Stylistic Walk Synthesis Based on Fourier Decomposition

Joelle Tilmanne and Thierry Dutoit

Numediart Institute, University of Mons, Mons, Belgium
joelle.tilmanne@umons.ac.be
www.tcts.fpms.ac.be\~tilmanne

Abstract. We present a stylistic walk modeling and synthesis method based on frequency analysis of motion capture data. We observe that two peaks corresponding to the walk cycle fundamental frequency and its first harmonic can easily be found for most walk styles in the Fourier transform. Hence a second order Fourier series efficiently represents most styles, as assessed in the subjective user evaluation procedure, even though it results in a strong filtering of the original signals and hence a strong smoothing of the resulting motion sequences.

Keywords: motion capture, synthesis, Fourier transform.

1 Introduction

A broad field of applications can be found for human motion analysis and synthesis: entertainment (games, animation, etc.), medical applications, sports, artistic performances, etc. When considering humanlike motion synthesis, most methods aim at modeling and synthesizing motion sequences in the temporal domain, using a wide range of methods [1] such as keyframe animation [2], procedural models, motion graphs[3], Principal Component Analysis [4], Hidden Markov Models [5,6,7], etc. However, another way of understanding and analyzing motion data is to study it in the frequency domain. This temporal to frequency domain transformation can be obtained through Fourier analysis.

Troje [8] models motion by a continuous component, the walk fundamental frequency and its first harmonic. He models the Cartesian coordinates of 15 body markers and reports that keeping only these three contributions retains 99% of the variance of the input data. Bruderlin et al. [9] use multiresolution filtering for decomposing motion into several frequency bands whose amplitudes can be modified in order to change the motion style. Unuma et al. [10] apply Fourier transform on motion and modify its aspect by changing the weights of the Fourier coefficients.

In the present work, we applied a Fourier decomposition to walk motion sequences, either normal walks or walks presenting exaggerated styles. The Fourier decomposition was applied to the angular representation of the motion (i.e. angles applied to the joints of a skeleton with constant limb lengths) rather than to the 3D coordinates of the skeleton joints, and to walks presenting various styles, expressed with very diverse influences on the resulting walk motion.

M. Mancas et al. (Eds.): INTETAIN 2013, LNICST 124, pp. 71–79, 2013.

2 Method

2.1 Stylistic Walk Frequency Content Analysis

We used motion capture data from our previously recorded databases (Tilmanne et al. [11]) which contain walk sequences at several speeds (*eNTERFACE'08* database) and with different styles (*Mockey* database) for 18-articulations skeletons. The tridimensional angle data is represented through the exponential map parameterization, and hence by three values for each articulation or joint angle. Figure 1 displays the amplitudes of the Fourier coefficients for a discrete Fourier transform of the 54 exponential map values representing the 18 body joint angles during one normal walk sequence of the *eNTERFACE'08* database. The Fourier transform is calculated on each one of the 54 exponential map values, and the 54 resulting Fourier transforms are superimposed in the figure.

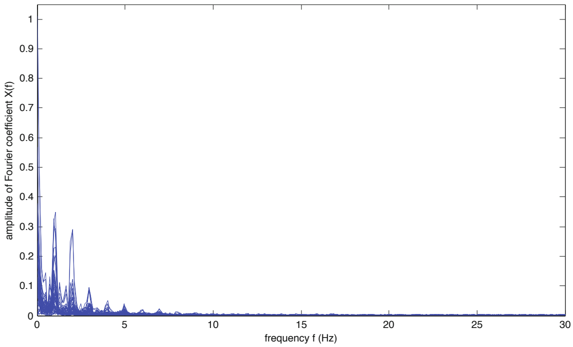

Fig. 1. Discrete Fourier transform of one normal walk sequence, for the 54 exponential map values of the 18 body joint angles

This discrete Fourier transformation clearly outlines the periodical nature of gait. Figure 1 displays a strong constant component for the frequency $f = 0$. This amplitude corresponds to the mean pose of the skeleton during the walk sequence. Outside from this first component, a noticeably higher peak appears at the fundamental frequency of the walk sequence, around 1Hz, which corresponds to the frequency of one complete walk cycle (= two steps) during the analyzed walk. The next highest peak is found at the first harmonic of the fundamental frequency, around 2 Hz, and corresponds to the frequency of one step.

