Using Music and Motion Analysis to Construct 3D Animations and Visualizations

Kuen-Meau Chen*, Siu-Tsen Shen **, Stephen D. Prior***

* Visualization & Interactive Media Lab, National Center for High-Performance Computing, Taiwan.
E-mail: morris@nchc.org.tw
** Department of Multimedia Design, National Formosa University, Taiwan.
Email: stshen@nfu.edu.tw
*** Department of Product Design and Engineering, Middlesex University, UK.
E-mail: s.prior@mdx.ac.uk

Abstract

This paper presents a study into music analysis, motion analysis and the integration of music and motion to form natural human motion in a virtual environment. Motion capture data is extracted to generate a motion library; this places the digital motion model at a fixed posture. The first step in this process is to configure the motion path curve for the database and calculate the possibility that two motions were sequential through the use of a computational algorithm. Every motion is then analyzed for the next possible smooth movement to connect to, and at the same time, an interpolation method is used to create the transitions between motions to enable the digital motion models to move fluently. Lastly, a searching algorithm sifts for possible successive motions from the motion path curve according to the music tempo. It was concluded that the higher ratio of rescaling a transition, the lower the degree of natural motion.

Keywords: *multimedia, choreography, rhythm analysis, motion analysis, computer animation.*

1. Introduction

To display natural motion of the human body is a very important topic in computer graphics. In recent years, by using computer animation techniques and combining with captured data of human movement, a virtual character can interact with a real person in real time. In the meantime, by taking advantage of the large amount of data generated from motion capture, it is more worthy for researchers to devote their efforts to analyze those data. For example, to re-edit or re-use motion capture data is a value-added application. Brand & Hertzmann created a statistical model named "style machine" which can generate new motion sequences (Brand & Hertzmann 2000).

Generally speaking, computer animation has two major parts, namely imaginative motion and performance motion. Imaginative motion is focused on cognitive thinking in terms of the intended expression, whilst performance motion emphasizes the meaning of the message that has been taken by execution. Therefore, the job of computer animation is not only to make graphics come alive, but also to convey emotional feeling through certain characters or motions. The elements in the presentation of animation include objects, motion, language, backgrounds, music, etc. Based on the request of narratives and scenarios, these visual images are all created by animation designers. Through sketches, digital recording and retrospective techniques, personalized characters can reach a certain level of accuracy and emotional expression in the development of computer animation. The relationship between emotional expression through motion and the tempo of music is closely associated, especially motion in design. In this research, it attempts to understand the interrelationship between choreography and the tempo of music through adjustment and modification. Therefore, it is interesting to find variations on the performance of subjects, under the emotional reflection through music in the design of character-based motion.

In this research, we have attempted to modify or adjust music tempo to understand the relationship between choreography and music tempo. We believe the emotion created by music profoundly influences motion design. Through the analysis of music and digital motions, this study tried to help motion creators conceive the composition of music rhythm and digital motions. It will help motion creators create a series of meaningful motions using music segments. Music performance is the combination of tempo, key, rhythm, pitch, melody and harmony. In the study, we brought up the relationship between music and motion. Through the expression of music, we generated different digital motion models accordingly.

The remainder of this paper is organized as follows: We first review related work in Section 2. Then we cover the system overview structure in Section 3. In Section 4, we describe how to carry out music analysis. Section 5 describes how we perform motion synthesis from the motion capture data and music analysis. We show the experimental results in Section 6. Finally, Section 7 concludes this paper.

2. Related Work

The research on the relation between music and motion can be categorized in two ways. On the other way, there are also applications that make use of motions to create melody reversely (Bevilacqua et al. 2001). Another research had focusing on correlating motions through the meaning of music, they analyzed dance structure with the assistance of music in order to digitize dance, which could help preserve the culture (Shiratori et al. 2004).

In computer animation, motions of characters have frequently been linked with music in order to convey various emotions. The synchronization of an animated image sequence with sounds helps deliver an artistic theme to the audience (Laybourne 1998). For most people, the salient point between music and dance is their tempo. Moreover, in the field of virtual choreography, many researches linked up music and motion with the tempo. However, extracting the tempo structures from rhythmic signals such as drum sounds and music is a difficult problem. There has been some work on musical beat detection. Therefore, we referred from previous work and adopted their algorithm to extract the dance tunes, which, has obvious accents.

Harper and Jerniga used recurrent timing networks to detect and predict periodicities in an onset stream (Harper et al. 2004). Hainsworth and Macleon first extracted musical change-points from the help signal and then use a particle-filtering algorithm to associate these with a tempo process (Hainsworrh et al. 2003). Simon Dixon proposed a robotic but more complicated method to detect tempo from either digital audio or MIDI files (Dixon 2001). Musical data is processed off-line to detect the salient rhythmic events and the timing of these events is analyzed to generate hypotheses of the tempo at various metrical levels. Based on these tempo hypotheses, a multiple hypothesis search finds the sequence of beat times which has the best fit to the rhythmic events. The main work in the beat tracking of audio data is done by Goto and Muraoka. They developed two beat tracking systems for popular music, the first system is for music containing drums and the second is for music without drums (Goto, M. 2001).

Motion synthesis researches could be divided into two categories. One approach relies on using motion simulated actions of body's joints. Since the human body's actions are very complex, it is very difficult to create natural movements from the degree of freedom information in each joint (Fang et al. 2003, Liu 2002, Popovic 2003). In addition, some research makes an approach of synthesizing from available motion database. Motion graph and motion synthesis from annotations have had success in animation synthesis, but the range of synthesized motion is still limited by the motion database.

