

DWT-DCT Based Individuals Identification Using Robust Gait Feature Images

K. M. Ibrahim Khalilullah¹, Delowar Hossain² and Md. Ekramul Hamid³

¹Lecturer, Department of Computer Science and Engineering, Bangladesh University of Business and Technology, Dhaka-1216, Bangladesh

²Ph. D Fellow, Intelligent Robotics Lab, Department of Electrical and Electronics System Engineering, Faculty of Engineering, University of Toyama, Japan,

³Professor, Department of Computer Science and Engineering, University of Rajshahi, Rajshahi-6205, Bangladesh

¹ibrahim@bubt.edu.bd, ²delowar_cse_ru@yahoo.com,

³ekram_hamid@yahoo.com

Abstract

Individual identification at a distance using gait features has newly gained growing interest from biometrics researchers. Most of the researchers have been shown that different covariate factors can affect different parts of the human body. In this paper, we propose a new approach that minimizes these difficulties, especially for carrying objects by combining static, dynamic, and part-based features. The Gait Features of the walking sequences are extracted by selecting only four sub bands of the Discrete Wavelet Transform (DWT) of the individual images. Moreover, Discrete Fourier Transform (DFT) and Discrete Cosine Transform (DCT) are implemented to extract lowest and middle frequency components that are used to create robust gait feature images (RGFIs). Then we select effective parts of the body from the Robust Gait Feature Images. After that, these parts of the body are trained using Principal Component Analysis (PCA) and Multiple Discriminant Analysis (MDA) to identify individuals. Experimental result shows promising performance in comparison with other methods.

Keywords: Gait feature, Sub band, Wavelet transform, Covariate factor, RGFI, FRR, FAR

1. Introduction

Biometrics means “life measurement” which is related to human characteristics and traits. In computer science, biometric authentication system is used for identification and access control. Individuals in groups that are under surveillance are identified using biometric authentication system. Biometric features play an important role to label and describe individuals because of their distinctive, measurable characteristics. Biometric features can be divided into physical and/or behavioral traits. Human identification based on biometrics refers to the automatic recognition of the individuals based on those traits. Physiological characteristics are related to the physical features of an individual. Examples include, but are not limited to fingerprint, face recognition, DNA, hand geometry, palm print, retina, iris. Behavioral characteristics are related to the unique pattern of a person’s behavior, most of them are typing rhythm, gait, and voice. Among them, gait has recently gained considerable attention from biometric researchers as a promising biometric feature.

The manner of how a person walks is called the gait. Since every individual has his/her own walking-scheme, gait is a distinctive feature. Human walking involves harmonized

movements of body parts, joints and the interaction among them [1]. As a result, the way of walking of a person is one of the unique features. For this reason, these features may be used for person identification. The key advantage of individual identification using gait comes from the fact that it can be done at a distance without the subject's knowledge, which, by itself, is an advantage when compared to other biometric features such as iris or face recognition. Murray, *et al.*, [2] medical studies explained that individual gaits are unique, differing from person to person and are difficult to disguise. In addition, Cutting, *et al.*, [3] showed that human gaits are so identical and we recognize individuals by their gait.

Person identification based on gait feature analysis typically progresses by extracting the silhouette of the walking person. Then the gait information is used for identification. The gait features used for recognition can be broadly divided into two categories, such as static and dynamic features. Wang, *et al.*, [4] showed in the recent work that we can achieve a promising recognition rate for person identification using stationary gait features. On the other hand, Cutting, *et al.*, [3] argued that gait-related features contribute significantly more in individual recognition than body-related features. Moreover, Stevenage, *et al.*, [5] supported the argument that dynamic features are more important than static features for the purpose of human identification. The body-related features represent the static information and the recent experiments proved that a prosperous recognition rate can be achieved for individual identification using stationary features.

But the main complexity is that body-related properties can be affected by various contexts such as different clothing and carrying objects; these covariate factors make difficulties for individual identification. Therefore, the body-related parameters are not potential as they are dependent on clothing, bags, and other factors [6, 7]. These covariate factors are not part of the gait information which would badly attack to the recognition performance and there is no solution provided for covariate factors in above experiments and therefore would be impractical for individual identification especially when people put in appearance with different clothing and carrying bags. This is the main challenging problem which complicates the identification problem. However, changes in appearance due to clothing variations and carrying bags are likely to change particular parts of the silhouettes and the ability of classification of subjects decreases with respect to these parts. Altab, *et al.*, [8] proposed a part-based method to improve recognition rate in this situation. The method used DFT for reducing noise from the silhouette image sequences and extracts eight parts of the human body.

