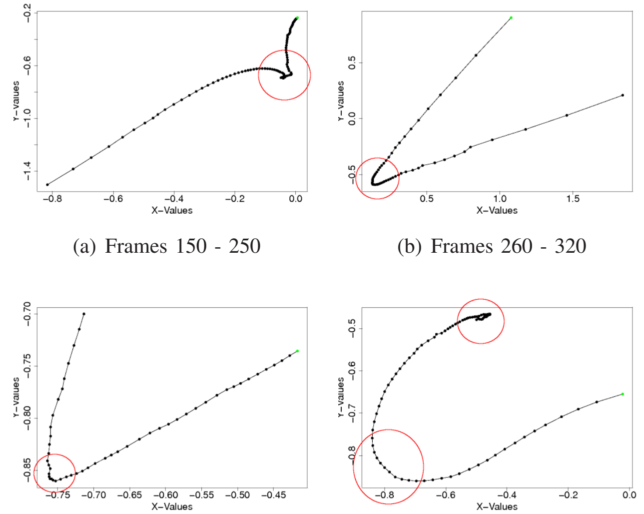


Fig. 4. Simplified visualisation of the turning motion algorithm

We are specifically interested in the forearms (*Antebrachium*) for our method. The limb motion is captured by two markers which are placed at the elbow joint and the wrist joint of the forearms. It is important to point out that this method is not limited to upper body motion and a similar application for the legs (e.g. *tibia*) or other parts is possible.

$$l_n \leftrightarrow \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} = \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix} + t \begin{pmatrix} x_{n+1} - x_n \\ y_{n+1} - y_n \\ z_{n+1} - z_n \end{pmatrix} \qquad t \in]-\infty, \infty[\quad (1)$$



(a) Frames 150 - 250 (b) Frames 260 - 320

(c) Frames 500 - 550 (d) Frames 600 - 700

Fig. 5. Raw, unsmoothed and unfiltered P_n data plots for *Aizu-bandaisan*. Areas of interest are cicled red. Starting point of sequences is marked in green.

Given points $a(x_1, y_1, z_1)$, $b(x_2, y_2, z_2)$ for the left and $c(x_3, y_3, z_3)$, $d(x_4, y_4, z_4)$ for the right arm we establish lines $l_1(a, b)$ and $l_2(c, d)$ through (1) in parametric form as shown in figure 4. Then by an orthogonal projection of the lines l_1 and l_2 we lower them from 3D into 2D space resulting in $l'_1(a', b')$ and $l'_2(c', d')$. We calculate this for all planes $i \in [(x-y), (y-z), (x-z)]$. Lines $l'_1(a', b')$ and $l'_2(c', d')$ intersects in a point $(P_n)_i$ $\forall$ frames n and plane i with the exception when both arms are actually parallel. For simplicity all further references of $(P_n)_i$ are denoted as P_n. The extraction process results in data P_n for each plane, which is highly non linear as seen in Fig. 6.

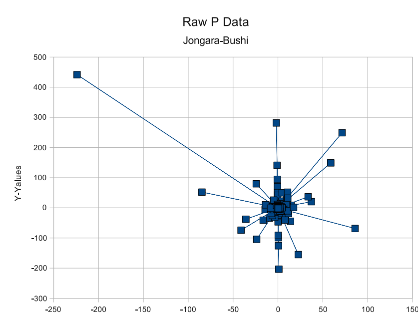


Fig. 6. Raw P_n data plot for *Jongara-Bushi* $\forall$ frames n of the entire dance sequence. Visualized as a radial graph without filtering of parallel arms.

To linearize and to increase robustness of the graphs P_n the following approach is applied. From our observations in section III we established that interesting points (*turning motions*) in a dance may slow down but do not stop like shown in Fig. 5(a) to 5(d). This means that around these critical points a higher, or denser, sampling occurs compared to other parts of the motion. In order to give a higher weight to these distributions we apply an euclidean distance (2) to points of a certain window size α resulting in clustered sample weights w_n (3) for each point P_n. In our experiments we found that a window size of 100 to 160 $[ms]$, $\alpha \approxeq 20$ [points], lead to stable results for all P_n of all three planes.