In the aspect of motion graph, it is difficult to generate the realistic human motion sequence. To synthesize motion only by the laws of physics is a challenge .The result of performance cannot meet our expectation. Motion capture is a reliable way of acquiring realistic motions. It is less inaccurate to acquire in motion capture than to synthesize by physics-base methods.

Motion graph is a directed graph that consists of both original motion and automatically generated transitions. In motion graph each node means a choice point connecting different sets of motions. By searching in the graph, the desired motion sequence will be synthesized (Kovar et al. 2002). Arikan proposed a framework in a more intuitive way of synthesizing appealing motion, like the way a director guides actors and actresses (Arikan et al. 2003). Pre-process is needed to allow the user to specify the annotations that reflect the content of motion database. Users can describe motions by means of the annotations on a timeline to specify when each annotation will happen. Arikan also proposed to generate human motions by cutting and pasting motion capture data, he thought the generation of human motion was a solving search problem (Arikan et al. 2002). It is mainly due to the need to use random search of a hierarchy of graphs. In this approach, the property of a hierarchy of graph will easily generate the motion sequences that satisfy a variety of constraints.

Matthew Brand uses the probability mathematic model in pattern recognition. By learning from motion sequences, each model can identify the common motion style element (Brand 2000). It then decides to interpolate or extrapolate between two style models. This approach generates new choreography and synthesizes motion in many styles. J.K. Hodgkins proposed a method built from a common skill: state machines, inverse kinematics, etc. (Hodgkins et al 1998). The control algorithm is designed for some kind of motion behavior. A state machine can represent corresponding changes in action. The algorithms complying with inverse kinematics will generate motion for several dynamic behaviors. Kathy Pullen created a method that allows an artist to have low-level control of the motion. Motion capture data is used to enhance a key-framing animation (Pullen et al. 2002). Frequency Analysis divides the data (both key-framed data and motion capture data) into frequency bands. To match the key-framed data with motion capture data is to find the path and nest the segments accordingly. Zordan simulated unexpected impacts into a motion capture-driven animation system to physically simulate the impact that responds to contact forces, search for the best plausible re-entry into the motion library, and synthesize transition motion that connects previous and next motion capture clips (Zordan 2005). Mueller proposed a more efficient retrieval algorithm for large data sets. The main concept is to analyze geometric features and to use them to segment motion data for synthesis (Mueller et al. 2005). During the synthesis, it searched all possible paths, but it might result in “playing back” the motion graph within that environment. Separate specification of key-frame values from key-frame timing. Features of the user's input are mapped to features of the key-framed motion. The key-frames are then distributed in time according to the timing of the user's input path.

Wang et al. developed methods to determine a visually appealing length for a motion transition (i.e., a segue between two sequences of character animation). They have two methods for different type of motions to compute blend length based on two hypotheses on the nature of blending, including geodesic distance and joint velocities (Wang et al. 2004). So et al. gave a new segmentation solution for extracting motion patterns from motion capture data by searching for critical key poses in the motion sequence (So et al. 2005). They established a rank for critical key poses that identifies the significance of the directional change in motion data. Yu et al. proposed a framework allows the user to retrieve motions via Lab annotation (Yu et al. 2005). They generated a corresponding Labanotation sequence as additional motion property for each motion clip. A similarity metric for Labanotation sequences is used to search the motions that have similar Laban descriptions. Retrieve motion segments that only match part of the query Laban sequence. Based on dynamic programming, these segments are stitched together to form a smooth output motion that is in an optimal sense of matching query Laban sequence.

The above literature reviews emphasize the knowledge of how to select the tempo of music and how to segment motion efficiently and effectively, in order to make motion smoother. However, more research work needs to be conducted into specific types of music driven motion and its possible influence. In this research we aim to focus on the design implementation in the integration of music driven motion and propose a practical method. Hopefully, this could provide an additional contribution in the field of animation design.

3. System Overview

Our system consists of three distinct modules: Music analysis, Motion analysis and Motion synthesis. Figure 1 shows the relationship between these modules.

The music analyzer includes two components, which are for tempo detection and tempo edition. The tempo detection module processes a musical tune and finds the musical beats (low frequency sounds). The tempo edition module allows the user to modify the tempo of a segment of music manually to get a better musical sequence, which could interact with synthesized motions.

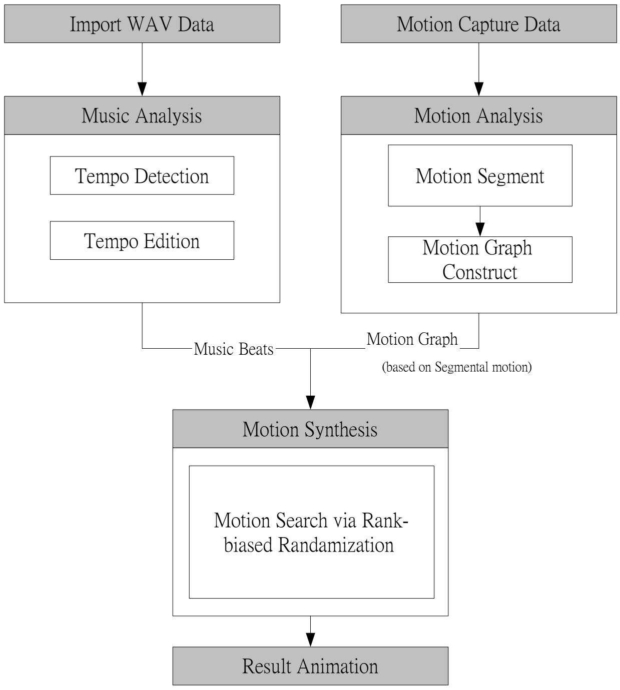


Figure 1 - System Overview.

