Dance Motion Segmentation Method based on Choreographic Primitives

Narumi Okada[1], Naoya Iwamoto[1], Tsukasa Fukusato[1] and Shigeo Morishima[2]
[1]Waseda University, JST, Tokyo, Japan
[2]Waseda Research Institute for Science and Engineering, JST, Tokyo, Japan

Keywords: Choreographic Primitives, Motion Segmentation, Motion Capture, Dance Motion.

Abstract: Data-driven animation using a large human motion database enables the programing of various natural human motions. While the development of a motion capture system allows the acquisition of realistic human motion, segmenting the captured motion into a series of primitive motions for the construction of a motion database is necessary. Although most segmentation methods have focused on periodic motion, e.g., walking and jogging, segmenting non-periodic and asymmetrical motions such as dance performance, remains a challenging problem. In this paper, we present a specialized segmentation approach for human dance motion. Our approach consists of three steps based on the assumption that human dance motion is composed of consecutive choreographic primitives. First, we perform an investigation based on dancer perception to determine segmentation components. After professional dancers have selected segmentation sequences, we use their selected sequences to define rules for the segmentation of choreographic primitives. Finally, the accuracy of our approach is verified by a user-study, and we thereby show that our approach is superior to existing segmentation methods. Through three steps, we demonstrate automatic dance motion synthesis based on the choreographic primitives obtained.

1 INTRODUCTION

The recent digitization of multimedia content has evolved streaming technologies, and the number of character animations has been increasing. Among them, dance animation has been attracting world-wide attention, because of their sophistication and artistry. Character dance animation is created by two main approaches as follows.

The first approach seeks to interpolate key-frames, i.e., a character's fundamental postures. This method enables the creation of high-quality animation by editing character postures. However, creating character dance motion by this approach is difficult, because it requires a match of the rhythm between music and motion.

The second approach seeks to capture a dancer's dance motion using a motion capture system. This method enables a smooth and high-fidelity representation of character dance motion. However, the method requires the added cost of a physical image capturing space. Moreover, dancer motion must be recaptured to create a new dance animation. Thus, this approach is costly and time-consuming.

From these considerations, an approach that creates realistic dance motion animation automatically and efficiently by reusing the available dance motion data is highly desirable. Such an approach could proceed by connecting dance motion segments that have been segmented from a series of previously captured dance motions; thus, a high-quality new dance motion can be easily generated.

Traditional human motion segmentation methods have focused on the periodic motion or angular velocity of each joint. However, these methods are not applicable to non-periodic and asymmetricaly motions such as a dance motion. In particular, a dance motion is composed of various elements of choreography, which we call "*choreographic primitives*" in this paper. Generally, professional human choreographers create an entire dance performance by combining choreographic primitives. Dividing a dance motion into choreographic primitives is expected to improve the quality of the dance motion synthesized by a data-driven approach.

In this study, our goal is to segment dance motion into the framework of choreographic primitives. First, we perform a perception-based

Okada N., Iwamoto N., Fukusato T. and Morishima S..
Dance Motion Segmentation Method based on Choreographic Primitives.
DOI: 10.5220/0005304303320339
In *Proceedings of the 10th International Conference on Computer Graphics Theory and Applications* (GRAPP-2015), pages 332-339
ISBN: 978-989-758-087-1

investigation using professional dancers to accurately detect the segmentation points of a particular dance motion. Then, we define rules for detecting the boundaries of choreographic primitives based on the results of the initial investigation. Finally, we verify the accuracy of our approach by a user-study. Our approach seeks to enable automatic, unsupervised dance motion segmentation according to the evaluated motion and specific features of the accompanying music. We further seeks to improve the accuracy of the detection of choreographic primitives.

The remainder of this paper is organized as follows. In Section 2, we review related work. In Section 3, we discuss the factors inherent in the segmentation process based on choreographic primitives. In Section 4, we describe the main ideas underlying the algorithms used in the proposed method. Section 5 presents the results. In Section 6, we conclude the paper and discuss limitations and future work.