$$d(p,q) = \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2} \tag{2}$$

$$w_n = \sum_{n=0}^{\frac{\alpha}{2}} (d(P_n, P_{n+1})) + \sum_{n=0}^{-\frac{\alpha}{2}} (d(P_n, P_{n-1})) \qquad \forall n \in [\frac{\alpha}{2}, n - \frac{\alpha}{2}] \tag{3}$$

To extract the poses as shown in the *dance master illustrations* we need to calculate another weighting function which minimizes the linear combination of acceleration a in $[\frac{m}{s^2}]$, velocity v in $[\frac{m}{s}]$ and power P in $[w]$ as shown in equation (4). We will use the weight e_n together with the euclidean distance weight w_n we determined earlier to extract special slow motions where the orientation or direction changes inside the sample window w_n.

$$e_n = \frac{1}{a} \cdot \frac{1}{v} \cdot \frac{1}{P} \tag{4}$$

We set the mass of the arms relative to the whole body proportions with $m = 2.2$ $[kg]$ per forearm and the time t_n in $[s]$ as given in equation (5) with t_{cap} depending on the capture rate in $[s]$ of each dance.

$$t_n = t_{cap} \cdot w_n \tag{5}$$

After solving the formula (4) with known theorems from mechanics we establish (6) and finally (7) after substituting t (5).

$$e_n = \frac{t}{w_n^2} \cdot \frac{1}{m \cdot w_n} \tag{6}$$

$$e_n = \frac{t_{cap}}{w_n} \cdot \frac{1}{m \cdot w_n} \tag{7}$$

Due to the complexity of the forearm motions the radial raw data P_n is as shown in Fig. 6. To simplify this for our purposes we further reduce the dimension by linearizing the data, similarly like a bit matrix, by equation q_n (8). For all further graphs we use the x-values $\forall$ frames $n \in [0, n] \in \mathbb{N}$. This means that one x-axis segment represents one frame for each q_n, M_n or T_n on the y-axis.

$$q_n = |P_n(x_n) \cdot P_n(y_n)| \tag{8}$$

Finally, we can define the *turning motions* $M_n \forall$ frames n as (9) and the *turning poses* $T_n \forall$ frames n as equation (10). We apply an experimentally determined scale factor $\theta = 1000$ to decrease the maximal amplitude.

$$M_n = \frac{e_n \cdot q_n}{\theta} \tag{9}$$

$$T_n = \frac{w_n \cdot e_n \cdot q_n}{\theta} \tag{10}$$

To find a good threshold to determine if a point occurrence in a graph is interesting we used a function to calculate a mean over all dances. It seems that a lower threshold value of $\gamma = 0.05$ is suitable to filter the noise which occurs in M_n and T_n.

TABLE II

EXTRACTION RESULTS FOR *Aizu-bandaisan* IN FRAMES n

DMP	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8
GT	193	345	520	678	808	900	1030	1073
TP	210	311*	526	679	796	907	1031*	1089

V. RESULTS

To evaluate our new method we use two different dance sequences, namely *Aizu-bandaisan* and *Jongara-bushi*. These dances vary greatly in their style, body movement and timing which gives us a good evaluation of the exactness of our algorithm under changing and difficult scenarios. All dances were captured using Vicon Motion iQ System equipment with 38 markers placed on specific places on the dancer as shown in Fig. 2.

To evaluate the poses by a dance master extraction quality we used a 3D Motion Viewer software by S. Kudoh [19] to visually hand pick the exact frames which correspond to the drawings (*dance master illustrations*) best. Consequently these frames are our Ground Truth (GT) to verfy the success of our results. In all tables all poses as indicated by *dance masters* are referred to as *Dance Master Poses (DMP)* and *turning poses (TP)* The motion capture equipment we used in this experiment captured frames in an interval of $t_{cap} = 0.008333$ $[s]$ for the current *Aizu-bandaisan* and $t_{cap} = 0.005000$ $[s]$ for the older *Jongara-Bushi* dance experiments.

Areas where the euclidean distance of the P-data is at its smallest distance, for a one second interval, are treated special and are evaluated on all planes as they might be a *turning motion* or *turning pose* indicator. These results are noted with a „*" in the tables. Just by looking at dense points without taking *turning motions* into consideration for *Jongara-bushi* we can already extract a reasonable amount of all *dance master poses*, which needs further study as it might yield interesting consequences.