The same profile of Fourier coefficient amplitudes was observed for all the speeds (slow, normal and fast) and 41 walkers of the *eNTERFACE'08* database. When analyzing more complex or elaborated walk styles, like the exaggerated walk styles present in the Mockey database, the profile of Fourier coefficients is sometimes more complex. Figure 2 illustrates the Fourier analysis of one continuous walk sequence for each one of the eleven Mockey styles. The red line corresponds to the mean walk cycle frequency of the analyzed sequence, calculated thanks to our automatic segmentation algorithm.

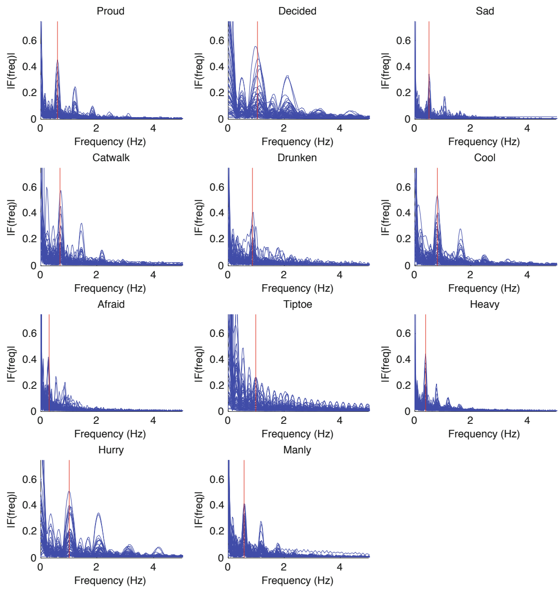

Fig. 2. Discrete Fourier transform of the eleven stylistic walk sequences from the Mockey database, for the 54 exponential map values of the 18 body joint angles. The red line represents the mean frequency of the walk sequence.

We observe that the two peaks corresponding to the walk cycle fundamental frequency and its first harmonic can easily be found for most styles. In general, however, the Fourier coefficient amplitudes display a more complicated profile over the eleven Mockey styles than for the normal walk of Figure 1. For some styles, like the *tiptoe* walk, the Fourier coefficients even display a completely different profile. Furthermore, the length of the available walk sequence can be an issue for an optimal evaluation of the Fourier coefficients and the Fourier coefficients cannot be calculated with the same precision for all the walk styles.

2.2 Fourier Based Stylistic Motion Modeling and Synthesis

Based on the Fourier analysis illustrated in Figure 1, it is obvious that the main components of a basic human walk can be modeled by taking into account only a few components of the Fourier transform. This approach has been introduced by Nikolaus Troje [12,8], who "linearizes" the periodic walk motion data by modeling it using only the first two components of the corresponding Fourier series. Each variable P_i of the motion data is represented by one continuous component and two cosines corresponding to the walk fundamental frequency and its first harmonic, as expressed in Equation 1:

$$P_i(t) = p_{i,0} + p_{i,1} \cos(\omega_0 t + \phi_{i,1}) + p_{i,2} \cos(2\omega_0 t + \phi_{i,2}) + err_i \tag{1}$$

Each motion variable is hence represented by five time-independent parameters $(p_{i,0}, p_{i,1}, \phi_{i,1}, p_{i,2}, and\ \phi_{i,2})$, plus a sixth parameter independent of the considered motion variable (the fundamental frequency ω_0). The err_i term represents the difference between the original variable P_i and its modeling and is hence set to zero for the synthesis stage.