From the above argumentative statements, we extend the previous idea to improve the identification performance by using DWT on individual image sequences and combining static, dynamic, and part-based features. The gait features of silhouette image sequences are calculated in frequency domain. For that at first we use Discrete Wavelet Transform instead of only Discrete Fourier Transform because of its better time-frequency localization properties. Wavelets have the great advantages of being able to put apart the fine details in a signal. Very small wavelets can be used to sequester very fine details in a signal, while very large wavelets can identify coarse details. Using DWT, we extract gait features of the walking sequences. Thereafter we apply DFT and DCT to extort appropriate frequency components. Two classical linear approaches are used for finding best data separation with dimensionality reduction. They are principal component analysis (PCA) and multiple discriminant analysis (MDA), which is efficiently used in human face identification. The Multiple discriminant Analysis is used to recover limitations of the Principal Component Analysis because of maximizing the difference between values of the dependent. Principal Component Analysis seek outs a projection that best represents the data in the least-square sense, while Multiple discriminant Analysis seeks a projection

that best separates the data in least-square sense. As a result, these combined techniques can reduce the data dimensionality and optimize the class separability simultaneously.

2. Gait Dataset with Various Covariate Factors

Relatively small databases were used in the early approaches. Obviously, the achievement and progress of a new application relies largely on the dataset use for evaluation. At the early stage, gait datasets were limited in terms of subjects or clothing variations. The small Soton database [9] includes some clothing variations such as trainers, rain coats, and trench coats; however, there are only 12 subjects. On the other hand, the large Soton database [9], the HumanID dataset [10], and the CASIA dataset [11, 12] have more than 100 subjects, but the variations in clothing and carrying objects are limited. As gait is a behavioral biometric, there is plenty of scope for intra-subject variation, including clothing or other covariate factors. Most of the algorithms have been used only single intra-subject variation like clothing or carrying conditions. Therefore, we construct our own large-scale gait dataset with varying clothing and different carrying condition from CASIA and SDUMLA-HMT [13] dataset. In SDUMLA-HMT dataset, the carrying object could be a back bag, a shoulder bag, a knapsack, a satchel, or a handbag chosen according to the subject's preference. The examples of the subject with different covariate factors are shown in Figure 1.

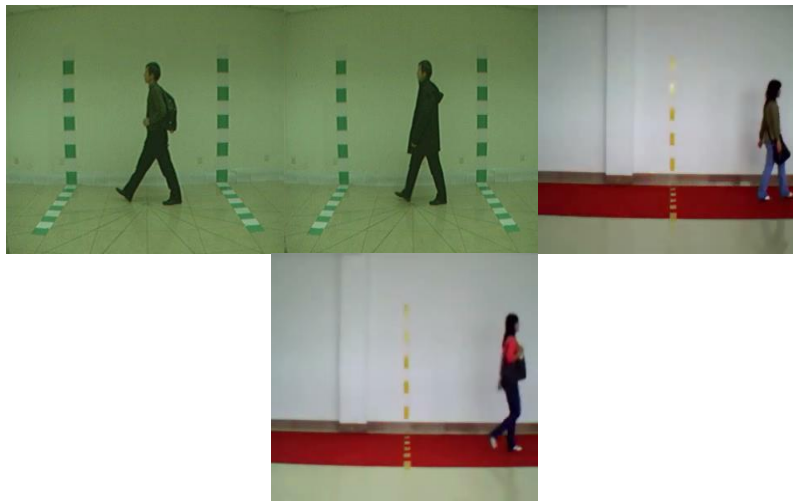


Figure 1. Original Video Images with Different Clothing and Carrying Objects

3. Feature Extraction

3.1. Pre-processing

In our proposed algorithm, silhouettes are the basis of gait based individual identification. After extracting silhouette images from video data or image sequence, shadow or small hole inside the object is found in most of the extracted silhouette images. It is appeared as part of the subject body in the binary human silhouette image as shown in Figure 2. The presence of the noise affects the gait feature extraction and reduces the identification accuracy. We reduce the noise by applying morphological image processing operations. Then the enhanced silhouette images is normalized to a fixed size and centralized as shown in Figure 3. After this preprocessing, the gait period is estimated in an individual's walking sequences. In pre-processing step, erosion and dilation operations are applied to the binary silhouette images respectively for enhancing image quality. The

diamond structuring element is used to remove the small hole in the erosion operation and square structuring element is used in the dilation operation.

