Nonlinear Dance Motion Analysis and Motion Editing using Hilbert-Huang Transform

Ran Dong
University of Tsukuba
Tsukuba, Ibaraki, Japan 305-8577
dong@cavelab.cs.tsukuba.ac.jp

Dongsheng Cai
University of Tsukuba
Tsukuba, Ibaraki, Japan 305-8577
cai@cs.tsukuba.ac.jp

Nobuyoshi Asai
University of Aizu
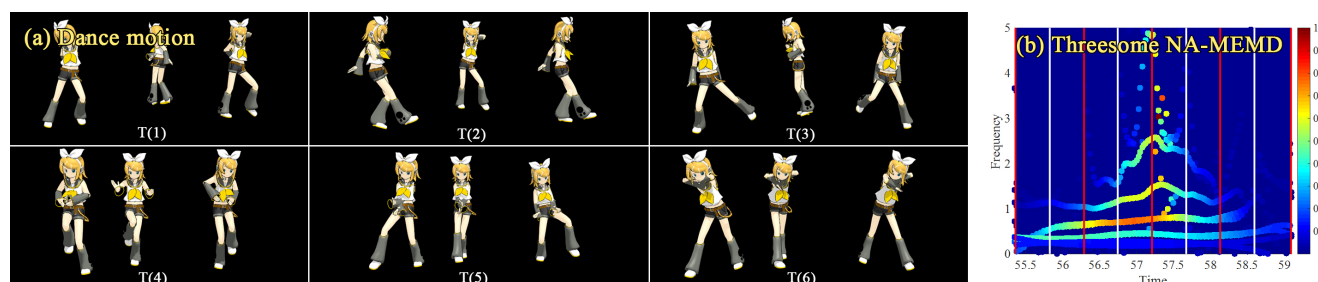Aizu, Fukushima, Japan 965-8580
nasai@u-aizu.ac.jp

Figure 1: (a) Six-time shots of a four-second Japanese threesome pop unit "Perfume" dance motions. (b) Hilbert spectrum (Energy (color)-Frequency (y-axis)-Time (x-axis)) of the dance motion data. The white and red lines are weak and strong beats, respectively.
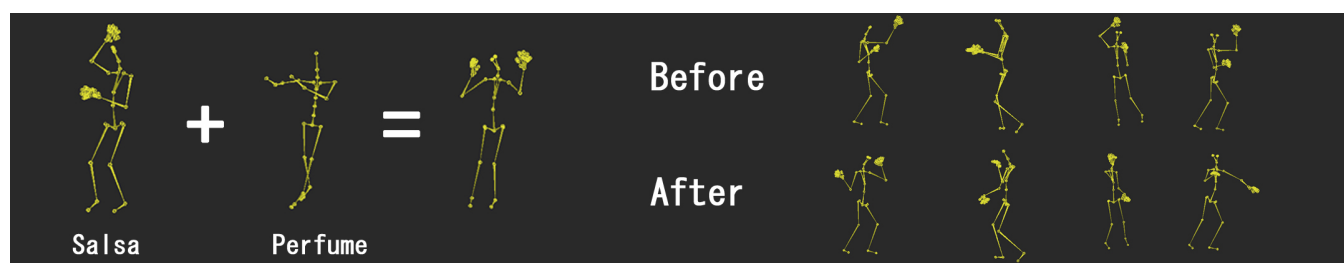

Figure 2: Salsa dance motion with Japanese threesome pop unit "Perfume" choreographies. Decomposing Perfume and Salsa dance into different distinct modes (IMFs) using NA-MEMD. Perfume upper body motion IMFs are blended with the Salsa motion IMFs. A new Perfume dance movements with Salsa steps are created.

ABSTRACT

Human motions (especially dance motions) are very noisy, and it is hard to analyze and edit the motions. To resolve this problem, we propose a new method to decompose and modify the motions using the Hilbert-Huang transform (HHT). First, HHT decomposes a chromatic signal into "monochromatic" signals that are the so-called Intrinsic Mode Functions (IMFs) using an Empirical Mode Decomposition (EMD) [6]. After applying the Hilbert Transform to each IMF, the instantaneous frequencies of the "monochromatic" signals can be obtained. The HHT has the advantage to analyze non-stationary and nonlinear signals such as human-joint-motions over FFT or Wavelet transform.