In his work, Troje applies the Fourier series decomposition to the Cartesian coordinates of the body joints. However, we used the angle data motion representation which we found more suitable for our synthesis purposes. In addition to avoiding synthesized motions implying varying limb lengths, this joint angle representation gets rid of one more non-kinematic factor influencing motion perception: skeleton proportions. Troje's work mainly focuses on gender recognition based on point light display walk sequences and shows that the male versus female walk recognition is based both on the static pose (ratio between shoulders and hip width for instance) and kinematics parameters. Since our final goal is to be able to control interactively the walk of a given virtual character, we do not want our style parameterization to be influenced by skeleton size related information. Basing our Fourier analysis on the angle data parameterization enables the synthesis model to rely on kinematic features only. As illustrated in Figure 1, the Fourier analysis of exponential map parameterization also displays peaks around the walk sequence fundamental frequency and first harmonic, as the Fourier analysis of cartesian coordinates used by Troje.

Let us consider a data matrix X of size 54 x N corresponding to one walk sequence of N frames parameterized by 54 exponential map values x_i $(i = 1, ..., 54)$ at each frame k. The first step of our Fourier series decomposition consists in determining the $p_{i,0}$ coefficient of Equation 1. This coefficient is very easily calculated since it corresponds to the mean of each one of the 54 x_i:

$$p_{i,0} = \frac{1}{N} \sum_{k=1}^{N} x_{i,k} \qquad\qquad i = 1, ..., 54 \tag{2}$$

The next step consists in determining the fundamental frequency ω_0 of the walk sequence to be modeled. The peaks can be observed in the Fourier Transform coefficients, and extracted very easily for the eNTERFACE'08 database walks. Nonetheless, as illustrated in Figure 2, the highest peak does not always correspond to the walk fundamental frequency, especially when more complex walk styles are considered. However, if we want the Fourier analysis to linearize our data and enable us to compare our different styles, a common representation must be kept. We therefore decided to apply Troje's modeling procedure to all of our walk styles, even when it implies the loss of important frequency contributions, since these additional contributions corresponded to different frequencies for each style (in opposition to the fundamental frequency and its first harmonic which are present in every walk style).

In order to determine which peak of the Fourier coefficient amplitudes corresponds to the fundamental walk frequency, even when it was not the highest one, we used the step durations extracted thanks to an automatic step segmentation. The fundamental cycle mean period T_{mean} is obtained by multiplying this mean step duration by two, and the inverse of the cycle mean period T_{mean} gives a mean cycle frequency $f_{mean} = \frac{1}{T_{mean}}$. As observed in Figure 2 (red lines), this f_{mean} frequency corresponds to the walk cycle fundamental frequency peak of the Fourier transform. We hence set $f_0 = f_{mean}$, and the fundamental angular frequency ω_0 is thus known: $\omega_0 = 2\pi f_0$.

Once the fundamental angular frequency ω_0 is known, the corresponding coefficients can easily be calculated. The complex Fourier coefficients $c_{i,1}$ and $c_{i,2}$ are then calculated by a projection of the data on the exponential axis corresponding to the fundamental frequency and to its first harmonic:

$$c_{i,1} = \frac{1}{N} \sum_{k=1}^{N} x_{i,k} * \exp^{-j\omega_0 k} \tag{3}$$

$$c_{i,2} = \frac{1}{N} \sum_{k=1}^{N} x_{i,k} * \exp^{-j2\omega_0 k} \tag{4}$$

The corresponding coefficients $p_{i,1}$, $\phi_{i,1}$, $p_{i,2}$ and $\phi_{i,2}$ are calculated thanks to $c_{i,1}$ and $c_{i,2}$:

$$p_{i,l} = 2 * abs(c_{i,l}) = 2 * \sqrt{real(c_{i,l})^2 + im(c_{i,l})^2}\ , \qquad i = 1, ..., 54,\ l = 1, 2 \tag{5}$$

$$\phi_{i,l} = phase(c_{i,l}) = \arctan \frac{im(c_{i,l})}{real(c_{i,l})}\ , \qquad i = 1, ..., 54,\ l = 1, 2 \tag{6}$$

Each of our 54 motion variables $P_i(t)$ being represented by five parameters plus one common parameter (fundamental frequency), one complete walk sequence of length N is represented by 271 values.

