INTRINSIC MODE FUNCTIONS FOR GAIT RECOGNITION

Prem Kuchi and Sethuraman Panchanathan

Research Center for Ubiquitous Computing
Arizona State University, Tempe AZ, 85287
{kuchi, panch}@asu.edu

ABSTRACT

Gait recognition is an attractive biometric as it is unobtrusive and can be used for recognition from a distance. A number of methods have been proposed by different researchers in the recent past for this purpose. Most of these methods analyze gait as a linear and stationary signal. However, recent research shows that gait is nonlinear and non-stationary. Hence, linear analysis would be insufficient for analysis of gait. In this paper, we present a novel recognition algorithm that derives the feature vector by performing nonlinear, non-stationary analysis of gait using a technique called empirical mode decomposition. We test the algorithm with both noise-free and noisy gait sequences and demonstrate its applicability.

1. INTRODUCTION

Gait recognition is becoming increasingly important for surveillance, aware-spaces etc. Unobtrusiveness and recognition at a distance are two most significant reasons why gait recognition is attractive. However, it has been difficult to develop effective techniques that could achieve this purpose. Recently, several research efforts in the field of biomechanics [1] has led to the conclusion that gait is both nonlinear as well as non-stationary. Consequently, we explored the use of a nonlinear, non-stationary signal analysis technique called the empirical mode decomposition for gait recognition. The results we obtained were highly encouraging. In the sections below, we discuss the theory of empirical mode decomposition, our experimental methodology and results of our method.

2. BACKGROUND

The recognition capability of human beings was first reported by several psychophysical studies in the late sixties [2, 3]. However, it was only in late nineties that machine recognition of gait became a possibility. Several methods were proposed, of which, some representative methods are described here. These methods can be divided into two basic categories: Structural methods and structure-free methods.

In structural methods the trajectories of specific points on the body or the angles between them are used to derive features for recognition. Bobick et.al. [4] proposed a gait recognition technique that recovers static body and stride parameters (different distances between various parts of the body). Tests on about 20 subjects showed that the technique held some promise. However, it needs to be seen how these static parameters scale up for larger databases. In [5], Lakany et.al. use the trajectories of specific locations, whose Fourier transform was used a the feature vector (FV). A Multi-Layer Perceptron (MLP) is used for classification and a Radial Basis Function (RBF) network was used to predict the data that is lost due to self-occlusion. In [6], Cunado et. al, Fourier magnitude spectra of joint angles weighted with phase information was used to generate the FV. A correct classification rate (CCR) of 90% was obtained using k-nearest neighbor rule on a database of ten subjects.

These structure-free methods characterize the motion pattern of the body without regard to its underlying structure. In [7], Little and Boyd used moments of the moving points weighted by the dense optical flow for characterizing motion. Phase of these scalar sequences was used as feature vector for recognizing gait. In [8], Huang et. al propose using the principal components of temporal information from optical-flow changes between two consecutive spatial templates for extracting features for gait recognition. Hayfron-Acquah et.al [9] use the symmetry pattern of human motion for recognition. The generalized symmetry operator was used for gait analysis. Using this approach they have achieved a CCR of 95%. However, there are studies that show that gait might not be symmetric [10].

One of the major problems of the above methods is that almost all of them use linear analysis to derive the feature vector. They do not consider the nonlinear, nonstationary nature of gait in the design of their algorithm. Also, all of these methods have been tested only on small databases. It is not clear how these methods would scale up for larger databases. We might lose information, probably vital, because of linear analysis of these nonlinear, non-stationary signals. Neural networks could perform nonlinear signal processing. However, they are used as *black boxes* and are typically used only for classification. A few other techniques use statistical methods like eigen analysis, which are insufficient for dealing with non-stationary signals. Hence, we propose a gait recognition methodology that uses nonlinear, non-stationary signal analysis for creating the FV.