In the present paper, we propose a new framework to analyze and extract some new features from a famous Japanese threesome pop singer group called "Perfume", and compare it with Waltz and Salsa dance. Using the EMD, their dance motions can be decomposed into motion (choreographic) primitives or IMFs. Therefore we can scale, combine, subtract, exchange, and modify those IMFs, and can blend them into new dance motions self-consistently. Our analysis and framework can lead to a motion editing and blending method to create a new dance motion from different dance motions.

CCS CONCEPTS

•**Computing methodologies** → **Motion capture;** *Motion processing;* •**Mathematics of computing** → *Mathematical analysis;*

KEYWORDS

Dance Motion, Motion Analysis, Motion Editing and Blending, Motion Synthesis

1 INTRODUCTION

We analyze and decompose a four-second threesome techno pop dance by Perfume as shown in Fig. 1. The threesome dancers' motions are almost synchronous and occasionally asynchronous. Applying the Noise-Assisted Multivariate Empirical Mode Decomposition (NA-MEMD) to all three dancers' joint angle data, the hip movements of one dancer are now clearly decomposed into five different almost harmonic IMFs as shown in the left panel (b) of Fig. 1 (Hilbert spectrum plot). In the center of the plot, at t=57.25 sec., two modes peaked at 2.2, and 5 Hz indicate the left and right leg rotational motions respectively, and now this four-second noisy motion data is clearly decomposed into five modes or IMFs and one residual motion entitled "trend". Note that these clear left and right leg rotations cannot be detected using FFT and wavelet analyses. When we apply the same NA-MEMD to Waltz, and Salsa dances, these dance motions also can be apparently decomposed into 9-11 different IMFs and one trend.

Appropriately using the NA-MEMD, dance motions can be completely decomposed into different IMFs that correspond to different nonlinear "monochromatic" or "choreographic" motions, and one trend that is "non-choreographic" motions. Consequently, these different "choreographic" motions or corresponding IMFs can be cut, pasted, eliminated, subtracted, added, scaled, and they also can be edited and blended with the different dance motions or corresponding IMFs smoothly. In Fig. 2, we replace the Salsa upper body IMFs with those of Perfume. Thus, Perfume dance can easily be converted into that with Salsa steps.

In this paper, first, we introduce previous works and review HHT briefly. Second, using HHT, we propose a new method and framework to analyze and edit dance motions. Third, using our proposed framework, we analyze a few dances (mainly Perfume dance) and adapt and blend those dances to create a new dance with different tastes. Finally, we conclude our paper.

2 RELATED WORKS

Various motion blending techniques based on motion capture data are proposed for character animation. For examples, Heck suggests a method using parametric motion graphs to generate and interpolate a new motion between some basic motion data [4]. Lee proposed a method editing existing human-like character motion data to obtain desired human actions [7]. However, previous researches focus only on interpolating motions, and they cannot blend different movements into new motions.

Unuma used Fourier coefficients to extract and edit motion capture data [10]. However, it is only possible to do so when animators have already known the characteristics of the targeted motions. It is inappropriate to use this method to edit dance motions composed of various unknown features.

To resolve this problem, we propose a framework using HHT to decompose dance motion into nonlinear signals (monochromatic waves) to extract their features. We also offer a method to edit and blend different dances to generate new dance motions using HHT.

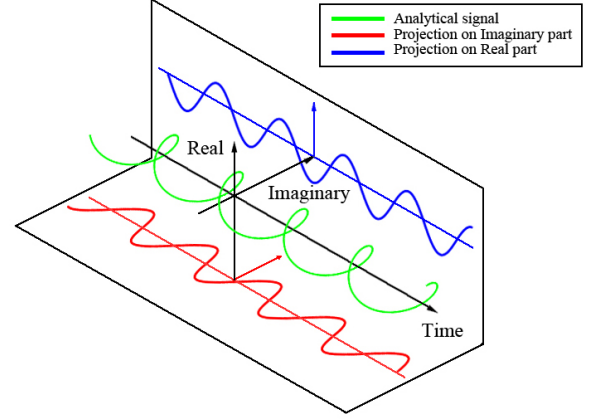## 3 HILBERT-HUANG TRANSFORM

3.1 Hilbert Transform

Figure 3: Analytical signals with a real and imaginary part. They are quadrature functions (cited from [1]).