A. Jongara-bushi - A Female Dancer

The result of our method is shown in Figure 7. This dance has twelve poses as indicated by the dance master. Previous methods were not able to successfully extract all poses as indicated by the dance teacher. Our method successfully extracts all poses by a dance master and establishes characteristic intermediate poses for LFO task/skill mapping as shown in table I.

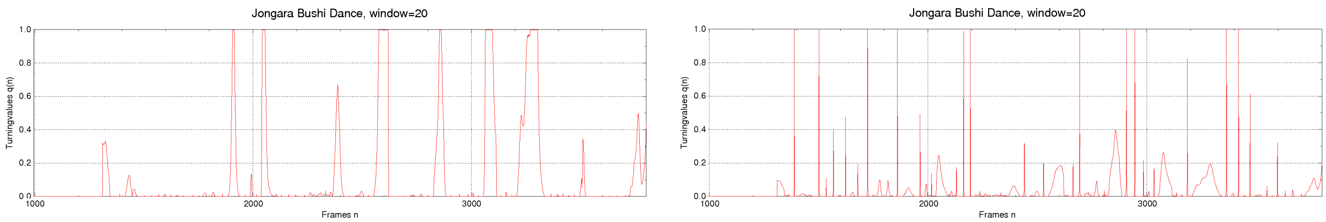
B. Aizu-bandaisan - A Female Dancer

The result of our method is shown in table II. This dance consists of eight poses as indicated by the dance master. Previous methods were able to extract the stop motions correctly (Four from eight, Shiratori et al.) but could not

TABLE I

EXTRACTION RESULTS FOR *Jongara-bushi* IN FRAMES n

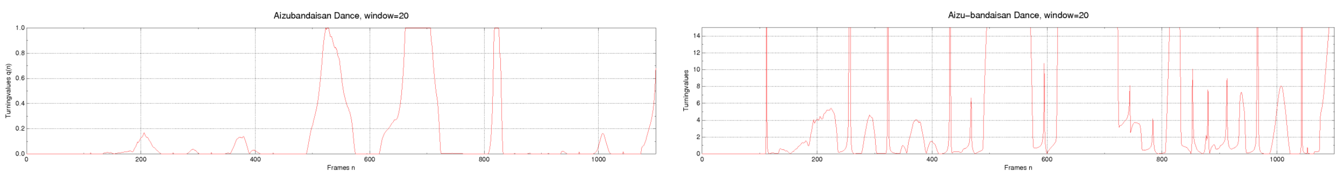
DMP	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}
GT	1319	1441	1610	1912	2119	2448	2608	2881	3088	3295	3560	3756
TP	1311*	1439*	1604*	1922*	2109*	2416*	2618	2858	3088	3272	3501	3759



(a) Linearized and weighted *turning poses* graph T_n (*x-y plane*)

(b) Linearized *turning motions* graph M_n (*x-y plane*) with $\theta = 300$ instead of 1000

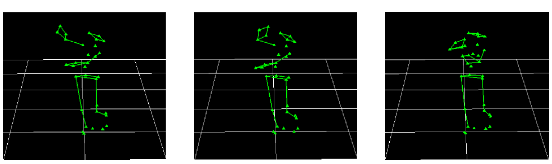
Fig. 7. Extraction results for Jongara-bushi for graph T_n and M_n shown in figure (a) and (b) respectively. Values from 0-1300 have been omitted as they were only used for calibration.



(a) Linearized and weighted *turning poses* graph T_n (*x-y plane*).

(b) Linearized *turning motions* graph M_n (*x-y plane*) without scaling θ and with a higher y-axis graph.

Fig. 8. Extraction results for Aizubandaisan for graph T_n and M_n shown in figure (a) and (b) respectively. Missing pose no. 6 (*dance master illustrations) is extracted from the corresponding y-z plan.*



(a) DMP #3 at the increase beginning, frame 492, $y = 0.06$

(b) DMP #3 at the peak, frame 526, $y > 1$

(c) DMP #3 at the decrease end, frame 573, $y = 0.05$

Fig. 9. Evaluation of the pose by a dance master, no. #3, range as shown in figure V-A, beginning (a), peak (b) and the end (c) for the *Aizu-bandaisan* dance.

achieve success during the dynamic motions. Our method can successfully extract these poses as shown in Fig. 8(a). Futhermore, it is also able to extract characteristic values for LFO skill mapping (peaks between the detected T_n peaks) as shown in Fig. 8(b) for the M_n graph.