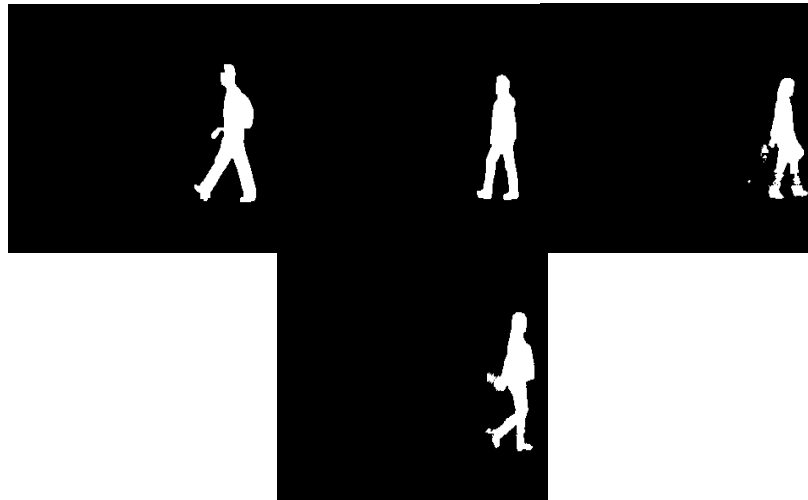


Figure 2. Binary Silhouettes with Noise e.g., Shadow or Small Hole Inside the Object



Figure 3. Normalized and Enhanced Binary Silhouette Images

3.2. Gait Period Estimation

Human walking pattern is a periodic signal. Since our proposed gait feature templates depend on the gait period, we must estimate the number of frames in each gait period. A single walking cycle can be regarded as that period in which a person changes location from the mid-stance (both legs are closest together) position to a double support position (both legs are farthest apart), then the mid-stance position, followed by the double support position, and finally back to the mid-stance position [14].

Based on Sarkar *et al.*, work [10] the gait cycle is estimated by calculating the number of foreground pixels in the lower part of the silhouette image. When both legs are closest together, the silhouette image holds the smallest number of foreground pixels. In double support position, the silhouette image contains the greatest number of foreground pixel. Since severe changes in the gait cycle are most obvious in the lower part of the body, gait period estimation makes use only of the lower half of the silhouette image, with the gait period (N_{gait}) being the distance between three consecutive minima. This gait period value is used in Robust Gait Feature Images calculations.

3.3. Robust Gait Feature Image

Robust Gait Feature Image (RGFI) is obtained by applying Discrete Wavelet Transformation, Discrete Fourier Transformation, and Discrete Cosine Transformation on

the binary silhouettes of a human over one gait period (N_{gait}). It is the time-normalized accumulative energy image of human walking in a complete cycle. To reduce the calculation complexity, we generate Gait Period Volume (GPV) from the number of binary silhouette images in one gait period. If a binary silhouette image is denoted as then the GPV is written as, $GPV = [S_1 S_2 S_3 \dots S_{N_{gait}}]$, where $S_1 S_2 S_3 \dots S_{N_{gait}}$ are the binary silhouette image sequences in one gait period. After applying DWT on GPV, we get approximation coefficients vector and detail coefficients vector. The approximation coefficients vector contains low-frequency components and detail coefficients vector contains high-frequency components. After this, DFT is applied on approximation coefficients vector and DCT is applied on detail coefficients vector. Average amplitude spectra from approximation coefficients vector is calculated at each pixel by DFT based on the gait period. DFT is used to extract 0-time frequency component and 1-time frequency component from the approximation coefficients vector. Although high-frequency elements contain noise rather than meaningful information, but all coefficients of the detail coefficients vector do not contain noise. For this, DCT is used to pick the detail coefficients vector to extract a few coefficients that represent 99% of the energy in the sequence. We can conclude that DFT is used to extract lowest frequency components and DCT is used middle frequency components. The subsequent calculation of the average amplitude spectra and RGFI for the approximation coefficients vector from GPV is summarized in the following algorithm:

Step-1: Repeat Step-2 for $x = 1$ to M

Step-2: Repeat Step-3 to Step-8 for $y = 1$ to N

Step-3: Initialize t and j ; $t = y$, $j = 1$

Step-4: Repeat while ($t \leq N_{gait} + y$)

(a) $G(j) = GPV(x, t)$

(b) $t = t + N_{gait}$

(c) $j = j + 1$

Step-5: Apply DWT on G to select only three sub bands, one of them is approximation coefficients vector and other two of them are detail coefficients vector.