An analytical signal $z(t) = z_r(t) + iz_i(t)$ varying in time is plotted in Fig. 3. The instantaneous amplitude and frequency of Fig. 3 can be expressed as follows:

$$A(t) = \sqrt{z_r^2(t) + z_i^2(t)} \qquad \omega_0(t) = \frac{d}{dt}tan^{-1}\frac{z_i(t)}{z_r(t)} \qquad (1)$$

It is known that the Hilbert Transform (HT) converts the real part of the analytical function into its imaginary part. However, the Hilbert Transform assumes the signal is a monochromatic wave, and the real part of the analytical function can be expressed as $x(t) = A(t)cos(\omega_0(t)t)$ as shown in Fig. 3. Thus, Hilbert Transform can be defined as follows:

$$z_i(t) = y(t) = \frac{1}{\pi}PV\int_{-\infty}^{\infty}\frac{x(\tau)}{t-\tau}\,d\tau = \frac{1}{\pi t}*x(t) \qquad (2)$$

In Eq. (2), the integration means the Cauchy principal value (PV) integration [1].

Since the Hilbert Transform (HT) assumes the signal is a monochromatic wave $x(t)$, it is not suitable to apply HT to the chromatic signals that are true for many real-world signals. Thus, Huang proposed a method entitled the Empirical Mode Decomposition (EMD) to decompose the chromatic signals into a set of monochromatic signals called the Intrinsic Mode Functions (IMFs).

The essential part of the HHT is the "Empirical Mode Decomposition (EMD)". A single-variate or multivariate chromatic signal can be decomposed into a finite set of "Intrinsic Mode Functions (IMF)" and a residual called "trend" [6]. After applying motion captured joint angle data into the Noise-Assisted Multivariate Empirical Mode Decomposition (NA-MEMD) [11], we can decompose dancers' "choreographic" motion primitives into separate IMFs and a trend that corresponds to dancers' postures including transitions in time.

3.2 Empirical Mode Decomposition (EMD)

In this section, we briefly introduce the EMD algorithm and the IMF definition (We do not explain the details of IMF and its residuals r in the present paper. For the details, please see [6]). The Hilbert Transform can be applied to the signals after decomposing the original signal into an IMF set and a "trend" [6].

3.3 EMD Algorithm

Intrinsic Mode Function (IMF) is defined as follow [6]:

- The number of extreme value is equal to that of zero-crossings, or the difference is 1.
- The average of maximum and minimum envelopes is 0 at any time.

Here is the EMD algorithm [6]:

(1) Calculate residual $r(t)$ (Let $r(t) = x(t)$ in the first time) as follows:

$$r(t) = x(t) - \sum_n C_n(t) \qquad (3)$$

(2) Initialize $c(t) = r(t)$ and extract the Intrinsic Mode Function (IMF)
  (a) Find maximum envelope $u(t)$ and minimum envelope $l(t)$ of $c(t)$
  (b) Subtract the average envelope from $c(t)$

$$c_{new}(t) = c_{old}(t) - \frac{u(t) + l(t)}{2} \qquad (4)$$

  (c) If the convergence condition SD (0.2-0.3) is satisfied, add $c(t)$ into the IMF set.

$$SD = \sum_n \frac{(c_{old}(t) - c_{new}(t))^2}{c_{old}^2(t)} \qquad (5)$$

(3) Repeat step 1 and 2 to extracts all IMFs.

3.4 Multivariate EMD (MEMD)

For multi-channel or multivariate analytical signals, it is impossible to use the EMD to decompose "chromatic" signals. The Multivariate EMD (MEMD) is proposed for multi-channel or multivariate signals such as dance motion data that are composed of all joint angles of a human body [9].

Due to MEMD Filter bank function that can remove white noises of the data, in NA-MEMD, Gaussian White Noises (GWN) is proposed to be packed into one extra data channel. Thus, the mode mixing that causes the HHT inaccurate can be reduced or eliminated [11].