Using the results from the weighted graph it is possible to extract the exact frames (graph peaks) but also a certain frame range where the characteristic motion *starts* (c.f. 9(a)) and *ends* (c.f. 9(c)). This is indicated in the graph by the beginning of a increase in y_n, peaking at a pose (c.f. 9(b)) by a *dance master* and falling off until the value is below the noise threshold again. Of course this depends highly on the scaling θ and lower threshold γ of the function M_n and T_n respectively. Results converge quickly to a pose as seen in the *dance master illustrations* for sufficient high values.

VI. DISCUSSION

In the case of *Aizu-bandaisan* for instance creating a Task Model with only *keyposes* and hierarchical B-splines is possible, but not optimal. The main problem lies in the fact that when later motions need modification and significant time scaling. This is specifically troublesome when dances need be performed slower or faster and large gaps between keyposes exist. This is for instance necessary if we want to see a very slow version of the dance to understand the movements in greater detail. This problem can now be avoided by our novel method which helps us to generate a number of characteristic states for trajectory mapping depending on the threshold γ we apply (rough to sensitive). These special states characterize naturally the interesting boundaries between states which will help to re-create dances with a minimal data size and greater detail.

Figure 10 gives more details about the context between *turning poses* T_n which result in a extraction of poses as shown in the *dance master illustrations* and *turning motions* M_n which give us the desired intermediate poses for skill modelling. We consider this approach as promising, due to the robust and successful extraction results for the *dance*

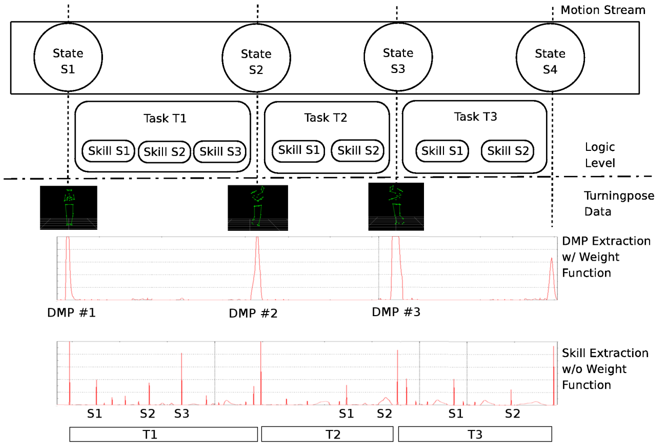


Fig. 10. Learning from Observation paradigm illustrated. States are poses from the *dance master illustrations*, tasks are state transitions from one pose to another and skills represent the trajectory of the state transitions. The use of the Turning pose algorithm in the LFO concept is indicated.

master illustrations. Furthermore, the simple weighting e_n basically just imitates the results from Shiratori et al. [3] by minimizing speed, velocity and power of the *turning motion*. Thus, the extracted intermediate states between the poses by a dance master are considered as important trajectory parameters (LFO skill).

Early experimental results of other dances outside of the japanese dance domain show also promising results. So are for instance DMP extraction ratios for the Macarena dance (latin/flamenco) well above eighty percent, which is a good indicator for the usefulness of this method in other domains. This was tested by using the exact same variables (θ, α, etc.) as with the previous dances.

VII. CONCLUSIONS AND FUTURE WORKS

This paper describes how we extracted Turning Poses from two given dances. We could establish all of our important goals, namely: **(1)** Extract all poses as shown in the *dance master illustrations* by using the *turning pose* method. **(2)** Understand the importance and purpose of the dynamic poses beside existing *keyposes*. **(3)** Our method is robust against different dances in the japanese dance domain. **(4)** Extract natural boundaries, characteristic *turning motions*, useful for intermediate poses to preserve sufficient trajectory information in a minimalistic form. **(5)** The computational complexity of this method is sufficiently low to allow real-time constraints.