Step-6: Apply DFT on the approximation coefficients vector to extract lowest frequency components.

Step-7: Calculate average amplitude spectrum (A) for DFT of the approximation coefficients vector.

Step-8: Create RGFI, $RGFI(x, y) = A$.

For the detail coefficients vector use the similar approach except that we apply DCT on G . After this, we extract a few coefficients from the DCT signal. Then we calculate energy (E) of the extracted coefficients f_c for the detail coefficients vector at the position of a pixel by the following equation:

$$E = \sqrt{\sum_{c=1}^k \left(|f_c|^2 \right)} \quad (1)$$

Finally we calculate RGFI for the detail coefficients vector. In both cases, the intensity of a pixel in RGFI reflects the energy of body occurrence at the position of this pixel for the estimated gait period. All RGFI for one gait period of a person is shown in Figure 4.

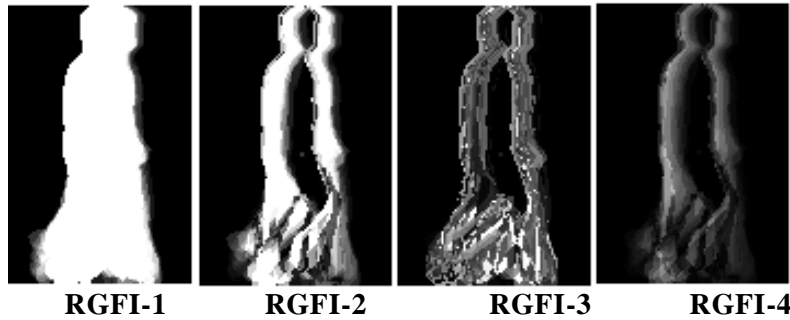


Figure 4. Robust Gait Feature Images

As certain parts of the body are desparingly affected by different carrying conditions, we divide the RGFI into three Sections to reduce the sensitivity to the carrying conditions. The definition of this body parts is inherited from known anatomical properties [15]. For a body height H , the RGFI is segmented according to the vertical position of the neck, waist- pelvis, and knee. The first part covers from the top of the head to the neck; the second part from the neck to the waist-pelvis, and the final part from the waist-pelvis to the floor. The process of features extraction is illustrated in the following Figure 5 through a block diagram:

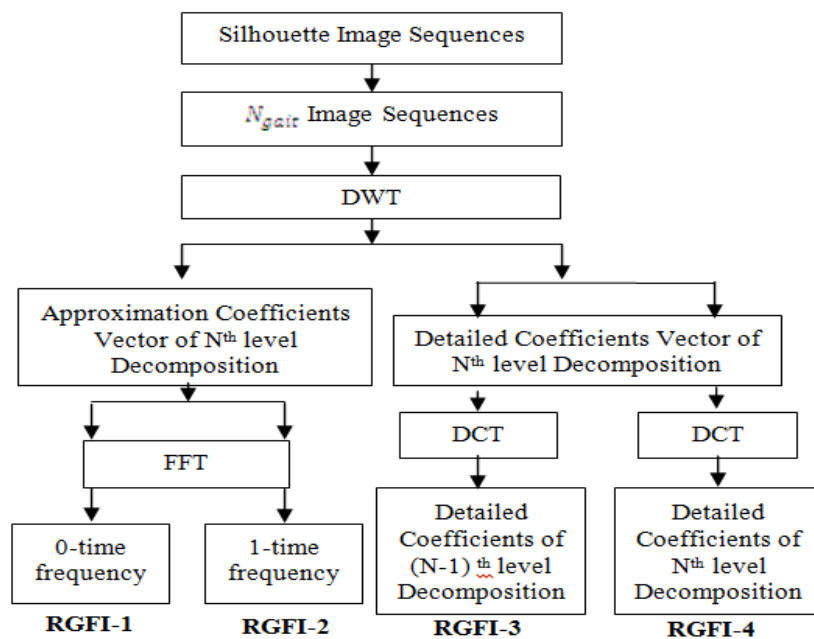


Figure 5. Block Diagram of Feature Extraction Process

4. Identification and Performance Evaluation

4.1. Feature Matching

The feature matching Section represents the level of similarity between the testing data and the training data. In this Section, the detail of the similarity measures in our proposed algorithm is explained. In this paper, we consider four Robust Gait Feature Images. Each RGFI is divided into three parts. Each part is considered as a training RGFI template and each part of all RGFIs is trained separately. Two classical linear approaches are used for finding transformations for dimensionality reduction. They are principal component analysis (PCA) and