As shown in Fig. 4, HHT can accurately capture five different choreographic motion primitives or IMFs of the left leg and right leg, while the Short-Time-Fourier-transform (STFT) only can capture vague high-speed motions around the center. It is desirable to observe dance movements in a frequency domain, and HHT certainly has some advantages over STFT, because the uncertainty principles limit the STFT resolutions.
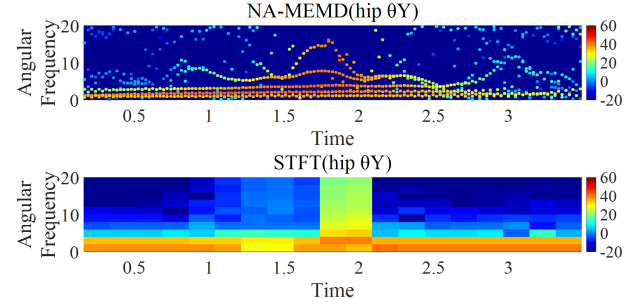

Figure 4: A four-sceond motion comparisons between HHT and Short-Time-Fourier-Transform (STFT). The motion is the same as that of Fig. 1.

4 PROPOSED FRAMEWORK OF ANALYZING AND DECOMPOSING DANCE MOTIONS

4.1 Dance Motions Analysis Framework

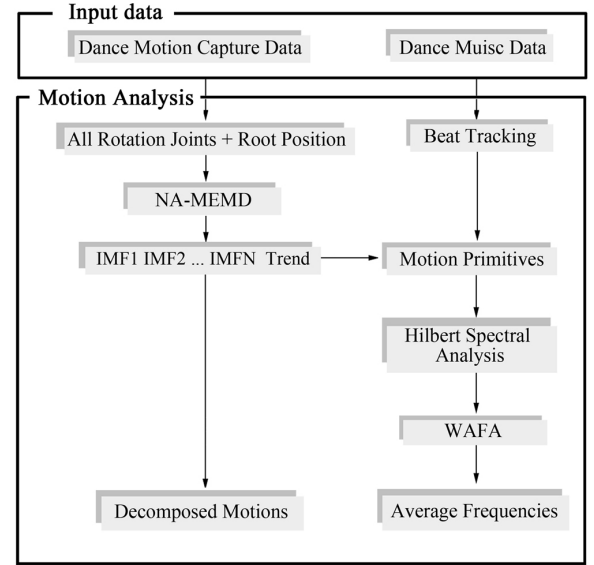Fig. 5 shows our proposed framework using both motion capture data and music data.

Figure 5: Proposed framework to decompose "choreographic" motion primitives and analyze dance motions using HHT.

Beat Tracking We used a beat tracking system suggested by Dan Ellis to segment the dance motions [2]. Our system assumes a regular tempo based on a fixed beat per minute (BPM). Then, we calculate the relevancy between beats and music to extract the beat itself. Please see the details in [2].

Motion Primitives To analyze and edit dance motions, it is necessary to segment dance motions into a set of short meaningful motions ("choreography"). A motion primitive is the most basic motion unit in human motions, and it can

be expressed as a time series of key poses. To analyze dance motions and extract motion primitives, first, we must track beats in dance music in our system because each primitive is synchronizing with the beats. To edit the dance motion primitives, the motions have to be segmented based on the beats.

NA-MEMD The motion-captured human motions are composed of many body joints with the Eulerian angles $\theta x, \theta y, \theta z$. To handle these data using the EMD, we generate the n-dimensional envelopes by projecting different joint signals along different directions in the n-dimensional space using the n-dimensional sphere. After we obtain the n-dimensional envelope, IMFs are obtained using the same algorithm, as introduced in subsections 3.3 and 3.4 [9]. Usually, motion captured kinematic data are extremely noisy. The NA-MEMD can significantly reduce the noises of the motion-captured data to obtain their IMFs as shown in Fig. 1(b) [11].

Weighted Average Frequency Algorithm Since we obtain the instantaneous frequency and amplitude in Eq. (1) using HHT, averaging them is necessary. It is because that the instantaneous frequency and amplitude in one flame are meaningless for one dance movement due to the nature of the instantaneous frequency and amplitude. So we have to observe and define the frequency in one smallest time unit i.e. one motion primitive time unit. The Weighted Average Frequency Algorithm (WAFA) is a method to average IMFs to obtain meaningful data. Besides, the real signals like motion captured data are always very noisy [8]. Applying WAFA to one motion primitive time length, the meaningful and denoised IMFs can be obtained.