3. INTRINSIC MODE FUNCTIONS

Instantaneous frequency (IF) is one of the ways in which a local characteristics of a non-stationary signal can be described. However, not all analytical signals can be described effectively using IF. If blindly applied to any signal, IF may result in two cases: (1) IF not being one of the frequencies in the Fourier spectrum, (2) IF going outside the band for a bandlimited signal [11]. Based on these observation, a class of functions called the intrinsic mode functions were proposed. A function is an intrinsic mode function if: "(1) The number of extrema and the number of zero crossings are either equal or differ at most by one. (2) At any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero"[12]. The above conditions satisfy the physically necessary conditions to define a meaningful instantaneous frequency. The method to decompose any signal into a set of IMFs, called *sifting* is described below.

3.1. Sifting

The signals found in nature do not fit the definition of intrinsic mode functions. A method for decomposing any general signal (which satisfies some basic properties) into a set of IMFs has been proposed by Huang et. al [12]. This method is described here. For the signal to be successfully decomposed into IMFs using this method it has to satisfy one assumption: It has to have at least two extrema – one maximum and one minimum. The method essentially involves two steps. Let $X(t)$ be the given signal.

1. Two smooth splines are constructed connecting all the maxima and the minima of $X(t)$ respectively to get its upper envelope $X_{max}(t)$ and $X_{min}(t)$. The extrema are found by determining the change of sign of the derivative of the signal. All the data points should be covered by the upper and lower envelopes.
2. The mean of the two envelopes is subtracted from the data to get a difference signal $X_1(t)$,

$$X_1(t) = X(t) - \frac{X_{max}(t) + X_{min}(t)}{2} \tag{1}$$

   Now the process is repeated for $X_1(t)$ till the resulting $C_1(t)$ satisfies the criteria of an intrinsic mode function.

3.1.1. Stopping Criterion

Typically, after a certain number of iterations, it results in a pure frequency modulated signal of constant amplitude. So we stop sifting by limiting the amount of standard deviation (SD), computed from two consecutive sifting results. Usually, SD is set between 0.2 to 0.3. This results in the first IMF $C_1(t)$. After this, we subtract $C_1(t)$ from the original signal to get the residue $R_1(t)$ and repeat the above two steps until we get a residue $R_n(t)$ that is smaller than a predetermined value or becomes a monotonic function. The original signal can be reconstructed using the following equation.

$$X(t) = \sum_{j=1}^{n} C_j(t) + R_n(t) \tag{2}$$

The method decomposes the data into n intrinsic mode functions, each with distinct time scale. The first component has the finest scale and this increases as we go to the successive modes. The decomposition is based on the local time scale of the data and yields adaptive basis functions. Hence it can be used for nonlinear and nonstationary signal analysis.

As we can see from the above section, the basis functions are derived from the signal itself. This is the major advantage of this procedure. Therefore the analysis is adaptive, in contrast to the traditional methods where the basis functions are fixed. Though in eigen analysis, the basis functions are derived from the signal, the method cannot be used for nonstationary signals. Since a majority of biological signals are nonlinear and nonstationary, this method could be effective (for e.g. electrogastrogram [13]).

4. EXPERIMENTS

We conducted gait recognition experiments on a database compiled by the Motor Control and Rehabilitation Research Laboratory at Arizona State University [14]. The database consists of kinematic signals of five participants with five sets of trials for each participant. The kinematic signals consisted of the (x, y, z) time-series of a set of fifteen markers that were placed on the participant (Marker locations shown in figure 1). The data was collected at a rate of 120 Hz.