2 RELATED WORK

Accompanying the continuous increase in the abundance of computer graphic content, numerous studies have considered the reuse of the existing motion data. In particular, automatic motion generation systems have been researched by Kovar et al. (2002), Arikan et al. (2002), and Beaudoin et al. (2008). These researches have sought to calculate the similarity of postures in motion databases and to connect the postures whose values are high, thereby creating new natural motions.

Recently research on automatic motion generation, especially for a dance motion have been proposed by Kim et al. (2003), Alankus et al. (2005) and Shiratori et al. (2006). These researchers have attempted to generate a new series of dance motions by connecting the dance motion segments derived from a database. Considering the temporal information of motion for synchronizing to the input music is necessary. To detect beats from the dance motion data, Kim et al. (2003) and Alankus et al. (2005) used the moments of directional changes of the motion. Shiratori et al. (2006) detected the local minimum points of *Weight Effort* which is the sum of the absolute angular velocities of the joints. They inspired by a theory developed by Laban et al. (1971). This theory is based on an impression of a human motion from the standpoint of *effort* and *shape*; *effort* represents the power of the movements, and *shape* represents the silhouette of the movements. Furthermore, Nakata et al. (2002) determined that relation between the strength of a motion and human feelings was more strongly correlated with *effort* than *shape*. Then, using the minimum points of *Weight Effort*, considering the stopping points of the bodily motion easily is possible.

From this viewpoint, the fundamental step is to collect motion segments in the form of a database to enable the reuse of the motion data and to create a new series of motions.

To construct a database, numerous studies have been performed for the detection and segmentation of motion images into a series of some type of a motion primitive. Osaki et al. (2000) proposed a method based on joint velocities. Barbic et al. (2004) proposed PCA (Principal Component Analysis) based method. Zhou et al. (2008) and Zhou et al. (2013) proposed a method based on an extension of kernel k-means clustering and Vögele et al. (2014) proposed the methods based on the positions of the head, wrists and ankles.

These methods are valid for cyclic motions such as walking; however, in dance motions, because of the complex posture and noncyclic motions, these methods are not applicable.

Dance motion segmentation frameworks have been mainly based on musical beats. Kim et al. (2007) and Fan et al. (2012) focused on the feature that dance choreography is made up of motions synchronized with the music and proposed a method that used the musical tempo. Lee et al. (2013) divided a dance motion based on points where the musical features change. However, to generate variety of a new series of dance motions, segmenting the existing dance motion into choreographic primitives is necessary. The length of each choreographic primitive would generally be different and the choreography often differs even if the same melodies exist in the music. Therefore, an accurate segmentation is difficult task without including motion information.

To detect choreographic primitives from a series of dance motion, Rennhak et al. (2010) used the acceleration, velocity and power of each joint. To divide a dance motion with a consideration for the time series, Masurelle et al. (2013) suggested a method segmenting the point when footstep impacts are detected, for the dance which has features in footsteps like salsa. These methods consider only motion information. Therefore, these are valid for the novice dancers' motions, because dancers who are not an experimented dancer will not be accurately synchronized in time with the music.

However, without incorporating musical information, when attempting to create a new motion, the output dance motion is not always matched to the musical beats.

To segment a dance motion into basic temporal primitives, Shiratori et al. (2004) used the velocities of the body parts in coordination with musical information. This method, succeeded in gaining a high-accuracy recognition rate of segmentation points. However, the segmentation rules were made up for a Japanese traditional dance, and the time series was not considered. Generally, a dance cannot always be segmented according to the timing when the body's velocities change, but the timing is decided by the time-series of the choreography.

Therefore, a segmentation method that can consider the musical tempo and the primitives of a dance motion and time series of motions has not yet been developed. Here we assume that a dance motion is made of segments of choreographic primitives that are synchronized with music. Therefore, to accurately segment a dance motion, considering the primitives of the choreography, while simultaneously accommodating the musical tempo would be natural. Owing to this segmentation, a natural stream of a dance motion is expected to be created by reusing the existing dance motion data.

