Gait Recognition using Empirical Mode Decomposition

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

Abstract

Gait recognition is identifying human beings by the manner in which they walk. It has been shown by several researchers that human beings have the ability to recognize other people by their gait. Machine recognition of gait is becoming increasingly important for surveillance, aware-spaces etc. 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 nonstationary. Hence, linear analysis would be insufficient for analysis of gait. In this paper, we present our research involving nonlinear, nonstationary analysis of gait, by using a technique called empirical mode decomposition. A novel method is proposed for recognizing gait using the resulting intrinsic mode functions. Preliminary investigations reveal that this method is very effective for gait recognition. In this paper, we discuss the method, experiments, and results of the experiments in detail.

1. Introduction

Gait has been defined as the manner of walking or stepping; bearing or carriage while moving (Webster's revised unabridged dictionary). It is determined by the weight, limb length, habitual posture, bone structure, age etc. of the person. All of these dependencies make gait individual to a person. We humans often recognize people we know by their gait, even when they are at a considerable distance from us. In fact this is one of the main advantages of gait when compared to other biometrics – it can be used for recognition at a distance at which other biometrics (eg. face, iris) fail or even cannot function. Another major advantage of gait recognition is that it is unobtrusive. Though these advantages make gait recognition very desirable, it has been difficult to develop effective techniques that could achieve this purpose.

Traditionally gait was analyzed as a linear and a stationary signal. However, recent research efforts in the field of biomechanics has led to the conclusion that gait is both nonlinear as well as non-stationary [13]. Using that as our guiding principle, we explored the use of a nonlinear, nonstationary 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

One of the first experiments on the gait recognition capability of human beings was reported by Cutting et. al [4]. From then on, there have been numerous algorithms for automatic gait recognition. Some representative methods are described here. Most of these methods use either the structure of human beings or use the low-level motion content in the image as features for recognition. Based on this, these methods can be basically divided into two basic categories Structural methods and structure-free methods. These are discussed briefly in the sections below.

2.1. Structural Methods

In these methods the body structure is recovered and a set of points on the body are tracked. Their trajectories or the joint angles are used to derive features that characterize, and thereby recognize the motion or the action performed by the body. Bobick et.al. [1] propose a gait recognition technique that recovers static body and stride parameters. These parameters were tested on a database of 15 to 20 subjects and encouraging results were obtained. In [8], Lakany et.al. use the trajectories of LEDs placed on the joints and head of the subjects for gait recognition. Components of the FFT of the trajectories are used as the feature vector. A Multi-Layer Perceptron (MLP) is used for classification. One of the problems in their method is the self-occlusion of the feature points. Lakany et.al. use a Radial Basis Function (RBF) network to predict the data that is lost due to self-occlusion.

2.2. Structure-free Methods

These methods characterize the motion pattern of the body without regard to its underlying structure. In [10], Little and

Boyd developed a description of instantaneous motion, *the shape of motion*, that is different for different types of moving figures and the type of motion. In [11], Liu et.al. propose to use *frieze patterns* for gait analysis. Frieze patterns are spatiotemporal patterns that are periodic along the time axis. They used a 3D humanoid model to study the change in the frieze patterns with respect to the camera viewpoint. Using this knowledge the viewpoint is determined. Moments were used for alignment of gait, which was used for recognition. In [6], Huang et. al propose using temporal information from optical-flow changes between two consecutive spatial templates for extracting features for gait recognition. The temporal template extracted is projected into eigen space to reduce dimensionality.

One of the major problems of the above methods is that all of them have been tested on only small databases. Also, they do not consider the nonlinear, nonstationary nature of gait in their solutions. The gait movement is like a horizontally swinging pendulum that swings forward, impacts a solid surface, followed by a short period of no movement, and then swings backward. The overall signal is nonstationary due to the impact of the foot on the ground. Though neural networks can be used for nonlinear signal processing, they are often used like *black boxes* and no insight into gait can be obtained by using them. Other techniques use Fourier methods for analysis, which are insufficient for dealing with nonlinear and nonstationary signals. In this paper, we use a method that is suited for nonlinear, nonstationary signal analysis for analyzing gait.

3. Empirical Mode Decomposition

The Empirical Mode Decomposition Method (EMD) was proposed by Huang et. al. in the study of nonlinear and nonstationary signals [5]. It involves decomposing the given signal into a series of Intrinsic Mode Functions (see 3.1) through a process called *sifting* (see 3.2). The major advantage of EMD is that the basis functions are derived from the signal itself. Hence the analysis is adaptive, in contrast to the traditional methods where the basis functions are fixed. Though in Principal component 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, EMD has been applied with great success on biological signals like electrogastrogram [9]. In the following sections we elucidate on the intrinsic mode functions and the sifting method.

3.1. Intrinsic Mode Functions

A good way to define the local characteristics of a nonstationary signal is its instantaneous frequency. However, instantaneous frequency, if blindly applied to any analytic signal, may result in a few paradoxes such as [3]: (1) The instantaneous frequency may not be one of the frequencies in the Fourier spectrum, (2) The instantaneous frequency may go outside the band for a bandlimited signal.

The functions that are symmetric with respect to the local zero mean and have the same number of zero crossings and extrema give a meaningful instantaneous frequency. Based on this 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"[5]. 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 is described below.