4.2 Limitations of the proposed framework

EMD is a decomposition technique without a mathematical proof. Therefore, there is no guarantee that dance motions can be decomposed into proper IMFs. Our proposed framework may have the following limitations:

An Over-Decomposed IMF In an EMD, a chromatic signal is decomposed into a set of "monochromatic" signals from high to low-frequency signals. However, dance motions are not exactly choreographed by monochromatic motions. Thus, one motion primitives can be over-decomposed into a few similar motion IMFs. To avoid over-decompositions, we have to combine over-decomposed IMFs with a few adjacent IMFs manually.

A Singular IMF The frequencies of decomposed IMFs decrease gradually with the decomposition order. However, a singular IMF is an IMF with an outlier frequency value [5]. Usually, the singular IMFs can be observed in a low-frequency dance motion, when taking their WAFA frequencies. Sometimes, the singular IMF is not associated with its original dance primitives physically. To avoid this singularity, it is necessary to use a proper noise in NA-MEMD. In general, it is ideal that adding GWN whose amplitude is about 8%-10% of the original signal into NA-MEMD. It can avoid the mode mixing and the singular IMF [11].

5 RESULTS

5.1 Dance Motion Analysis

The threesome techno pop dance unit Perfume has been performed for more than ten years and considered to be matured and experienced dance group. Perfumes' dance is considered to be extremely difficult. However, at a glance, it seems catchy, simple and easy to dance. According to Perfume choreographer Mikiko, she expressed that the image of Perfume music is inorganic and closer to near future than the modern style. As a result, their dances have been arranged to express a puppet to obtain unworldly movements.

The motion captured Perfume dance is compared with other motion-captured dances such as Waltz, and Salsa using the proposed HHT method. Here, Table 1 shows the lengths and frame rates of four different dance motion captured data.

Table 1: Motion Captured Data for Analysis

Dance Style	Perfume	Waltzy	Salsa
Time(s)	60.5	52.6	28.5
Frame rate	40	30	40

We apply the NA-MEMD to three hip Euler angles $\theta x, \theta y, \theta z$ for full dance motions to compare their features. Figures 6-8 show HHT spectrum of the one hip Eulerian angle for three different dances respectively. Here, the angle of hip θx for Perfume, Waltz and θz for Salsa is plotted because they are the most significant angles in the dances. The beats are not displayed to avoid being overwhelmed. Each IMF average frequencies that are averaged for all dance time are listed in the left column of the figures.
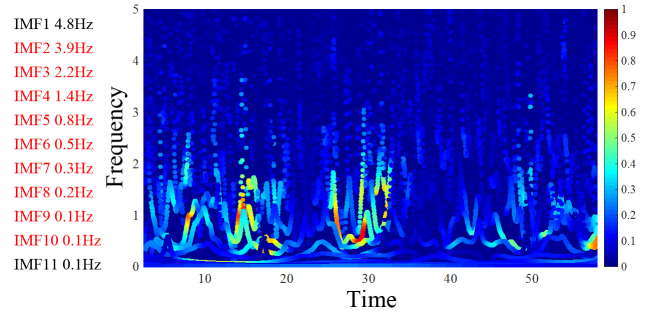

Figure 6: NA-MEMD Hilbert Spectrum of Perfume (hip θx).

In Figs. 6, 7, and 8, the Perfume, Waltz, and Salsa dance motions are respectively decomposed into 11, 11, and 9 IMFs ranging from 0.1 to 4.8Hz, from 0.1 to 3.8Hz, and from 0.4 to 4.8Hz. The Salsa dance Hilbert spectrum shows IMF frequencies are noisy and concentrated around 2-4Hz. These results indicate that the Salsa motion is very up-tempo and the dance style is rather free, unorganized, and noisy.