3 INVESTIGATION

In this section, we discuss the perception-based investigation we performed using dancers to determine choreographic primitives. On the basis of the results, we defined rules for dance motion segmentation with a consideration for choreographic primitives and musical information and improved the quality of the dance motion composed of the existing motion data.

3.1 Subjective Experiment

For dance motion animation, we employ 28 character joints, as shown in Figure 1. We use 1 motion data whose length is about 1 minute captured from a professional female dancer with music synchronized the motion from the web page http://perfume-global.com/project.html.

Generally, professional human dancers create the entire dance motion synchronized to 8 counts (2 bars of music). Here, we assumed that the boundaries of the choreographic primitives occur at the musical beats. We recruited 5 professional dancers(2 men and 3 women) as participants, and asked them to assign the counts wherein the choreographic primitive boundaries occurred from the 112 musical beats. The dance motion is analysed using skinned mesh animation, because evaluating choreographic primitives based only on skeletal animation is difficult.

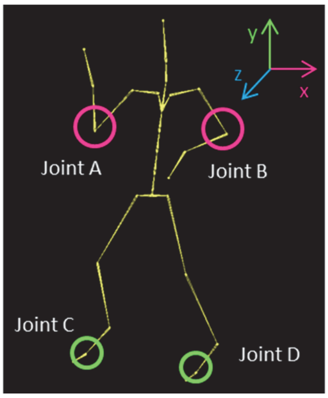

Figure 1: Skeletal model based on 28 joints. For the perception-based investigation, we focused on elbows and toes (Section4).

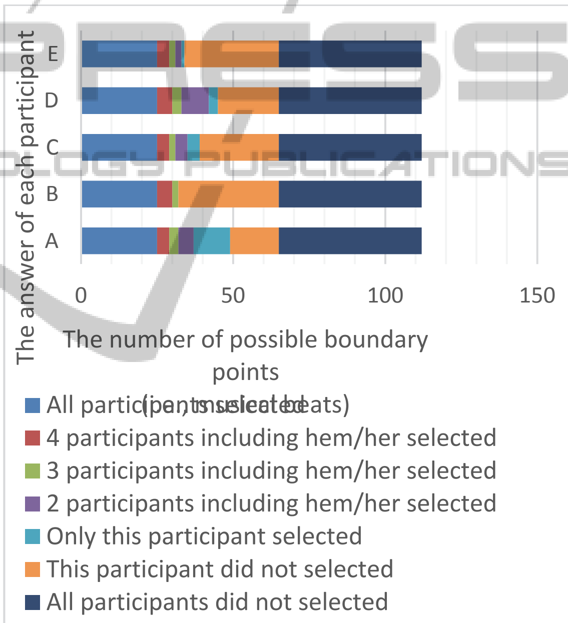

Figure 2: The results of the investigation. 5 professional dancers selected segmentation points from 112 possible boundary points (i.e., musical beats). The participants A, B are male, and C, D, E are female.

3.2 Results of the Investigation

Figure 2 shows the results of the subjective experiment, in which the number of each participant selecting a particular boundary point (i.e., musical beat) is represented. The results suggest the existence of an empirical rule for the selection of choreographic primitives because all participants selected 25 points as choreographic primitive boundaries and none of the participants selected 47 points.

In this paper, we focus on some of the most significant factors for dance motion segmentation, rather than on the individual variations of the answers provided. Therefore, we assumed that those beats selected by all dancers were the accurate segmentation points representing the choreographic primitive boundaries.

3.3 Extracting the Important Factors

We examined the tendencies of the segmented points that all dancers had selected as choreographic primitive boundaries. Based on this examination, we determined that the segmentation timing was caused by the three factors as follows; musical beats, motion symmetry, and the intervals of footsteps. We focused on not only music-based detection, but also the motion features. In doing so, we determined that two types of choreographic primitives were in evidence. The first type reflected symmetry movement, whereas the second type represented asymmetry movement, as shown in Figure 3. If we take note of footsteps, each choreographic primitive consisted of footsteps whose intervals were of the same length, as shown in Figure 4.

Figure 3: Examples of symmetrical dance postures (left) and asymmetrical postures (right) during a dance performance.