Using these 271 values and Equation 1, new walk sequences of any chosen length can be synthesized. This second order Fourier decomposition is illustrated in Figure 3 for 3 of the 54 original angle data values of the motion captured walk sequence. The original data used to calculate the Fourier coefficients is presented in the first subplot and its approximation through the second order Fourier series defined by Equation 1 is illustrated in the last subplot. The individual contributions of the continuous, fundamental frequency and first harmonic components are presented in subplots two to four.

However, even if new walk sequences of any chosen length can be synthesized, for a given style each step will remain exactly identical during the whole walk sequence. Furthermore, this decomposition process is equivalent to a very strong filtering since only the fundamental frequency and its first harmonic are kept. This approach hence results in a heavy smoothing of the synthesized data. However, new styles can be synthesized by taking new values in the 271-dimensional Fourier coefficient space and using them to synthesize walk sequence according to Equation 1 or by interpolation between the existing styles. If the new styles remain close to the space area determined by our original walks, they will lead to believable walk sequences.

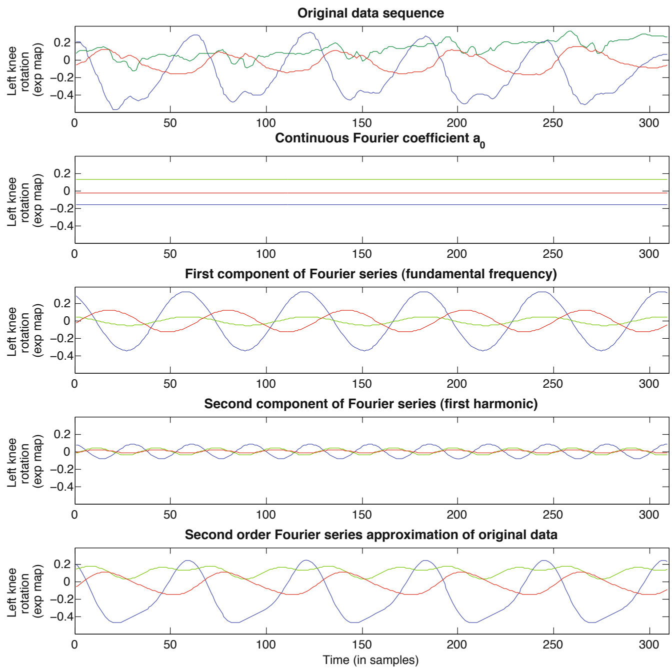

Fig. 3. Decomposition of a motion captured walk sequence into a second order Fourier series, illustrated for the three values of the rotation of the left knee, expressed in the exponential map parameterization (*expmap* 1, 2 and 3 illustrated in blue, green and red respectively). The original motion capture sequence is illustrated in the first subplot. The three next subplots represent respectively the continuous component of the Fourier decomposition, the fundamental frequency component and the first harmonic. The last subplot presents the approximation of the original data by the second order Fourier series.

3 Qualitative User Evaluation

In order to evaluate the quality of the Fourier synthesis process, a qualitative user evaluation was conducted. Thirty-seven subjects took part in the unsupervised web-based evaluation: 16 males and 21 females, with an age of mean 35.6 and standard deviation 12.6. The evaluation was divided into three series of 15 tests, with a maximum of 45 evaluations per participant.

In each test iteration, the participant was presented two videos at the same time. He could play each video as many times as he wanted. He was asked to position the synthesized stylistic walk on a scale from 0 to 4 compared to the original stylistic walk (0 being "not realistic" and 4 being "as natural as the original walk"). In the video, the motions to be assessed were applied on a simple blue stick-figure character.