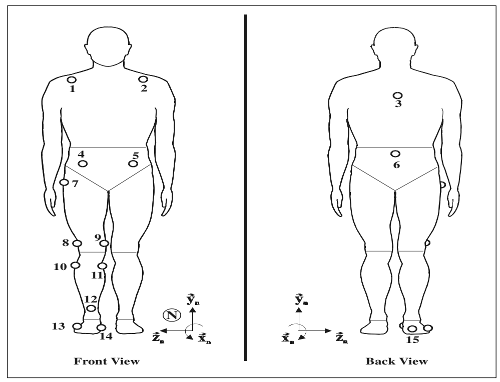

Fig. 1. Marker locations used for data capture

One gait cycle (series of events that occur between a right heel strike to a right heel strike or its left counterpart) of raw data (without any filtering) was used for our recognition experiments. The trajectory of only one marker was used for recognition (y-axis displacement of the 14$^{\text{th}}$ marker). The trajectory of one gait cycle is shown in the top most subplot in the figure 2. It shows the y-axis (vertical) displacement of the 14$^{\text{th}}$ marker.

4.1. Decomposition into IMFs

Each of these kinematic signals were decomposed into their respective IMFs using the *sifting* procedure. Since different people walk with different speeds, and also, each person walks with slightly different speeds for each gait cycle, we need to normalize the gait cycles to a specific length. We achieve this by using dynamic time warping [15]. We then perform empirical mode decomposition as described in the section 3.1.

5. RESULTS AND DISCUSSION

The intrinsic mode functions resulting from the empirical mode decomposition of the gait trajectories are shown in the figure 2 (the second subplot to the last one are the IMFs). Careful observation of these IMFs for different participants revealed two important facts.

1. The last IMF is similar in all participants and is large in magnitude. In fact, this component has the largest magnitude. Figure 3 illustrates this with the mean of all the last IMFs and amount of deviation that might occur from the mean. The deviation is not significant between different subjects and different trials of each subject.

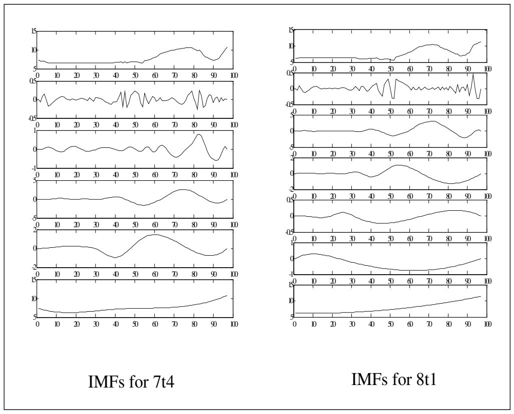

Fig. 2. IMFs of participants 7 and 8

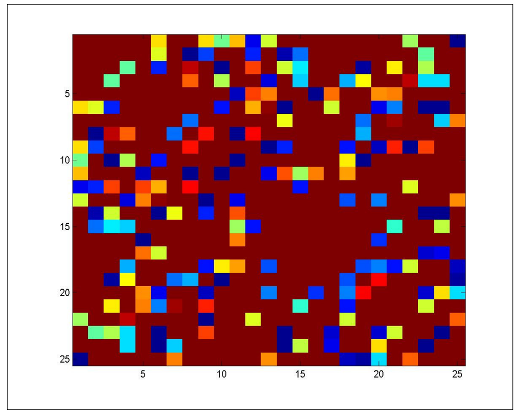

Fig. 4. Similarity between different First IMFs

2. The first IMF is different for different participants and different for different trials of the same participant. Figure 4 shows the correlation between different first IMFs. As we can see, the correlation between different trials of the same participant and between different participants is nearly same.

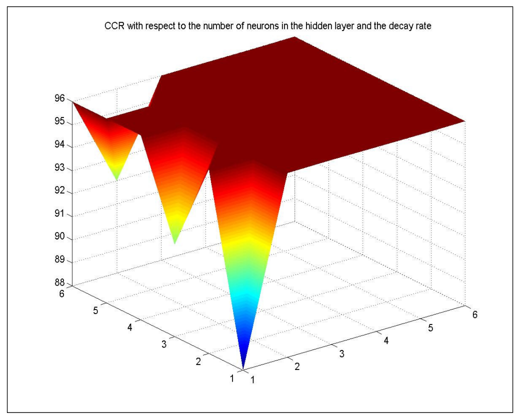

Fig. 5. CCR with respect to the number of neurons in the hidden layer and the decay rate.