From our beat tracking analysis, Perfume, Waltz, and Salsa dance motion BPMs (Beat Per Minutes) are, respectively, 130, 84-93, and 150 (we estimate it as a slow type Salsa) [3]. The BPM of Perfume is about 130 and is as fast as that of Salsa. In addition, the highest WAFA frequency of IMF (IMF 1) of both Perfume and Salsa is the
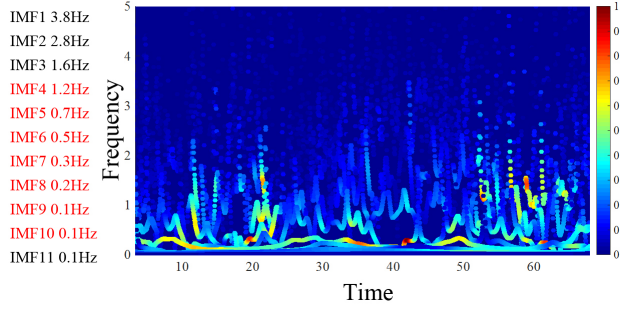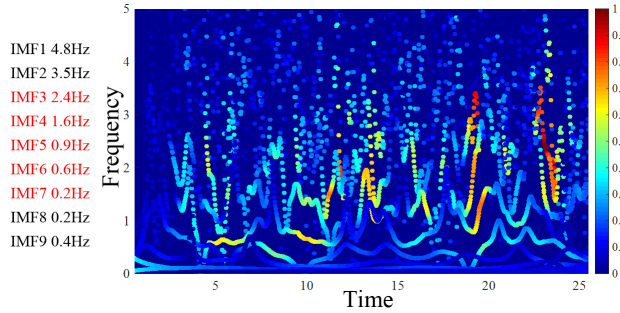
Figure 7: NA-MEMD Hilbert Spectrum of Waltz (hip θx).

Figure 8: NA-MEMD Hilbert Spectrum of Salsa (hip θz).

same and is 4.8 Hz, and much greater than the Waltz. The most distinct Hilbert power spectrum of both Perfume and Salsa is located around 1-2 Hz, and higher than the Waltz of which the distinct Hilbert power spectrum are located around 0-1 Hz.

On the other hand, the IMF number of Perfume is 11 and is as many as that of Waltz which is a rather slow tempo and classical dance. The Waltz is considered to have many distinct choreographic primitives, and each IMF should correspond to different choreography or motion primitives (The way how the different dance motions correspond to the different IMFs are not shown here due to the limit of the paper length).

One of the most interesting features of the HHT analysis applied to the dance motions can be analyzed from the WAFA frequencies of IMFs listed as the red numbers in the left columns in Figs. 6-8. The WAFA frequencies are averaged over for whole dance time. From the bottom to the top of those red numbers, we can quickly to find the relation as follows: frequency of IMF (n) = frequency of IMF (n+1) + frequency of IMF (n+2). From the bottom to top, they form the Fibonacci sequences. In Fig. 6 of Perfume dance, from the bottom to top, 0.1+0.1=0.2, 0.1+0.2=0.3, 0.2+0.3=0.5Hz…, etc., and except the top two and bottom IMFs, they all form Fibonacci sequences. These Fibonacci relations indicate the dancers are moving their body strictly in a regular tempo and rhythms. Comparing the Perfume (Fig. 6) and Salsa (Fig. 8), the Perfume dancer follows more strictly Fibonacci relation than the Salsa dance.

Following the Hilbert-Huang spectrum analysis, we can summarize the Perfume dance as follows:

(1) the Perfume dance IMF spectrum have as high frequency as those of up-tempo Salsa.

(2) the Perfume dance has as many IMFs as those of the traditional Waltz

(3) Perfume dance average IMF frequencies follows more Fibonacci relations than the up-tempo Salsa.