The motion analyzer was made up of two components, the motion graph constructor and the motion path searcher. The motion graph constructor detects transition candidates and then selects transition points. After that, it creates the transition and prunes the motion graph. After constructing the motion graph, the motion paths are converted to different motion sequences. Afterwards, we can look for a motion sequence within the motion graph. After the procedure of music and motion analysis, the motion synthesizer adjusts the time of motion, based on music tempo, which is called motion time warping. Then, a motion path, which synchronized the music beats, is synthesized.

The details of each module are explained in the following sections.

4. Music Analysis

4.1 Beat Tracking

The purpose of music analysis is to extract the music beats. In most cases, the tempo is the regular pattern of beats in the music. In most dance tunes, the beats often appear in the low-pitch sound with low frequency, such as drums. Therefore, we locate the peaks of the low frequency volume curve, and regard them as the beats. Figure 2 shows the steps of tempo extraction.

Figure 2 - Tempo detection workflow.

The easiest way to extract low-frequency sounds is to apply a low-pass filter to the audio data. There are different low-pass filters. We use the Butterworth low-pass filter that is characterized by the magnitude response that is maximally flat in the pass-band and monotonic overall. It designs an order-n low-pass filter with a normalized cutoff frequency. It returns the filter coefficients in length n+1 row vectors b and a, with coefficients in descending powers of z.

$$H(z)=\frac{B(z)}{A(z)}=\frac{b(1)+b(2)z^{-1}+\cdots+b(n+1)z^{-n}}{1+a(2)z^{-1}+\cdots+a(n+1)z^{-n}}\quad\ldots\ldots(1)$$

Cutoff frequency is the frequency range where the magnitude response of the filter is described within. For Butterworth, the normalized cutoff frequency must be a number between 0 and 1, where 1 corresponds to the Nyquist frequency, in radians per sample. In our case, we choose the cutoff frequency and the filter order for different music respectively.

The measurement of sound volume, decibel, can be defined as:

$$Volume = 10 \bullet \log_{10}(Y)\quad\ldots\ldots(2)$$

The decibel (dB) is used to measure sound volume level, which represents the comparative intensity of sound. This measurement is closer to the human sense of sound magnitude.

We computed each frame’s volume (frame size = 1024, overlap = 128).

Before searching the local maxima of volumes, we quantized the volume and time vectors. When the volumes’ local maximum is greater than the threshold, it is regarded as where the beat occurs.

4.2 Tempo Changing

We also implement the method, WSOLA, in order to support the function of altering tempo in any part of music so that users can interact with the motions (Verhelst et al. 1993). This method was originally for tackling the problem of time-scale modification of speech.

The common problem while altering the music tempo is how to connect the waves naturally without gaps and keep their pitches. The method, WSOLA, provides a good solution to this problem. In figure 3, the blue curve is the original waveform and the black curve denotes a range of a frame. (1) is the first frame, and (1') is the second frame, and so on. For example, if we want to shorten the original waveform by half, we have to prune half of the frames away. So we take off the frame (1') right half after the first frame (1) and find the waveform which is most similar to the second one from a specific range of the third frame (2). The assumption is that there shouldn't be any abrupt change of waveform in a short-time period hence we certainly can find a similar enough waveform in that duration. （t） represents the gap of the waveform. This is shown in figure 4 below:

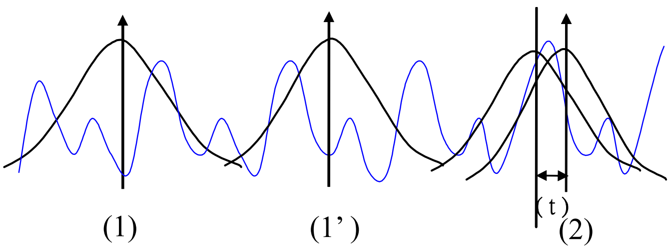


Figure 3 - Original Waveform.

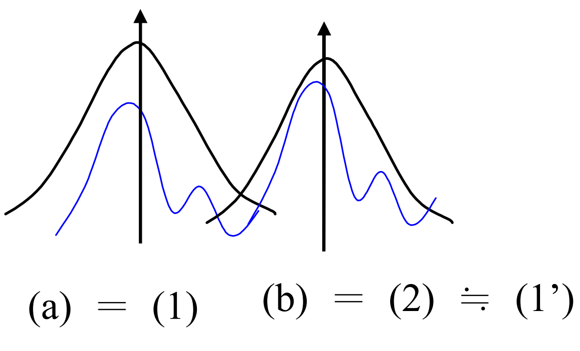


Figure 4 - Resultant waveform using WSOLA.