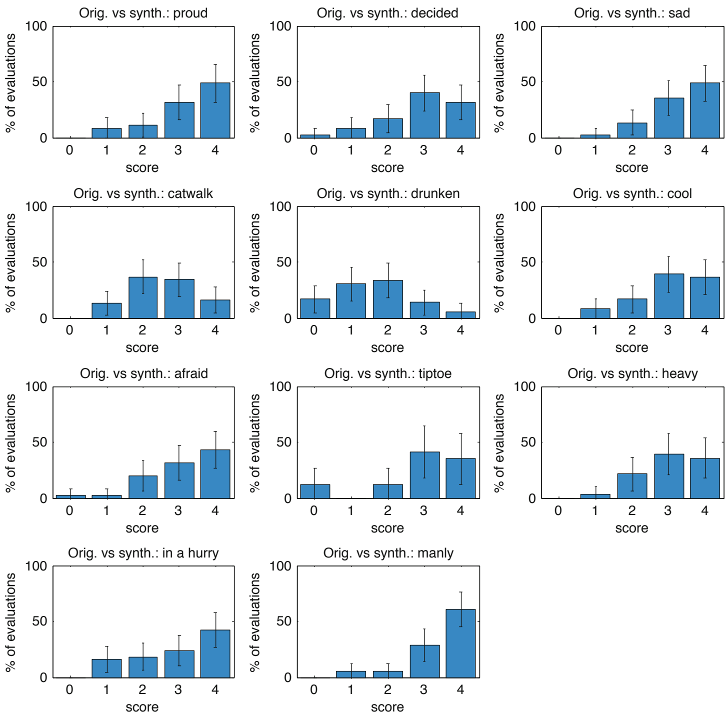

Fig. 4. Scores given by the evaluators to the synthesized walk sequences (Fourier-based approach) compared to original walk sequences, with 95% confidence intervals, for each on of the eleven styles of the Mockey database. A score of four corresponds to a synthesized walk as natural as the original walk, and a score of zero to a non realistic synthesized walk.

Two different method were assessed during this evaluation, and among the 67 tests of the complete evaluation set, 21 video pairs aimed at assessing the perceived naturalness of the Fourier series based synthesis. 390 evaluations were performed on these 21 walk pairs. The second method assessed was based on a principal component analysis (PCA) and is beyond the scope of the present paper.

Figure 4 displays the style by style results of the evaluation and Figure 5 the mean score per style. With a mean score of 2.91 but with wide variations between the extreme scores, as style specific mean scores range from 1.60 (drunken walk) to 3.45 (manly walk), the naturalness perception varies widely.

The overall score is quite good since evaluators were sensitive to the exaggerated nature of the original walk styles. Since the two cosines based recomposition smoothed part of the stylistic content of the walk, some evaluators orally reported that they found the synthesized sequences more believable than the original walk sequences. The results displayed in Figure 4 illustrate the fact that this method has some interest for periodical walk styles displaying a simple Fourier transform profile, but that it does not work in walks displaying important variations

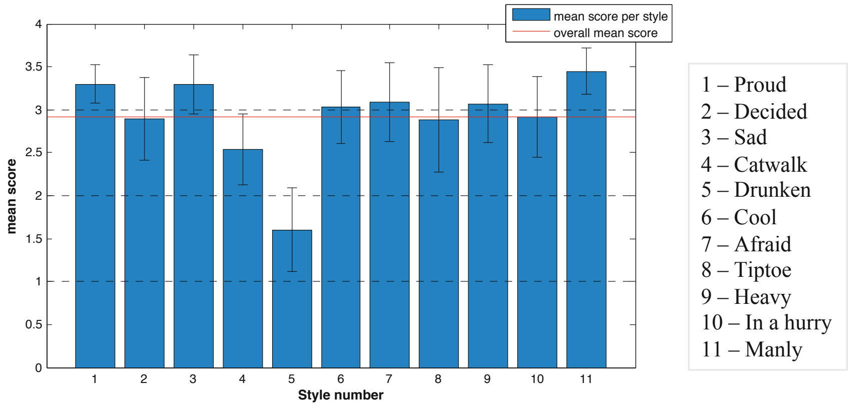

Fig. 5. Mean scores given by the evaluators to the synthesized walk sequences (Fourier-based approach), with 95% confidence interval, for each on of the eleven styles of the Mockey database. The overall mean score is displayed as a red line.

or more complex structures than periodical motions of the limbs aligned to the walk cycle frequency and its first harmonic.