We can assume the number of dance choreography is proportional to the IMF number. We know the Salsa is up-tempo and freestyle dance. The Perfume dance is as up-tempo as the Salsa but has as much kind of IMF or motion primitives as that of Waltz. Also, Perfume dance motion is strictly in regular tempo and rhythm. These may explain the reason why people comment "Perfume's dance is visually simple but difficult to dance", and the choreographer Mikiko commented her dance as "puppet-like".
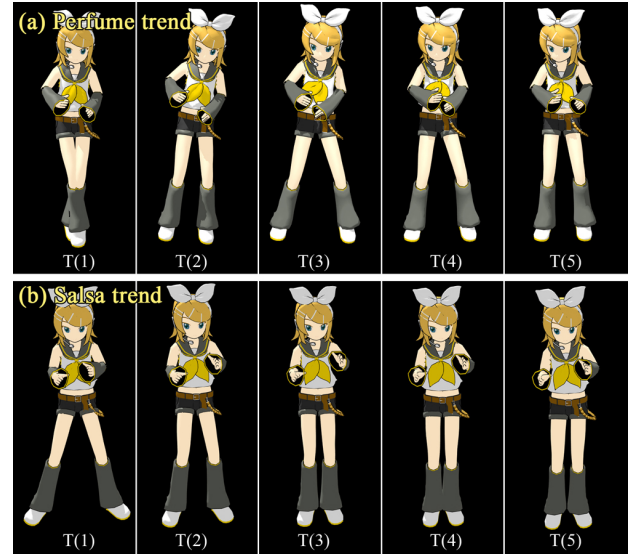
5.2 Dance Motion Editing

Figure 9: (a) Perfume trend, and (b) Salsa trend extracted from 55-second motion data by NA-MEMD.

The EMD can decompose chromatic signals into a finite set of monochromatic signals entitled IMFs and its residual entitled Trend. Applying the NA-MEMD to dance motion captured data, the different IMFs can be associated with the different motion primitives or choreographic motions, and the trend can be associated with the posture of the dance. We show dance motions by using free software MikuMikudance. Fig. 9 shows Perfume and Salsa trends, these trends have almost no distinct body movements and only show the gradual time transitions of their postures. In the figure, a significant difference between Perfume and Salsa trends can be found, especially, on the upper and lower bodies.

Because the EMD can decompose the motions into the distinct different motion primitives as "monochromatic" signals and trend, we can cut, paste, add, subtract, and scale the IMFs and trend appropriately, to produce new dance motions consistently. Also, we can add, blend, exchange IMFs from different dances and create a new dance with different tastes.
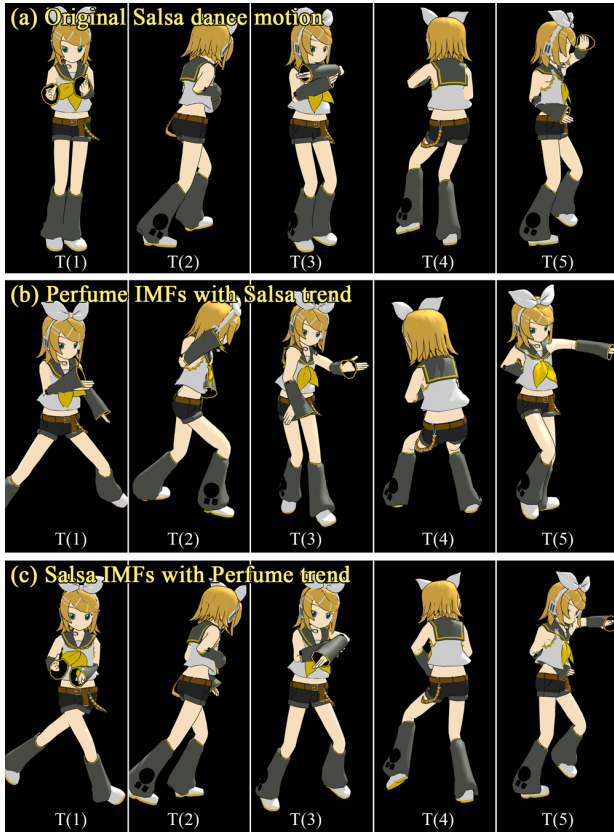
Figure 10: (a) Original Salsa dance motion and (b) Perfume dance motion with Salsa trend (c) Salsa dance motion with Perfume trend.

First, we decompose threesome pop unit Perfumes' dance into IMFs, select one of three dancers' IMFs, and generate a new dance using Salsa trend as shown in Fig. 10 (b). For instance, as shown in Fig. 10 (b), the dancer has the entirely different posture comparing with the original one as illustrated in T (3) of Fig. 10 (a). As a result, a new dance motion has been created. The new dance has both Perfumes' distinct primitive motions (puppet-like choreography) and active and flexible posture transitions of the Salsa as indicated in Fig. 10 (b).