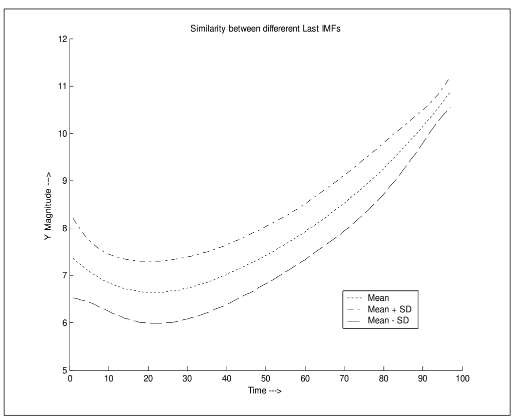

Fig. 3. Similarity between different Last IMFs

Since the largest component of the kinematic signal is common for all participants, distinguishing between them is difficult. Hence, that component has to be discarded when differences between participants' gait is to be determined. Also, the first component appears random and does not add any information as far as distinguishing between different people is concerned. Hence that component was also discarded. The rest of the components were added together and their fourier transform is taken to serve as the FV.

The length of the FV we used for recognition was 100. We used a single hidden layer MLP for classification of these FVs. Different number of neurons in the hidden layer and the decay rate were determined by experimenting. Figure 5 shows the CCR for different number of neurons (y - axis, representing [100 150 200 250 300 350] neurons) in the hidden layer and the decay rate (x - axis, representing decay rates of [0.01 0.03 0.05 0.07 0.08 1.0]). Cross validation scheme with leave one out strategy was used in our experiments. Using this figure, we selected 250 neurons in the hidden layer and a decay rate of 0.03. We obtained a best average correct classification rate of 96% using our FV. Comparative results with a few other FVs (with the same neural network configuration) are shown in table 1. Using only the sum of the IMFs (without the first and the last IMF) as the FV gives a CCR of 92%. Classifying the trajectory *as is* gives a 20% CCR (chance percentage), where as using the FFT of the trajectory as the FV (similar to the one used in [5]) gives a CCR of 64%. Though the sample size of five participants is very small, we are optimistic about the recognition capability of this technique because we used only the y-axis of a single marker to get this recognition rate. We believe, that using trajectories of multiple markers and using y-z coordinates instead of only y coordinates would allow us to extend this technique for large databases.

	1	2	3	4	5	Avg CCR
Trajectories	20	20	20	20	20	**20**
FFT	80	80	60	40	60	**64**
EMD	100	80	100	100	80	**92**
EMD FFT	100	80	100	100	100	**96**

Table 1. CCR when each of the trails were left out for different FVs and Average CCR for different FVs

	1	2	3	4	5	Avg CCR
Case 1	100	80	100	100	100	**96**
Case 2	100	60	100	100	100	**92**
Case 3	100	80	100	100	100	**96**

Table 2. Algorithm performance in the presence of noise

For gait recognition in real world, the trajectories need to be extracted from conventional 2D video. A question arises on how the proposed algorithm will perform in such a scenario, as no matter how good the tracking algorithm is, the trajectories extracted using it will be more noisy when compared to motion-captured trajectories. So, we simulated that condition by adding noise to the trajectory sequences and evaluated the performance of the algorithm. We added gaussian noise with η equal to 10% of the standard deviation of the FV and μ equal to the mean of the FV. We tested the algorithm with three scenarios; Case (1) No noise was added to both training and testing sequences, Case (2) noise was added only to test sequences and Case (3) noise was added to both training and testing sequences. The results of the test are tabulated in table 2. As we can see, the performance of the algorithm is not deteriorated in the presence of noise. Hence, we believe, the proposed algorithm can be used with real world video where the trajectories are obtained using off the shelf tracking video.