multiple discriminant analysis (MDA), which is efficiently used in human face identification [16]. Huang, *et al.*, [17] merge PCA and MDA to achieve the best data representation and the best class discriminability simultaneously. In the feature matching procedure, we use this combination approach. Given n d -dimensional training RGFI templates $\{G_1, G_2, \dots, G_n\}$, PCA minimizes the function

$$Q_{d'} = \sum_{k=1}^n \left\| \left(\text{mean} + \sum_{i=1}^{d'} a_{ki} e_i \right) - G_k \right\|^2 \quad (2)$$

where $d' < d$, $\text{mean} = \frac{1}{n} \sum_{k=1}^n G_k$, and $\{e_1, e_2, \dots, e_{d'}\}$ is a set of unit vectors. $Q_{d'}$ is minimized when $e_1, e_2, \dots, e_{d'}$ are the d' eigenvectors of the scatter matrix S having the largest eigenvalues, where

$$S = \sum_{k=1}^n (G_k - \text{mean})(G_k - \text{mean})^T \quad (3)$$

The d' -dimensional feature vector H_k is obtained from the G_k as follows:

$$H_k = M_{pca} G_k = [a_1, a_2, \dots, a_{d'}]^T = [e_1, e_2, \dots, e_{d'}]^T G_k, \quad k = 1, \dots, n \quad (4)$$

Suppose that the n d' -dimensional principal component vectors $\{H_1, H_2, \dots, H_n\}$ belong to c classes. MDA finds a transformation matrix W that maximizes the ratio of the between-class scatter matrix S_B to the within-class scatter matrix S_W

$$Q(W) = \frac{|S_B|}{|S_W|} = \frac{|W^T S_B W|}{|W^T S_W W|} \quad (5)$$

The within-class scatter S_W is defined as

$$S_W = \sum_{i=1}^c S_i, \quad (6)$$

where,

$$S_i = \sum_{H \in \nu_i} (H - \text{mean}_i)(H - \text{mean}_i)^T \quad (7)$$

and

$$\text{mean}_i = \frac{1}{n_i} \sum_{H \in \nu_i} H \quad (8)$$

where ν_i is the training template set that belongs to the i^{th} class and n_i is the number of templates in ν_i . The between-class scatter S_B is defined as

$$S_B = \sum_{i=1}^c n_i (\text{mean}_i - \text{mean})(\text{mean}_i - \text{mean})^T \quad (9)$$

where,

$$\text{mean} = \frac{1}{n} \sum_{H \in \nu} H \quad (10)$$

and ν is the whole training template set. $Q(W)$ is maximized when the columns of W are the generalized eigenvectors that correspond to the largest eigenvalues in

$$S_B W_i = \lambda_i S_W W_i \quad (11)$$

There are no more than $c-1$ nonzero eigenvalues, and the transformation matrix is formed using corresponding eigenvectors V_1, \dots, V_{c-1} . The $(c-1)$ -dimensional

feature vector ζ_k is obtained from the d' -dimensional principal component vectors H_k :

$$\zeta_k = M_{mda} H_k = [V_1, \dots, V_{c-1}]^T H_k, \quad k = 1, \dots, n. \quad (12)$$

For each training gait template, its gait feature vector is obtained as follows

$$\zeta_k = M_{mda} M_{pca} G_k = TG_k, \quad k = 1, \dots, n. \quad (13)$$

The obtained feature vectors represent the n templates for individual identification. Similarly, the feature vector ζ is calculated from input RGFI template. We now determine which RGFI template class provides the best description for the input RGFI template. This is done by minimizing the Euclidean distance,

$$\varepsilon_k = \|\zeta - \zeta_k\|^2 \quad (14)$$

The input RGFI template is considered to belong to a known class if ε_k is below an established threshold θ_ε . Then the person is identified as a known person.