Second, a new dance is created combining one Perfumes' trend and Salsa IMFs as shown in Fig. 10 (c). Significant differences can be observed between the original and newly created dance motions. For instance, it is clear that distinct leg motions can be observed in T (1) due to the Perfume trend in Fig. 9.

6   CONCLUSIONS

The human motions like dance motions are very noisy and extremely difficult to analyze. The Hilbert-Huang transform using the NA-MEMD can clearly decompose the noisy dance motions into distinct "monochromatic" IMFs and may have a substantial advantage over the other methods such as the short-time-Fourier-transform (STFT) and Wavelet analysis, etc. We propose a dance motion analysis system based on HHT, and we also discover some

new dance motion characteristics using our method such as Fibonacci relations in the Hilbert spectrum, especially in the Perfume dance motions.

The Fibonacci relation in the averaged frequencies can be interpreted as follows: Because the human body has a dynamic link structure when the dancer moves each part of the body at a particular frequency, we expect that the relationship of a Fibonacci sequence can be observed in the Hilbert spectrum. We only discuss this relation briefly, and further studies are required in the future. Apparently, such characteristics cannot be discovered using other methods as we discussed in Section 3. However, further researches are required for the detailed dance motion analyses using the HHT.

Because the EMD can decompose dance motions into "monochromatic" motions that roughly correspond to "motion primitives" with the different frequencies and its residual, i. e., Trend. For examples, we show the Perfume dance motions with the Salsa steps in Fig. 2 and the Perfume dance with the Salsa trend, and the Salsa dance with the Perfume trend as shown in Fig. 10. Thus, new dance motions with different tastes can be created. The editing and blending methods shown here are still limited and very primitive. Some feet-sliding, body collisions, and body penetrations can be observed using our method. These have to be carefully removed. Further researches are required for the detailed dance motion editing and blending using the HHT in the future.

ACKNOWLEDGMENTS

REFERENCES
[1] Ronald Newbold Bracewell and Ronald N Bracewell. 1986. *The Fourier transform and its applications*. Vol. 31999. McGraw-Hill New York.
[2] Daniel PW Ellis. 2007. Beat tracking by dynamic programming. *Journal of New Music Research* 36, 1 (2007), 51–60.
[3] Rich Gross. 2012. Tempo Indications and Beats Per Minute (BPM) Reference for Social Dance Genres. http://www.hollywoodballroomdc.com/recommended-tempos-for-dance-music/. (January 2012).
[4] Rachel Heck and Michael Gleicher. 2007. Parametric motion graphs. In *Proceedings of the 2007 symposium on Interactive 3D graphics and games*. ACM, 129–136.
[5] Jianzhao Huang, Jian Xie, Feng Li, and Liang Li. 2013. A THRESHOLD DENOISING METHOD BASED ON EMD. *Journal of Theoretical and Applied Information Technology* 47, 1 (2013).
[6] Norden Eh Huang. 2014. *Hilbert-Huang transform and its applications*. Vol. 16. World Scientific.
[7] Jehee Lee and Sung Yong Shin. 1999. A hierarchical approach to interactive motion editing for human-like figures. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*. ACM Press/Addison-Wesley Publishing Co., 39–48.
[8] J Niu, Y Liu, W Jiang, X Li, and G Kuang. 2012. Weighted average frequency algorithm for Hilbert–Huang spectrum and its application to micro-Doppler estimation. *IET Radar, Sonar & Navigation* 6, 7 (2012), 595–602.
[9] Naveed Rehman and Danilo P Mandic. 2009. Multivariate empirical mode decomposition. In *Proceedings of The Royal Society of London A: Mathematical, Physical and Engineering Sciences*. The Royal Society, rspa20090502.
[10] Munetoshi Unuma, Ken Anjyo, and Ryozo Takeuchi. 1995. Fourier principles for emotion-based human figure animation. In *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*. ACM, 91–96.
[11] Naveed ur Rehman, Cheolsoo Park, Norden E Huang, and Danilo P Mandic. 2013. EMD via MEMD: multivariate noise-aided computation of standard EMD. *Advances in Adaptive Data Analysis* 5, 02 (2013), 1350007.