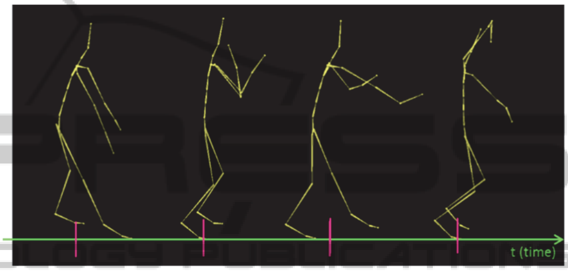

Figure 4: Examples of a single segment. Each footstep has equal time intervals.

4 DANCE MOTION SEGMENTATION

Based on the factors indicated by the investigation (Section 3), we defined the rules for the segmentation of choreographic primitives. In this study, the rules consist of the three factors: musical beats, motion symmetry and the intervals of footsteps. Then, we suggested a method for obtaining segmentation points automatically. When we performed investigation (Section 3), we use skinned character, though we use stick figures in this section.

4.1 Musical Beats

A dance motion is generally synchronized with music; therefore, musical information is essential for acquisition of choreographic primitives. As observed in the previous investigation (Section 3), 8 counts of the musical beats are particularly important because when choreographers create a dance motion, they are usually aware of the 8 counts. We therefore calculated the frames of 8 counts, and all the calculated 8 counts are segmentation points. In this study, the input music tempo is known.

4.2 Motion Symmetry

To detect segmentation points based on changes from symmetrical (asymmetrical) to asymmetrical (symmetrical) motions in a dance motion series,

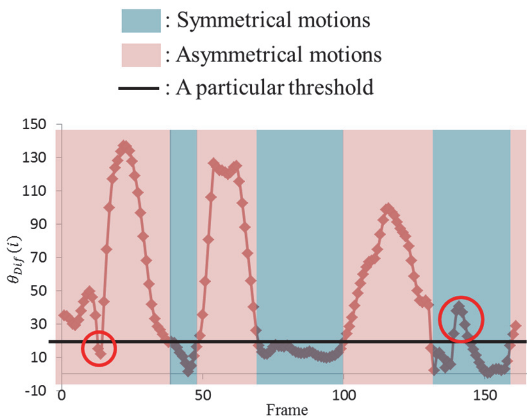

Figure 5: Detecting symmetrical motions. The red circles are regarded as false detection because the length of sequences are less than 7 frames.

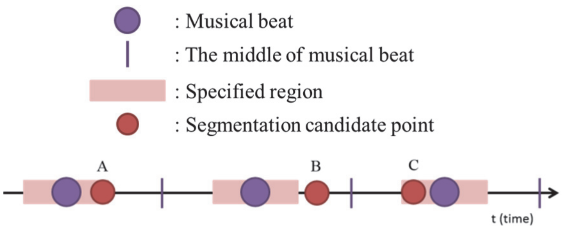

Figure 6: Refinement of the point candidates using musical beats. Candidate point A and C is shifted to the nearest musical beat, because these are located within a specified region near the beats.

knowing whether the motion is symmetrical is, of course, necessary. However, detecting symmetrical motions based on posture is difficult because some motions appear to be symmetrical, but are not symmetrical in reality. For example, some right and left legs movements are similar in a time series, although they are stepping alternately. Then, we focus on the elbows, because we assume that the shape of the elbows are important when we judge the symmetry of figures. Unlike the legs, arms have many symmetrical movements by frames, and the degree of freedom is one, except for the twist. Therefore, detecting symmetrical motions easily is possible. We calculated the difference between the degrees of the right and left elbows, $\theta_{Dif}(i)$, which is given as follows:

$$\theta_{Dif}(i) = |\theta_R(i) - \theta_L(i)| \tag{1}$$

where, $\theta_R(i)$ is the degree of the right elbow and $\theta_L(i)$ is that of the left elbow in the i-th frame. If $\theta_{Dif}(i)$ is under a particular threshold, the posture in the i-th frame is regarded as symmetrical.