5. Motion Synthesis

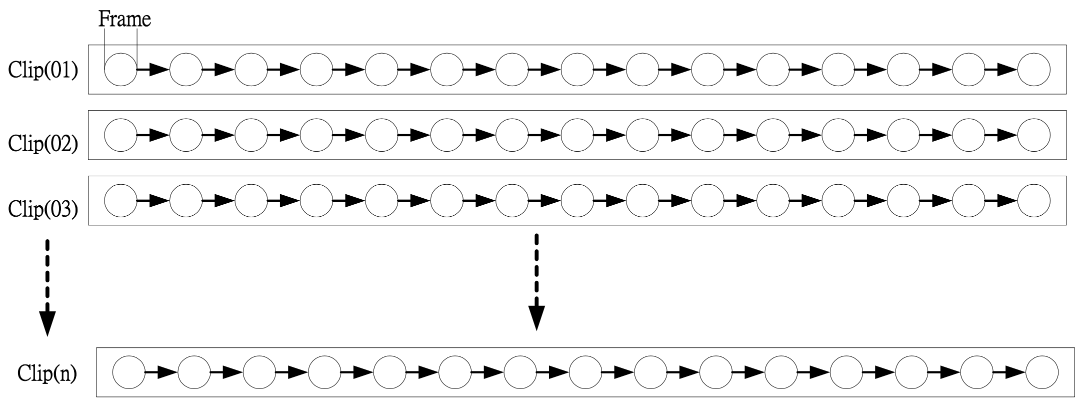
The important of how to synthesize a motion sequence with motion capture database (Kovar et al. 2002). Furthermore, we produce motions according to musical tempo (Alankus et al. 2005). We followed this approach and made some revisions by clipping motion segments in the database.We want to match motion segments to both the beginning and the end of the beats. However, it is unnatural to synthesize a motion sequence if we only match one motion to one segment.

5.1 Segment Motion Clip

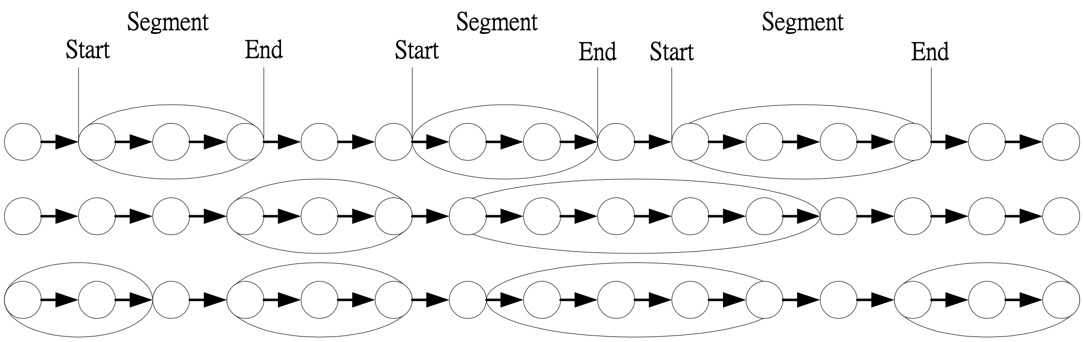
First, we cut each motion clip apart into many segments for later use in synthesizing the segment-based motion sequences. Each segment includes a complete motion unit. In addition, the motion segments were categorized into two groups, “big action” and “tiny action”. We only allow one “big action” in a correspondent beat set. It is necessary to rescale the length of segments but not to go beyond the limitation when the beat durations couldn’t match the segment lengths. If the length of a motion segment couldn’t fit the tempo duration even after rescaling, we applied “tiny action” in this case. This condition is shown in Fig.1. In our approach, a beat shouldn’t match with more than one “big action”.

5.2 Construct Motion Graph

There are many frames in each clip. We have located some frames, called nodes, in the clip, to cut the clip apart into many segments. We make connections between notes and synthesize motions based on them. See Fig 5.



(a). Motion capture database



(b). several segment

Figure 5 - (a) There are many clips in the motion capture database. (b) Each clip was divided into several segments.

In the next step, we recorded possible motions in the nodes of the present segment. The existence of a connection between two nodes means that we can transit from one motion to the other. To avoid unnatural connections, we chose two nodes with joint positions close to each other so that we could obtain good interpolation between them. These motions derived from interpolation are taken as “tiny actions”.

Next, we explain how to evaluate the similarity of two frames, F_a and F_b. F_b was transformed x_0 units along the X-axis, and z_0 units along the Z-axis and rotated an angle theta around the Y-axis.

$$\min_{\theta, x_0, z_0} \sum_i w_i \left\| P_i - T_{\theta, x_0, z_0} P'_i \right\|^2 \qquad (3)$$

Where the linear transformation T_{θ, x_0, z_0} rotates a point **p** about the y-axis by θ degrees and then translates it by (x_0, z_0).

This optimization has a closed-form solution:

$$\theta = \arctan \frac{\sum_i w_i (x_i z'_i - x'_i z_i) - \frac{1}{\sum_i w_i}(\overline{x}\,\overline{z'} + \overline{x'}\,\overline{z})}{\sum_i w_i (x_i x'_i + z_i z'_i) - \frac{1}{\sum_i w_i}(\overline{x}\,\overline{x'} + \overline{z}\,\overline{z'})}$$

$$x_0 = \frac{1}{\sum_i w_i}(x - \overline{x'}\cos\theta - \overline{z'}\sin\theta)$$

$$z_0 = \frac{1}{\sum_i w_i}(\overline{z} - \overline{x'} \sin\theta - \overline{z'} \cos\theta)$$

Where $\overline{x} = \sum_i w_i x_i$

Equation (3) defines the similarity of two nodes. The similarity between nodes was calculated by equation (3) and we can link two nodes depending on their similarity.