Additional development will include testing of our new method with robot and robot simulators. Also we want to find out how strong the impact of turning poses are on a dance for a subjective viewer compared to previous hierarchical B-spline methods. Also a closer analysis and comparison of the dimensional reduction by other means, e.g. Principal Component Analysis (PCA), seems interesting.

Dances outside the japanese dance domain are also worth studying to establish if this method is usable for other dance genres. Recent results with latin dances suggest that also good extraction ratios can be achieved in this new dance domain. All software created for this project has been open sourced and can be found in the internet [20] for further reference.

REFERENCES

[1] A. Hutchinson and D. Anderson, *Labanotation, or Kinetography Laban: the system of analyzing and recording movement*. Phoenix House, London, 1955.

[2] K. Ikeuchi, T. Shiratori, S. Kudoh, H. Hirukawa, S. Nakaoka, and F. Kanehiro, "Robots that learn to dance from observation," *IEEE Intelligent Systems*, vol. 23, no. 2, pp. 74–76, 3-4 2008.

[3] T. Shiratori, "Synthesis of dance performance based on analyses of human motion and music," Ph.D. dissertation, The University of Tokyo, 2006.

[4] 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 *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*, 29 2007-Nov. 2 2007, pp. 3251–3257.

[5] G. W. S. Lee and S. Y. Shin, "Scattered data interpolation with multilevel b-splines," vol. 3, no. 3. IEEE Transactions on Visualization and Computer Graphics, 1997.

[6] J. Lee and S. Y. Shin, "A hierarchical approach to interactive motion editing for human-like figures," vol. 99. ACM Siggraph, 1999, pp. 39–48.

[7] K. Ogawara, "Robot instruction of everyday manipulation tasks from human demonstration," Ph.D. dissertation, The University of Tokyo, 2001.

[8] A. Nakazawa, S. Nakaoka, K. Ikeuchi, and K. Yokoi, "Imitating human dance motions through motion structure analysis," in *In Proc. of International Conference on Intelligent Robots and Systems*, 2002, pp. 2539–2544.

[9] S. B. Kang and K. Ikeuchi, "Toward automatic robot instruction from perception: Recognizing a grasp from observation," *IEEE International Journal of Robotics and Automation*, vol. 9, no. 4, pp. 432–443, 1993.

[10] T. Shiratori, A. Nakazawa, and K. Ikeuchi, "Detecting dance motion structure using motion capture and musical information," in *Proc. International Conference on Virtual Systems and Multimedia*, 2004.

[11] Ikeuchi and Takashi, "Towards an assembly plan from observation: Task recognition with polyhedral objects," 1991.

[12] Y. Kuniyoshi, M. Inaba, and H. Inoue, "Learning by watching: Extracting reusable task knowledge from visual observation of human performance," *IEEE Transactions on Robotics and Automation*, vol. 10, pp. 799–822, 1994.

[13] 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. J. Rob. Res.*, vol. 26, no. 8, pp. 829–844, 2007.

[14] S. Nakaoka, A. Nakazawa, F. Kanehiro, K. Kaneko, M. Morisawa, and K. Ikeuchi, "Task model of lower body motion for a biped humanoid robot to imitate human," in *Dances, IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2005.

[15] L. Kovar, M. Gleicher, and F. Pighin, "Motion graphs," in *SIGGRAPH '02: Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, vol. 21, no. 3. New York, NY, USA: ACM Press, July 2002, pp. 473–482. [Online]. Available: http://dx.doi.org/10.1145/566570.566605

[16] L. Kovar and M. Gleicher, "Automated extraction and parameterization of motions in large data sets," *ACM Trans. Graph.*, vol. 23, no. 3, pp. 559–568, 2004.

[17] E. Hsu, K. Pulli, and J. Popović, "Style translation for human motion," in *SIGGRAPH '05: ACM SIGGRAPH 2005 Papers*. New York, NY, USA: ACM, 2005, pp. 1082–1089.

[18] 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*, 2002, pp. 1390–1397.

[19] S. Kudoh, "Balance maintenance for human-like models with whole body motion," Ph.D. dissertation, The University of Tokyo, 2004.

[20] B. Rennhak, "Source code of the turning poses algorithm," Availalbe under http://github.com/rennhak/keyposes, 2010.