To avoid the false detection of symmetry, we defined a minimum choreographic primitive length of 7 frames because some frames happened to be judged as symmetrical on the way of the choreographic primitive. Setting a minimum length allows us to better consider the temporal continuity. In accordance with these considerations, the points at which a motion changes from symmetrical (asymmetrical) to asymmetrical (symmetrical) are regarded as segmentation candidate points. Figure 5 shows an overview of this algorithm.

4.3 Intervals of Footsteps

We detected footsteps on the assumption that if the length of frames between one footstep and the next is stable, then these footsteps are closely associated with a single choreographic primitive.

To detect footsteps, we use the velocities of foot joints (joints C and D in Figure 1) based on Shiratori et al.'s (2004) study because foot velocity is nearly zero when the foot is on the ground. However, using velocity information alone results in many false detections, because the motion capture data suffers from numerous fluctuations

To avoid these errors, we incorporate some constraints for foot coordinate detection.

- To condition of touching the ground is assumed when the foot velocity is less than the threshold value.
- The condition of leaving the grounds is assumed when the foot acceleration is greater than the threshold value.
- If the Y coordinate (as shown in Figure 1) is under the threshold value, the foot is judged as touching the ground, even if it was judged as leaving the ground by the other constraint.

In a dance motion, footsteps are usually synchronized with the musical beats, and we use the moment when the foot is about to touch the ground as a candidate frame.

4.4 Refinement of Candidates using Musical Beats

The candidate segmentation points based on motion symmetry and footsteps (Sections 4.2 and 4.3) are not always synchronized with the musical beats. Therefore, we refine candidate points based on choreographic primitive boundaries that are synchronized to musical beats. We employ the method developed by Shiratori et al. (2004) to shift the candidate points based on musical beats. If a musical beat is located within a specified region near a candidate point based on motion symmetry or footsteps, we shift the candidate point to the beat. However, if no musical beats are observed within the specified region, we regard the boundary point as a false detection. If no musical beats are observed in the specified region around the candidate point based on footsteps, we use the frame as it is. Thereafter, we calculate the interval of the next footstep at all the detected frames, if the interval has the same length as the next interval, the area around these footsteps is regarded as belonging to a single segment. Although there are segmentation candidate points in the area, we revise the points as false detections. In our method, the length of the specified region is set to one half of the musical tempo. Figure 6 shows an overview of this algorithm.

5 RESULTS

5.1 Results and Verification of Our Method

The segmentation candidate points detected by the musical beats and symmetry movement are shown in Table 1. The metrics of the evaluation were the "*Recall Rate*" and "*Precision Rate.*" In accordance with the result of the investigation report in Section 3, the frames on 8 counts are always segmented; thus, the result of using the musical beats shows 100% precision rate. However, the recall rate is still low, that is, the choreographic primitives are not detected.

For the detection of the intervals of footsteps (Section 4.3), all detected points are successfully placed in the same segment. Using only footsteps, some points are not in the same segment. Then, by prioritizing the musical beats, we succeeded in improving the segmentation accuracy.

To verify the accuracy of our method, we compared the results of our method with those of the two existing methods. One is the most popular segmentation method for dance motions that uses the musical information of 4 beats (Kim et al. (2007) and Fan et al. (2012) used). The other is proposed by Shiratori et al. (2004) that uses motion velocities and musical information.

As the dataset, we use the same dance motion which we used in Section 3 consisting of 25 choreographic primitives (decided by the 5 professional dancers reported in Section 3.2). To compare the accuracy of the results, we revised the detected segmentation points by existing methods using the musical beats. Table 2 shows the respective segmentation results.

The results indicate that our method attended a higher accuracy than the existing methods. The recall rate of our method was 80.0%, which is a sufficient accuracy for segmentation. Although the precision rate is still 62.5%, the results of our method represent an improvement, owing to the consideration of the motion symmetry and footstep intervals. This will lead to an improved generation of the new series of a dance motion. The results of the method proposed by Shiratori et al. (2004) indicate low recall and precision rates, because the using motion data was not Japanese dance and the boundaries of choreographic primitives are not always placed at the changing points of body velocities.

Table 1: Accuracy of the candidate points.