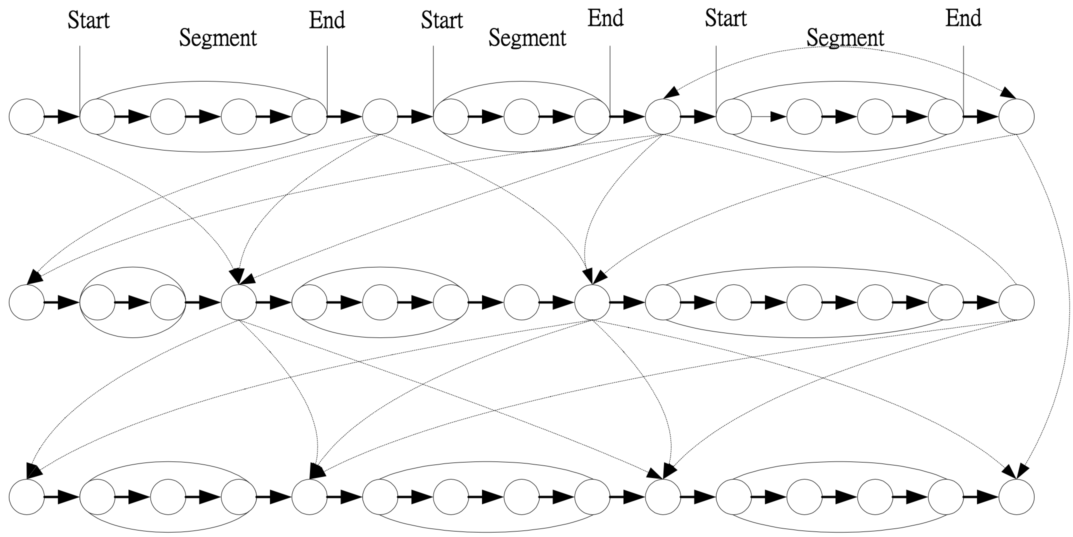


Figure 6 - Establishes transition between each clip segment to link the relations.

5.4 Motion Synthesis

We want to synthesize a sequence of motions which is the best match to the beats based on the motion graph we constructed before. Nevertheless, it makes little sense to find the best match from the overall paths, because of the enormous number of paths connecting all couple nodes among clips. For this reason, we solve this problem by the rank-biased randomization method. See Fig. 7

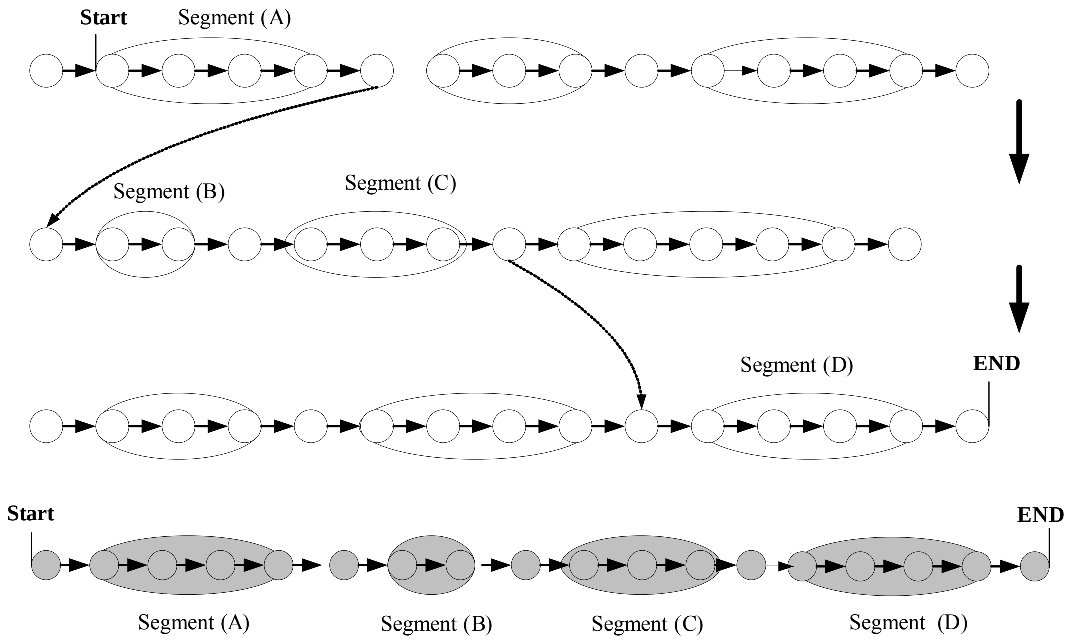


Figure 7 - (A) In the relation of transition in segment, we use the rank-biased randomization to synthesize the motion sequence. (B) Search the segments and match the beats in length. (C) After the search, there may be a big action with some small actions, or just a big action, or all small actions, but no big action between two beats.