4 Conclusion

Fourier analysis enables us to study the profile of different walk styles into the frequency domain. We apply Troje's approach to stylistic walks sequences, but adapt it and use it on exponential map angle data rather than Cartesian coordinates, so as to have a style representation independent of skeleton size and hence of the walker's morphology.

As can be observed in the Fourier analysis illustrated in Figures 1 and 2 and in the qualitative evaluation results, the walk modeling with a second order Fourier series applies better to normal walks than to stylistic walks which display more complex Fourier coefficients patterns and tend to display more than the two basic peaks observed in normal walks. This is especially the case for walk styles such as the *drunken* walk, where part of the style lies in the fact that two successive steps are very different from each other, which cannot be modeled with a periodic time series. However, we were able to apply the same procedure to all the walk styles and to analyze the resulting synthesized sequences. As can be foreseen, the synthesized sequences appear smoothed, since keeping only the contributions of two frequencies in addition to the constant component is equivalent to a strong filtering. With this approach, no cycle to cycle variation is possible since the walk sequence is described as a strictly periodic pattern, and even the small periodic variations that compose the motion are filtered out. It has to be noted that such a modeling technique is based on the periodical nature of gait, and hence cannot be adapted to non-periodical motions.

However, this Fourier series decomposition remains interesting to analyze the motion and its basic components. It is worth noticing that for most styles, a recognizable walk, even if not realistic, can be synthesized with as few as two cosines and one constant for each variable.

References

1. Forsyth, D.A., Arikan, O., Ikemoto, L., O'Brien, J., Ramanan, D.: Computational studies of human motion: Part 1, tracking and motion synthesis. Foundations and Trends in Computer Graphics and Vision 1(2-3), 77–254 (2005)
2. Williams, R.: Animator's Survival Kit: A Manual of Methods, Principles, and Formulas for Classical, Computer, Games, Stop Motion, and Internet Animators. Faber & Faber (2002)
3. Kovar, L., Gleicher, M., Pighin, F.: Motion graphs. In: ACM SIGGRAPH 2008 Classes, p. 51. ACM (2008)
4. Glardon, P., Boulic, R., Thalmann, D.: PCA-based walking engine using motion capture data. In: Computer Graphics International, pp. 292–298. IEEE (2004)
5. Brand, M., Hertzmann, A.: Style machines. In: Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, pp. 183–192. ACM Press/Addison-Wesley Publishing Co. (2000)
6. Li, Y., Wang, T., Shum, H.Y.: Motion texture: A two-level statistical model for character motion synthesis. In: Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques, pp. 465–472. ACM (2002)
7. Tilmanne, J., Dutoit, T.: Continuous control of style and style transitions through linear interpolation in hidden markov model based walk synthesis. In: Gavrilova, M.L., Tan, C.J.K. (eds.) Transactions on Computational Science XVI. LNCS, vol. 7380, pp. 34–54. Springer, Heidelberg (2012)
8. Troje, N.F.: Retrieving information from human movement patterns. In: Shipley, T.F., Zacks, J.M. (eds.) Understanding Events: How Humans See, Represent, and Act on Events, vol. 1, pp. 308–334. Oxford University Press (2008)
9. Bruderlin, A., Williams, L.: Motion signal processing. In: SIGGRAPH 1995 Proceedings, pp. 97–104 (1995)
10. Unuma, M., Anjyo, K., Takeuchi, R.: Fourier principles for emotion-based human figure animation. In: SIGGRAPH 1995 Proceedings, pp. 91–96 (1995)
11. Tilmanne, J.: Joelle tilmanne's webpage, `http://tcts.fpms.ac.be/~tilmanne/`
12. Troje, N.: The little difference: Fourier based gender classification from biological motion. Dynamic Perception, 115–120 (2002)