Methodology	Recall[%]	Precision[%]
Musical 8 beats	60.0	100
Motion symmetry	40.0	40.0

Table 2: Verifying the accuracy of our method.

Methodology	Recall[%]	Precision[%]
Musical 4 beats	60.0	50.0
Shiratori et al. (2004)	4.0	7.7
Our method	80.0	62.5

When we focus on the false detection of segmentation points, 12 normal beats were detected, 3 points are one count later than the accurate points, and 1 point was the timing for which 4 of the 5 dancers had segmented. Few segments whose lengths are the same as one count are expected to exist. If two such points reside next to each other, only one of the two should be selected. Other detected points that were selected as accurate points by 5 dancers appear to represent a dancers' individual judgement.

5.2 Synthesizing Dance Motion using the Proposed Segmentation Method

We also performed an experiment to verify that our method is valid for reusing motions through a user-study, because we assumed that the appropriate segmentation of the dance motion enables the creation of a new series of dance motions with high-quality. We recruited 12 participants consisting of 10 men 20-30 years of age and 2 women 20-25 years

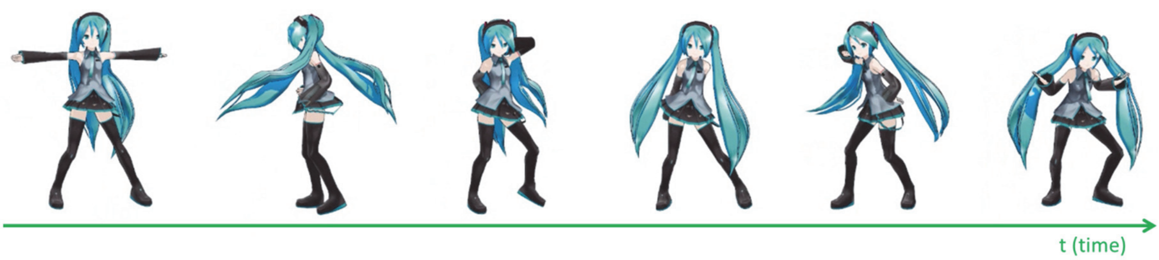

Figure 7: Example of the generation of a new series of dance motion.
You can watch the examples of a new series of dance motions in the demo as the paper attachment.

Table 3: Results of the evaluation experiment for the automatic creation of a new dance motions.

Segmentation Method	Percentage[%]
Musical 4 beats	33.3
Our method	66.7

of age. We showed them a series of dance motions formed by connecting the short motion segments detected by proposed method and musical beats method wherein segmentation is regularly performed every four counts. The participants were asked to select what they considered the better of the two resulting motions. The segments were selected at random to do justice to both segmentation methods. Figure 7 shows as example of a generated new series of dance motions.

Table 3 shows the results of the experiment. The results show that the 66.7% of the participants felt that the dance motion based on the proposed approach provided more natural dance transitions, because of the choreographic primitives used.

Some participants indicated dissatisfaction with the dance motions based on the musical beats method, for the reason; the character changes to another motion in the middle of a choreographic primitive. Therefore, when we generate a new series of dance motions, the segmentation phase is highly significant, which reflects the general appreciation shown in the experiment for the results of our method.

6 CONCLUSIONS

In this paper, we proposed a segmentation method for a dance motion that established choreographic primitives based on an investigation into the perceptions of actual dancers. We defined segmentation rules based on the musical beats, the symmetry of motion, and the timing of footsteps. Higher accuracy percentages were obtained in our method than those in existing methods.

In the proposed method, if the music at times deviates from the standard 8 counts, the existing 8 counts cannot be detected accurately; e.g., in case of songs that have four counts before their chorus. However if the timing of the beginning of the chorus is input, irregular counts can be detected. Therefore, we intend to apply the proposed segmentation method to any genre of music. Though if the music has except for 4 counts in its bar, the way of making the choreography may be different. To apply our method for these music is our future work.