In the initial stage of path searching, we arbitrarily pick up one node as the beginning. The following procedure was divided into two steps. First, we find another segment whose length is similar to the beat duration and then link it to the beginning. The equation for counting similarity is:

$$distortion = \min_{s_1 \dots s_n} \left[\left(\sigma \sum_{i=1}^{n} s_i \right) - L_{beats} \right]^2 \quad \text{...(5)}$$

S_i represents the length of the segment, and L_{beats} represents the length of two beats. σ is the ratio of rescaling. Segment searching should continuously link more segments to this beat set if the distortion is still over the threshold. We keep the one with smallest distortion and repeat this process to find a sequence of segment until the distortion is below the threshold. It is worth noting that the path should be deleted if more than one “big actions” were added into a beat set in our approach.

Second, we arrange all possible paths in an order based on their distortion and preserve first N paths. For all paths, the ending node of the segment added into a path last is able to link to the starting node of some segments in other paths. Therefore, we keep first n paths in order and search for their next possible segments whose start node fills the bill. Repeat step 1 and stop when all beat sets have their corresponding motion segments. See Fig. 8

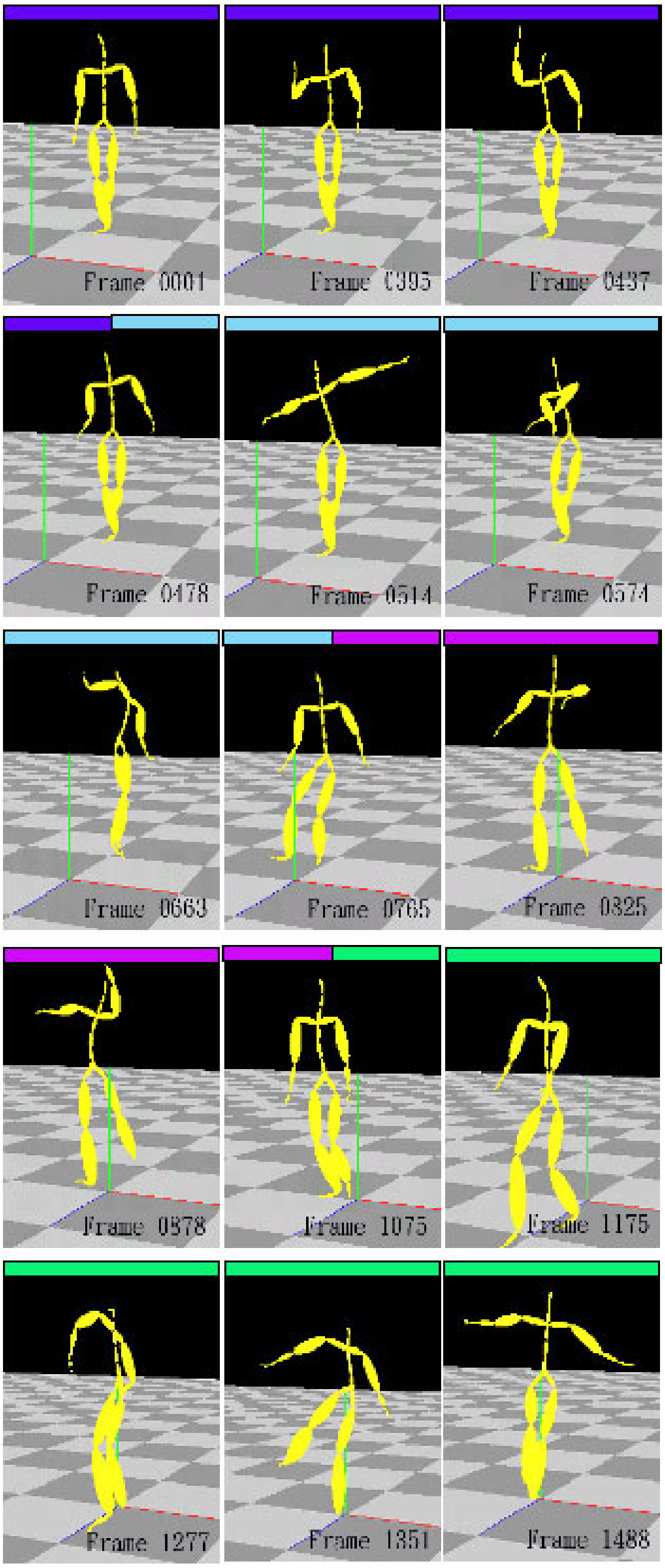


Figure 8 - The frame pictures with two-colors bar separate two neighbouring segments.

6. Experiment

6.1 Beat Detection Experiment

In the experiment on beat-detection, we chose some segments from three songs, which belong to three different types of music. The first song, A, is a pop song, the second one, B, is more like an aboriginal music used in a memorial ceremony and the last one, C, is a mixed dance music. The following table lists their recording-information.

Table 1 - Recordings information

	A	**B**	**C**
Recording type	44100Hz, 16-bit, stereo	22050Hz, 8-bit, mono	44100Hz, 16-bit, stereo
Recording format	Windows PCM	Windows PCM	Windows PCM
Recording length	31.266 sec	11.716 sec	5.561 sec

To measure the correctness, we manually marked where all the beat locations are, and called them reference-beats hereafter. Comparatively, the beats detected by our system were denoted as caught-beats. Two error-measuring criteria were designed as follows.

Redundant rate: The ratio of redundant beats number to caught-beats number. (Redundant beats are detected by our approach but could not be found in reference-beats group.)

Redundant rate = (No. of redundant beats) / (No. of caught beats)

Missing rate: the ratio of missed beats number to reference-beats. (Missing beats are not detected by our approach but in reference-beats group.)