4.2. Experimental Results and Comparison

In this paper, our experiment is emphasized on the different carrying conditions. A robust database with different carrying conditions and clothing variation are generated from CASIA and SDUMLA-HMT databases. We use 40 subjects in the training and testing dataset. In this experiment, two different walking sequences are considered. One of the two sequences, individuals with normal conditions, is used for training. Another gait dataset with clothing variations (*i.e.*, jacket, casual dress) and different carrying conditions (*i.e.*, back bag, shoulder bag) is used in the testing dataset. The DWT is applied at each pixel based on the gait period. According to the algorithm after performing some operations on selected four sub bands of the DWT, we get RGFIs. These features include both stationary (*e.g.*, body shape, walking posture) and potential (*e.g.*, stride, arm swing) components. Each feature image is divided into three parts *e.g.*, upper, middle, and lower part. All parts are trained separately in the PCA subspace using the feature images and dimension-reduced features are used as input in the multiple discriminant analysis for better identification.

The performance of individual identifications is evaluated by calculating false rejection rate (FRR) and false acceptance rate (FAR). The FRR and FAR of different number of subjects are shown in Figure 6 and Figure 7, respectively.

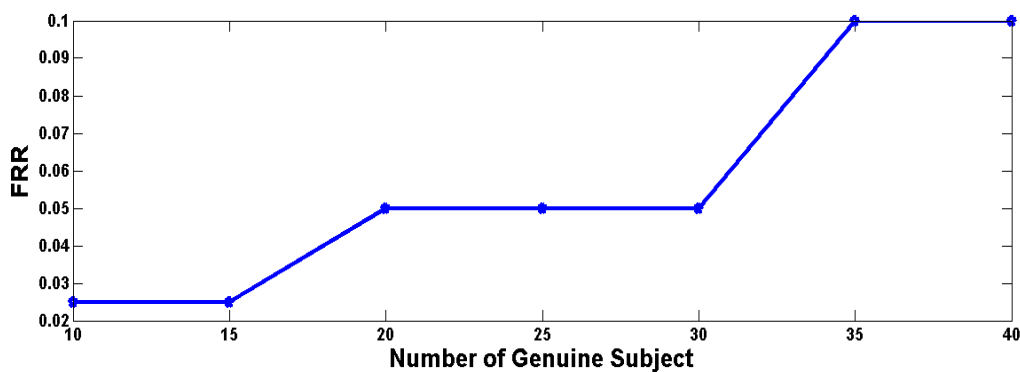


Figure 6. False Rejection Rate for Performance Evaluation

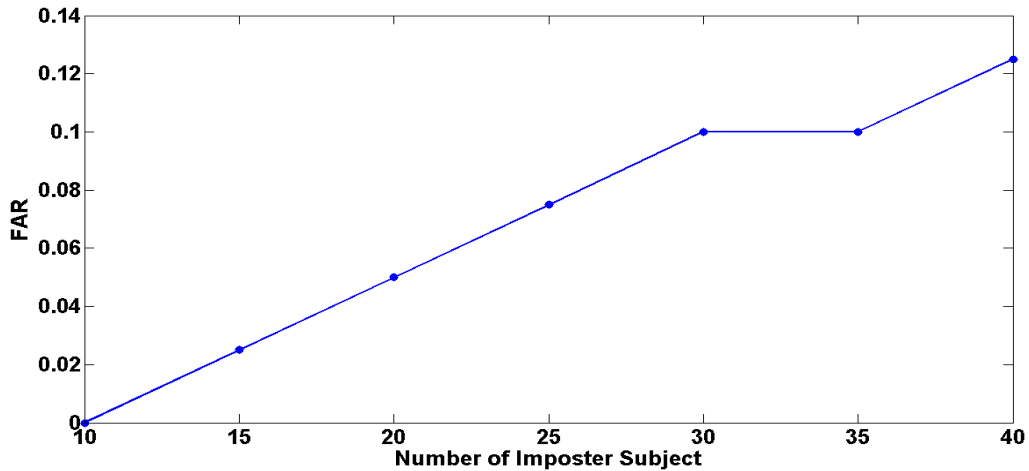


Figure 7. False Acceptance Rate for Performance Evaluation

The accuracy of individual identifications is compared with other three algorithms [14, 18, and 19] in which they used various covariate factors especially different carrying conditions. The DSST-PCA method [14] classified 88.9% of the walking sequences. An average recognition rate of 73.4% is achieved by KNN-classifier based method [18] whereas the baseline algorithm performance varies from 73% on the simplest case to 10% and 30% on the hardest experiments [19]. The DWT-DCT based method gives better results than other three methods as shown in Figure 8.