The segments were sorted randomly when comparing the new series of dance motions generated by the dance segmentation methods considered. If we incorporate sorting rules, generating more natural dance transition is possible. On a future work, we plan to construct sorting rules based on similarities of posture and to synchronize a dance motion with the atmosphere of the input music. Finally, we intend to construct an automatic dance motion generation system based on the accurate segmentation of the dance motion data collected from the Internet.

REFERENCES

Alankus, G., Bayazit, A. A., and Bayazit, O. B., 2005. Automated Motion Synthesis for Dancing Characters. *Computer Animation and Virtual Worlds*, 16(3-4), 259-271.

Arikan, O., and Forsyth, D., 2002. Interactive Motion Generation from Examples. *ACM Transaction on Graphics*, 21(3): 483-490.

Barbic, J., Pan, J. Y., Faloutsos, C., Hodginsj, K., and Pollard, N., 2004. Segmenting motion capture data into distinct behaviors. *In Proc. of Graphics Interface*, 2(6): 185-194.

Beaudoin, P., Coros, S., Panne, M., and Poilin, P., 2008. Motion-motif graphs. *In Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, 117-126.

Fan, R., Xu, S., and Geng, W., 2012. Example-based Automatic Music-Driven Conventional Dance Motion Synthesis. *IEEE Transactions on Visualization and Computer Graphics*, 18(3): 501-515.

Kim, G., Wang, Y., and Seo, H., 2007. Motion Control of a Dancing Character with Music. *In Proceedings of the 6th IEEE/ACIS International Conference on Computer and Information Science (ICIS)*, 930–936.

Kim, T.H., Park, S.I., and Shin, S.Y., 2003. Rhythmic Motion Synthesis based on Motion-Beat Analysis. *ACM Transaction on Graphics*, 22(3): 392-401.

Kovar, L., Gleicher, M., and Pighin, F., 2002. Motion Graphs. *ACM Transaction on Graphics*, 21(3): 473-482.

Laban, R., and Ullmann, L., 1971. The Mastery of Movement. *Pre Textos.*

Lee, M., Lee, K., and Park, J., 2013. Music Similarity-based Approach to Generating Dance Motion Sequence. *Multimedia Tools and Applications,* 62(3): 895-912.

Masurelle, A., Essid, S., and Richard, G., 2013. Multimodal Classification of Dance Movements using Body Joint Trajectories and Step Sounds. *In Proc. 14th International Workshop on Image and Audio Analysis for Multimedia Interactive Services (WIAMIS)*, 1-4.

Nakata, T., Mori, T., and Sato, T., 2002. Analysis of Impression of Robot Bodily Expression. *The Journal of Robotics and Mechatronics*, 14(1): 27-36.

Osaki, R., Shimada, M., and Uehara, K., 2000. A Motion Recognition Method by Using Primitive Motions. VDB 5 Proceedings of the Fifth Working Conference on Visual Database Systems: Advances in Visual Information Management, 117-128.

Rennhak, B., Shiratori, T., Kudoh, S., Vinayavekhin, P. and Ikeuchi, K., 2010. Detecting Dance Motion Structure Using Body Components and Turning Motions. *In Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2010)*, 2264-2269.

Shiratori, T., Nakazawa, A., and Ikeuchi, K., 2004. Detecting Dance Motion Structure through Music Analysis. *In Proc. IEEE International Conference on Automatic Face and Gesture Recognition (FGR 2004),* 857-862.

Shiratori, T., Nakazawa, A., and Ikeuchi, K., 2006. Dancing-to-Music Character Animation. *Computer Graphics Forum*, 25(3): 449-458.

Vögele, A., Krüger, B., and Klein, R., 2014. Efficient unsupervised Temporal Segmentation of Human Motion. *In Proceedings of the ACM SIGGRAPH / Eurographics Symposium on Computer Animation*.

Zhou, F., Torre, F., and Hodgins J. K., 2008. Aligned cluster analysis for temporal segmentation of human motion. *In IEEE Conference on Automatic Face and Gestures Recognition*. 1(2).

Zhou, F., Torre, F., and Hodgins J. K., 2013. Hierarchical aligned cluster analysis for temporal clustering of human motion. *IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI)*. 35(3): 582-596.