3.2. Sifting

In 3.1 we explained that the IMFs have physcially meaningful instantaneous frequencies. However, 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 [5]. This method is described here. The signal has to have atleast two extrema – one maximum and one minimum to be successfully decomposed into IMFs using this method. 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.

3.2.1 Stopping Criterion

In practice, after a certain number of iterations, the resulting signals do not carry significant physical information, because, if sifting is carried on to an extreme, it could result in a pure frequency modulated signal of constant amplitude. To avoid this we can stop the sifting process at this stage and we do it by limiting the amount of standard deviation

(SD), computed from two consecutive sifting results. The SD is given by

$$SD = \sum_{t=0}^{T} \left[\frac{|X_{1(k-1)}(t) - X_{1(k)}(t)|^2}{X^2_{1(k-1)}(t)} \right] \qquad (2)$$

Usually, SD is set between 0.2 to 0.3. The above process 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.

$$X(t) = \sum_{j=1}^{n} C_j(t) + R_n(t) \qquad (3)$$

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.

4. Experiments

For our experiments we used the dataset of kinematic signals of five participants compiled by the motor control and rehabilitation research laboratory at Arizona State University [2]. The kinematic signals consisted of the (x, y, z) time-series of a set of fifteen markers that were placed on the participant, collected at a rate of 120 Hz. The maker locations are shown the figure 1. Five sets of data was collected for each participant. We used the raw data generated

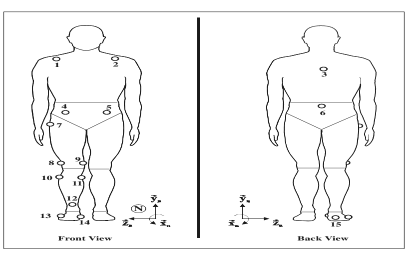

Figure 1: Marker locations used for data capture

from the equipment without any filtering. One gait cycle[1] of each trial was extracted for our purpose. The trajectory of one gait cycle is shown in the top most subplots in the figure 2. They show the y-axis displacement of the 14$^{\text{th}}$ marker.

[1]Each gait cycle is the series of events that occur between a right heel strike to a right heel strike, or between a left heel strike to a left heel strike

4.1. EMD of Gait

We performed empirical mode decomposition on the kinematic signals. For our experiments, we used the y-axis displacement of the 14$^{\text{th}}$ marker for each participant. Now gait cycles for each person, and each trail of a person, are of different length. In order to facilitate comparison, they need to be of the same length. We achieve this by normalizing using dynamic time warping [12]. We then perform empirical mode decomposition as described in the section 3.

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 top most subplots are the original signals). Careful observation of these IMFs for different participants

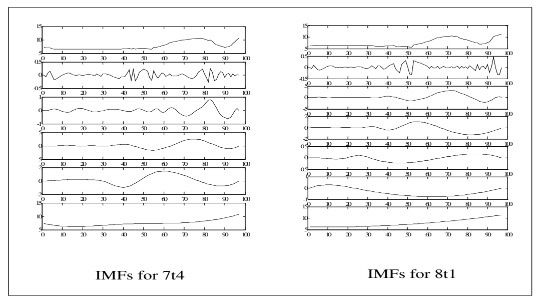

Figure 2: IMFs of participants 7 and 8

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.

2. The first IMF is different for different participants and different for different trials of the same participant.

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 is also discarded. The rest of the components were added together and their Fourier transform is taken to serve as the feature vector (FV).

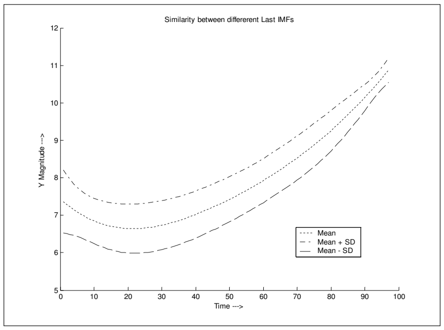

Figure 3: Similarity between different Last IMFs

	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

The length of the FV we used for recognition was 100. We used a single hidden layer MLP for classification of these FVs. Cross validation scheme with leave one out strategy was used in our experiments. We obtained an average correct classification rate of 96% using our FV. Comparative results with a few other FVs are also shown in table 1. Using the sum of the IMFs (without the first and the last IMF) as the FV yielded 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 [8]) gives a CCR of 64%. Though the sample size is very small of five participants, 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 just y coordinates would allow us to extend this technique for large databases.

6. Conclusions

We proposed a novel method for gait recognition using EMD. We found that EMD is extremely suitable for this purpose and provides great insights into the nature of gait. 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. For gait recognition in real-world situations, the trajectories of specific points have to be extracted from the conventional 2D video. Once these trajectories are extracted, the proposed method can be applied to distinguish between different people. 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 [7].

References

[1] A. Bobick and A. Johnson. Gait recognition using static, activity-specific parameters. In *Proc. of the CVPR*, volume 2, pages 423–430, 2001.

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

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

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

[5] 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*, 454:903–995, 1998.

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

[7] 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.

[8] H. Lakany and G. Hayes. An algorithm for recognising walkers. In *Proc. of the International Conf. on AVBPA*, pages 112–118, 1997.

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

[10] J. Little and J. Boyd. Describing motion for recognition. In *Proc. of International Symp. on Computer Vision*, pages 235–240, 1995.

[11] Y. Liu, R. Collins, and Y. Tsin. Gait sequence analysis using frieze patterns. In *ECCV (2)*, pages 657–671, 2002.

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

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