6. CONCLUSIONS

We proposed a novel algorithm for gait recognition in this paper. We found that EMD is extremely suitable for this purpose and provides great insights into the nature of gait. The algorithm also performs well in the presence of noise, because of which, it is applicable for real-world data. We used the trajectory in only one direction, of only one marker. The results of this method are extremely encouraging and motivates us to further analyze IMFs for recognition. However, the robustness of the method on large databases is yet to be evaluated. Our future work will include using data from multiple markers and automatically generating trajectories from 2D video. We are also building a gait database for testing our algorithm on a larger number of gait patterns [16].

7. REFERENCES

[1] C. Peng, J. Hausdorff, and A. Goldberger, "Fractal mechanisms in neural control: Human heartbeat and gait dynamics in health and disease," in *Self-Organized Biological Dynamics and Nonlinear Control*, J. Walleczek, Ed. 2000, Cambridge University Press.

[2] J. Cutting and L. Kozlowski, "Recognizing friends by their walk: Gait perception without familiarity cues," *Bulletin of Psychonomic Society*, vol. 9(5), pp. 353–356, 1977.

[3] M. Murray, "Gait as a total pattern of movement," *American Journal of Physical Medicine*, vol. 46(1), pp. 290–332, 1967.

[4] A. Bobick and A. Johnson, "Gait recognition using static, activity-specific parameters," in *Proc. of CVPR*, 2001, vol. 2, pp. 423–430.

[5] H. Lakany and G. Hayes, "An algorithm for recognising walkers," in *Proc. of International Conference on Audio and Video-Based Biometric Person Authentication*, J. Bigun, G. Chollet, and G. Borgefors, Eds. 1997, pp. 112–118, Springer Verlag.

[6] D. Cunado, M. Nixon, and J. Carter, "Using gait as a biometric, via phase-weighted magnitude spectra," in *Proc. of International Conference on Audio and Video-Based Biometric Person Authentication*, J. Bigun, G. Chollet, and G. Borgefors, Eds. 1997, pp. 95–102, Springer Verlag.

[7] J. Little and J. Boyd, "Describing motion for recognition," in *Proc. of International Symposium on Computer Vision*, 1995, pp. 235–240.

[8] P. Huang, C. Harris, and M. Nixon, "Human gait recognition in canonical space using temporal templates," *IEE Proc. on Vision, Image and Signal Proceesing*, vol. 146(2), pp. 93–100, April 1999.

[9] J. Hayfron-Acquah, M. Nixon, and J. Carter, "Recognising human and animal movement by symmetry," in *Proc. of ICIP*, 2001, pp. 290–293.

[10] T. Karaharju-Huisman, S. Taylor, R. Begg, J. Cai, and R. Best, "Gait symmetry quantification during treadmill walking," in *Proc. of the Seventh Australian and New Zealand Intelligent Information Systems Conference*, 2001, pp. 203–206.

[11] L. Cohen, *Time-Frequency Analysis*, Prentice Hall, 1995.

[12] N. Huang, Z. Shen, S. Long, M. Wu, H. Shih, Q. Zheng, N. Yen, C. Tung, and H. Liu, "The empirical mode decomposition and hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. of the Royal Society of London*, vol. 454, pp. 903–995, 1998.

[13] H. Liang, Z. Lin, and R. McCallum, "Artifact reduction in electrogastrogram based on the empirical model decomposition method," *Medical and Biological Engineering and Computing*, vol. 38(1), pp. 35–41, 2000.

[14] M. Carhart, *Biomechanical analysis of compensatory stepping: implications for paraplegics standing via functional neuromuscular stimulation*, Ph.D. thesis, Arizona State University, 2000.

[15] C. Myers and L. Rabiner, "A comparative study of several dynamic time-warping algorithms for connected word recognition," *The Bell System Technical Journal*, vol. 60(7), pp. 1389–1409, 1981.

[16] P. Kuchi, R. Hiremagalur, H. Huang, M. Carhart, J. He, and S. Panchanathan, "DRAG: A database for recognition and analysis of gait," in *Proc. of the ITCOM, Internet Multimedia Management Systems IV*, 2003.