Missing rate = (No. of missed beats) / (No. of reference beats)

Table 2 listed these two error ratios of three music pieces.

Table 2 - Beat-detection error rate

	A (POP Music)	**B (aboriginal music)**	**C (mixed dance music)**
Redundant rate	0.05	0.00	0.33
Missing rate	0.09	0.00	0.00

From the above results, we noticed that almost all beats were detected by this beat-detected approach but there were non-beats mistaken as beats. The error occurred in the step of finding local maximum because there are more than one local maximum in a period of waveform. Our solution to this problem is to shift beats above a given threshold energy.

6.2 Motion Synthesis Experiment

In the experiment of synthesizing the motion by means of the beats that are extracted from music, we used motion capture library supported by CMU (Carnegie Mellon University) Graphics Lab. We chose the dance category in the motion capture library as experimental data. There are twenty-nine motion clips in this experiment. We labeled each clip as several segments. In all motion clips, we generated one hundred and two labeled segments and about four thousand transitions between different segments. The average length of motion clips is approximately six hundred frames. A motion graph based on the segments was constructed by those segments and transitions.

We took three different kinds of music to synthesize motion. The mean and variance of the length between beats are described in the following table.

Table 3 - Motion clip attribution

	A (POP Music)	B (aboriginal music)	C (mixed dance music)
Mean (frame)	1684	267.158	769.646
Variance (frame)	97,973	1,683	15,000

We found that the performance in the synthesis was greatly affected by the segment which was generated from the interpolation of the transition of motion. Finding the transition between clips is a hard problem. We identified frames which were similar enough to make a transition between them. If we created a transition between a run and a jump, the transition motion would have been unrealistic.

So, among the segments generated from the transitions, we look for those clips with rescale ratio closed to one. In the process of synthesis, we rescale the length of the segment so it matches the length of beats. Therefore, we evaluated our system by the ratio of rescaling the length of segment. If the rescaling ratio is larger than 1, it means stretch a segment, if the ratio is smaller than 1, it means to contract or to shorten a segment.

We have used the three songs A, B, and C to synthesize thirty different motion sequences. The rescaling ratio in different music is described in the following table.

Table 4 - Ratio of rescaling evaluation

	A (POP Music)	B (aboriginal music)	C (mixed dance music)
Transition segment	1.74	0.73	1.10
Other segment	1.50	0.50	1.40
Average	1.62	0.62	1.25

Average rescaling rate = (Ratio of rescaling transition segments * No. of transition segments + Ratio of rescaling other segments * No. of other segments) / (No. of transition segments + No. of other segments)

Therefore, we can find the higher ratio of rescaling a transition, the lower degree of nature, because the segments that are generated from a transition were stretched out. In motion clip A, we found the fact that the transition segment was longer than other kinds of music. It had made unnatural transition segments more obvious. The low ratio of scale resulted from shortening the length of transition segment in motion clip B would introduce the rapid transition between frames. The result in music C is acceptable because the ratio of rescaling of transition segment is near to one. That is to say, the transition segment would not be stretched or contracted too much.

7. Conclusion

From the results of experiments, the synthesized motion matched the beats well. Between two beats, there may be a major action along with some minor actions, or only a major action. So, the beats will appear between the motions.

The length of the segment in the motion graph plays a key role in synthesizing a music-driven motion. If there were a lot of differences between the tempo duration and the length of the segment, the ratio of rescaling in the result would have increased, and the transition segment would have been affected. The motion synthesis with music in our experiment is dependent on the motion graph, which is constructed from some style of data. The transition between frames is important. If we could find some type of music can be analyzed in motion graph’s beat set, we could expect the result of synthesis by the mean of the segment length in motion graph. The more natural transition segments we could generate, the better the synthesis would be. So we use the ratio of rescaling segment as the measure of the performance. Besides, when we found the sequence of segments between the beat set, we could apply the same ratio of rescaling to the remaining segments. Nevertheless, for those sequence segments not matching well with beats set, we would save them and try to find a better solution in our future work.

Acknowledgements

We are grateful to CMU Graphic Lab and AXIS 3D Co. for providing us with character models and captured motion data. This research was supported by the National Center for High-Performance Computing, Taiwan.

References

[1] Alankus, G., Bayazit, A. A. & Bayazit, O. B. (2005) Automated motion synthesis for virtual choreography. *In Journal of Computer Animation and Virtual Worlds*.

[2] Alonso, M., David B. and Richard, G. (2004) Tempo and beat estimation of musical signals. *Proceedings of the 5th International Symposium on Music Information Retrieval (ISMIR)*, Barcelona, Spain, Universitat Pompeu Fabra

[3] Arikan, O., and Forsyth, D. 2002. Interactive motion generation from examples. *In Proceedings of SIGGRAPH 2002*, pp. 483–490.

[4] Arikan, O., Forsyth, D., O'Brien, J. (2003) Motion Synthesis from Annotations. *In Proceeding of ACM SIGGRAPH 2003*, pp402-408.

[5] Bevilacqua , Frédéric., Naugle, Lisa and Valverde , Isabel.(2001) Virtual dance and music environment using motion capture. *In Proceeding of the IEEE Multimedia Technology And Applications Conference* (MTAC), Irvine.

[6] Brand, M., & Hertzmann, A. (2000) Style machines*. In Proceedings of ACM SIGGRAPH 2000*, pp. 183-192, ACM Press.