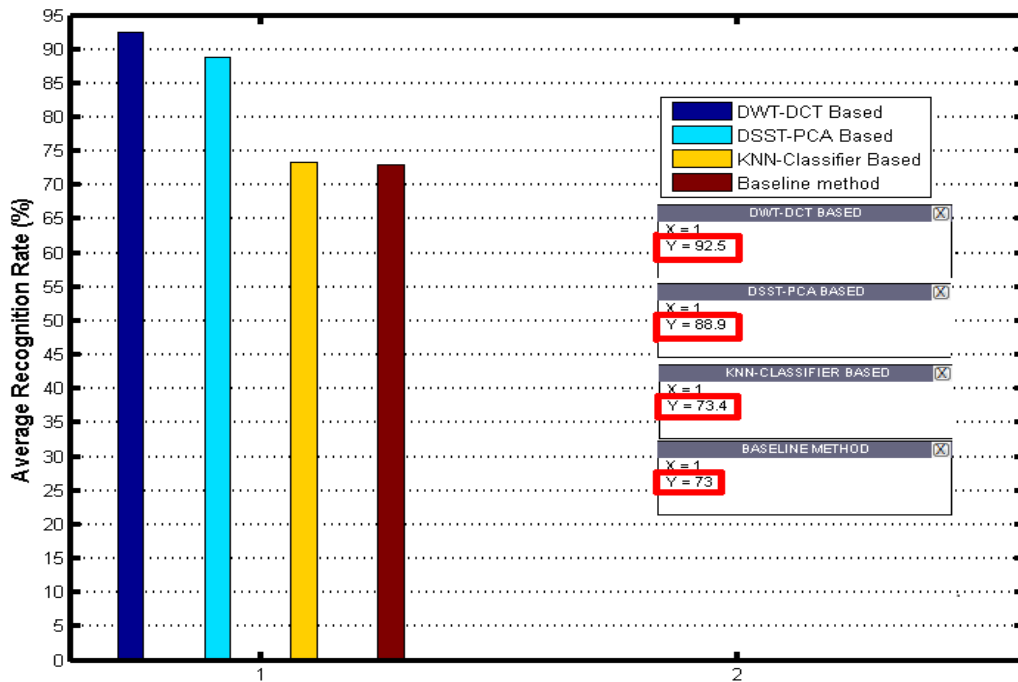


Figure 8. Bar Graph for Comparison of Methods

5. Conclusion

In this paper, a gait recognition algorithm is proposed for individual identification using DWT-DCT and PCA-MDA based technique. In the feature extraction phase, we have used DWT and DCT. The PCA and MDA have been used in the identification phase. The performance of the proposed algorithm has been evaluated experimentally

using the constructed robust dataset; confirm that the proposed method is largely robust in comparison with other methods. Furthermore, we have selected the four sub bands of the DWT willingly, just using some previous experiences. We didn't use any technique to select appropriate sub bands. Therefore another future direction would be to apply some new techniques for selecting adaptive sub bands that might improve accuracy of the identification.

Acknowledgements

I would like to thank to the provider of CASIA and SDUMLA-HMT Gait databases. These databases have made my research work very easy.

References

- [1] C. BenAbdelkader, R. Culter, H. Nanda and L. Davis, "EigenGait: motion-based recognition of people using image self-similarity", in Proc. of International Conference Audio and Video-Based Person Authentication, (2001), pp. 284-294.
- [2] M. Murray, A. Drought and R. Kory, "Walking Patterns of Normal Men", Journal of Bone Joint Surgery, vol. 46-A, (1964), pp. 335-360.
- [3] J. E. Cutting and D. Proffitt, "Gait perception as an example of how we may perceive events", Intersensory perception and sensory integration, (1981), pp. 249-273.
- [4] L. Wang, T. Tan, H. Ning and W. Hu, "Fusion of static and dynamic body biometrics for gait recognition", IEEE Transactions on Circuits and Systems for Video Technology, vol. 14, no. 2, (2004), pp. 149-158.
- [5] S. V. Stevenage, M. S. Nixon and K. Vince, "Visual Analysis of Gait as a Cue to Identity", Applied Cognitive Psychology, vol. 13, no. 6, (1999), pp. 513-526.
- [6] G. V. Veres, M. S. Nixon and J. N. Carter, "Modelling the Time-Variant Covariates for Gait Recognition", in Proc of 5th Intl conf on Audio-and Video-Based Biometric Person Authentication, (2005), pp. 597-606.
- [7] D. K. Wagg and M. S. Nixon, "On automated modelbased extraction and analysis of gait", Proc. 6th IEEE Int'l Conf. Automatic Face and Gesture Recognition, (2004), pp. 11-16.
- [8] Md. A. Hossain, Y. Makihara, J. Wang and Y. Yagi, "Clothing-invariant gait identification using part-based clothing categorization and adaptive weight control", Pattern Recognition, vol. 43, no. 6, (2010), pp. 2281-2291.
- [9] J. Shutler, M. Grant, M. Nixon and J. Carter, "On a large sequence-based human gait database", in: Proceedings of the Fourth International Conference on Recent Advances in Soft Computing, Nottingham, UK, (2002), pp. 66-71.
- [10] S. Sarkar, J. Phillips, Z. Liu, I. Vega, P. Grother and K. Bowyer, "The humanid gait challenge problem: data sets, performance, and analysis", Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 2, (2005), pp. 162-177.
- [11] S. Zheng, J. Zhang, K. Huang, R. He and T. Tan, Robust View Transformation Model for Gait Recognition", International Conference on Image Processing (ICIP), Brussels, Belgium (2011).
- [12] S. Yu, D. Tan and T. Tan, "A Framework for Evaluating the Effect of View Angle, Clothing and Carrying Condition on Gait Recognition", In Proc. of the 18'th International Conference on Pattern Recognition (ICPR), vol. 4, (2006) August, pp. 441-444, Hong Kong, China.
- [13] Y. Yin, L. Liu and X. Sun, "SDUMLA-HMT: A Multimodal Biometric Database", The 6th Chinese Conference on Biometric Recognition (CCBR 2011), LNCS 7098, (2011), pp. 260-268, Beijing, China.
- [14] Y. Pratheepan, J. V. Condell and G. Prasad, "The Use of Dynamic and Static Characteristics of Gait for Individual Identification", 13th International Machine Vision and Image Processing Conference - IMVIP, (2009).
- [15] W. Dempster and G. Gaughran, "Properties of body segments based on size and weight", American Journal of Anatomy, vol. 120, no. 1, (1967), pp. 33-54.
- [16] P. N. Belhumeur, J. P. Hespanha and D. J. Kriegman, "Eigenfaces vs. fisherfaces: recognition using class specific linear projection", IEEE Trans. Pattern Anal., Intell., vol. 19, no. 7, (1997) March, pp. 711-720.
- [17] P. S. Huang, C. J. Harris and M. S. Nixon, "Recognizing humans by gait via parametric canonical space", Artif. Intell. Eng., vol. 13, (1999), pp. 359-366.
- [18] I. Bouchrika and M. Nixon, "Exploratory Factor Analysis of Gait Recognition", In 8th IEEE International Conference on Automatic Face and Gesture Recognition, (2008), pp. 1-6.
- [19] P. J. Phillips, S. Sarkar, I. Robledo, P. Grother and K. W. Bowyer, "Baseline Results for the Challenge Problem of Human ID Using Gait Analysis", 5th International Conference on Automatic Face and Gesture Recognition, (2002) May.

Authors



K. M. Ibrahim Khalilullah, obtained his B.Sc. and M.Sc. degree in Computer Science and Engineering from University of Rajshahi, Rajshahi-6205, Bangladesh in the year of 2006 and 2007 respectively. He is now working as a Lecturer in the department of Computer Science and Engineering, Bangladesh University of Business and Technology, Dhaka-1216, Bangladesh. His research interests include Multispectral Signal Analysis, Multidimensional Multirate Signal Processing, Biomedical Signal and Image Processing, Biometric Systems, Robotics and Computer Vision.



Delowar Hossain, obtained his B.Sc. and M.Sc. degree in Computer Science and Engineering from University of Rajshahi, Rajshahi-6205, Bangladesh in the year of 2009 and 2010 respectively. He is a PhD student in the department of Electrical and Electronics System Engineering at Toyama University in Japan. His research interests are in Robotics, Image and Signal Processing.



Ekramul Hamid, received his B. Sc and M. Sc degree from the Department of Applied Physics and Electronics, Rajshahi University, Bangladesh. After that he obtained the Masters of Computer Science from Pune University, India. He received his PhD degree from Shizuoka University, Japan. During 1997-2000, he was a lecturer in the Department of Computer Science and Engineering, Rajshahi University. Dr. Hamid was working as Assistant Professor at the King Khalid University, Abha, KSA from 2009 to 2011. He worked as a visiting researcher in Shizuoka University, Japan in 2012 and 2014, respectively. He is currently working as Professor and Chairman in the Department of Computer Science and Engineering, Rajshahi University. His research interests include Digital Signal Processing, Analysis and synthesis of speech signal, Speech Enhancement and Image Processing.