[7] Fang, A., Pollard N. (2003) Efficient Synthesis of Physically Valid Human Motion. *In Proceedings of ACM SIGGRAPH 2003.*

[8] Goto, M. (2001) An audio-based real-time beat tracking system for music with or without drum-sounds. *Journal of New Music Research*, Vol. 30, No. 2, pp. 159–171.

[9] Hainsworrh, S. and Macleod M. (2003) *Beat tracking with particle filtering algorithms*. In 2003 IEEE Workshop on Applications of Siena1 Processing to Audio and Acoustics. October 19-22. New Paltz. NY pp91~94.

[10] Harper, R. and Jernigan, M.E. (2004) *Self-adjusting beat detection and prediction in music.* In Acoustics, Speech, and Signal Processing, 2004. Proceedings, pp245~248.

[11] Hodgins, J. K. and Wooten, W. L. (1998) Animating human athletes. *In Robotics Research: The Eighth International Symposium*. Y. Shirai and S. Hirose (eds). Springer, pp. 95-101.

[12] Kovar, L., Gleicher, M., and Pighin, F. (2002) Motion graphs. *In Proceeding of ACM SIGGRAPH 2002*, pp. 473-482 ACM Press.

[13] Laybourne, K. (1998) *The Animation Book: A Complete Guide to Animated Filmmaking-From Flip-Books to Sound Cartoons to 3-D Animation*. Three Rivers Press.

[14] Liu, C. K., Popovic, Z. (2002) Synthesis of Complex Dynamic Character Motion from Simple Animations. *In Proceedings of ACM SIGGRAPH 2002.*

[15] Mueller, M., Roeder T. and Clausen M. (2005) Efficient Content-Based Retrieval of Motion Capture Data. *In Proceeding of ACM SIGGRAPH 2005*, pp. 677-685.

[16] Popovic, J., Seitz, S., and Erdmann, M. (2003) Motion Sketching for Control of Rigid-Body Simulations. *ACM Transactions on Graphics*, Vol. 22, No. 4, pp. 1034–1054.

[17] Pullen, K. and Bregler C. (2002) Motions capture assisted animation: Texturing and synthesis. *Proceedings of SIGGRAPH*, pp 501 – 508.

[18] Qi Wang, Mie Nakatani and Shogo Nishida. *An approach to automatic motion synthesis harmonized with music for multiple objects*. In IEEE Transactions on Electronics, Information and System 2003

[19] Shiratori , T.,. Nakazawa, A. and Ikeuchi , K. (2004) *Detecting dance motion structure through music analysis.* In Automatic Face and Gesture Recognition Proceedings, 2004.

[20] So, C. K. F. and Baciu, G. (2005) Entropy-based Motion Extraction for Motion Capture Animation: Motion Capture and Retrieval. *Computer Animation and Virtual World*, Vol. 16, Issue 3-4.

[21] T. Yu, X. Shen, Q. Li and W. Geng. *Motion Retrieval Based on Movement Notation Language*, Computer Animation and Virtual World 2005, Vol. 16, Issue 3-4.

[22] Verhelst, W. and Roelands, M. (1993) *An overlap-add technique based on waveform similarity (WSOLA) for high quality time-scale modification of speech*. Acoustics, Speech, and Signal Processing 1993.

[23] Wang, J. and Bodenheimer, B. (2004) Computing the Duration of Motion Transitions: An Empirical Approach. *In Proceeding of ACM SIGGRAPH 2004*, pp. 335-344.

[24] Zordan, V., Chiu, B., Majkowska, A. and Fast, M. (2005) *Dynamic Response for Motion Capture Animation*. In Proceeding of ACM SIGGRAPH 2005, pp. 697–701.

Biographical note

Kuen-Meau Chen is a Assistant Researcher and leads the project group Interactive Multimedia Lab within the National Center for High-Performance Computing (NCHC), Taiwan, He is a Ph.D. candidate in the Department of Industrial and Commercial Design at the National Taiwan University of Science and Technology (NTUST). He received his M.S. degree in Industrial Design with honors from the National Yunlin University of Science and Technology (YunTech). His research interests include multimedia design, animation, shape grammar, as well as user interfaces, virtual reality and computer aided design.

E-mail: morris@nchc.org.tw

Dr Siu-Tsen Shen has studied widely, gaining her Masters degree in Industrial Design Research from the Design Academy of Eindhoven, and her PhD in Design from Goldsmiths College, University of London. She has been a guest lecturer at Middlesex University, UK and is currently an Assistant Professor in Multimedia Design at the National Formosa University, Taiwan. Her research interests lie in the areas of HCI, User-Centred Design, Cross-Cultural Research, User Interface Design, and Design Team Formation using Personality Type. She has taught on a number of different programmes including Introduction to Multimedia Design, Creative Thinking & Design Methods, Cognitive Psychology and Emerging Technology.

Email: stshen@nfu.edu.tw

Dr Stephen D. Prior is a Principal Lecturer in Product Design and Engineering at Middlesex University, London. He received a BEng Mechanical Engineering degree from Middlesex Polytechnic, London in 1987 and a PhD in Rehabilitation Robotics from Middlesex University, London in 1993. He is a Chartered Mechanical Engineer a Corporate Member of the IMechE and a Registered Practitioner with the Higher Education Academy. His research interests are in the areas of Mechatronics, Robotics, Unmanned Ground Vehicles and Team Selection & Formation in Design Education.

E-mail: s.prior@mdx.ac.uk