ISSN 2320 - 2602
Volume 2, No.12, December 2013
International Journal of Advances in Computer Science and Technology
Available Online at http://warse.org/pdfs/2013/ijacst032122013.pdf

Extraction of Shape-Invariant Gait Features for Human Gait Recognition

Md. Shariful Islam[1], Md. Shafiul Azam[1]
[1]Dept. of Computer Science and Engineering, Pabna University of Science and Technology, Pabna, Bangladesh,
sharif_cse_ru@yahoo.com, shahincseru@gmail.com

ABSTRACT

In appearance based human gait recognition, minimizing the variations of different cofactors such as viewing angles, carrying objects, clothing are the key challenging problems. To extract the shape-invariant gait features we have adopted a qualitative spatial and temporal reasoning based feature extraction technique. First we extract the space time shapes (STS) silhouettes from a gait sequence including three important parameters space, time and viewing angle. Although the viewing angle is fixed but other two parameters are used in accounting the variations due to different cofactors with the time. Thus, these two parameters are fascinating in reasoning inference rules to generate the final cognitive maps of each sequence. Experiments are conducted on the benchmarking CASIA dataset B and we have got much better result compared to others classical gait recognition approaches.

Key words : Spatial, Gait, Recognition, , Shape, Invariant.

1. INTRODUCTION

Human gait is now proved as a promising biometric feature for individual recognition at a distance without the subject cooperation [1]. In early medical science [2] reported that normal gait has eight sub-phases and 24 different components that are unique for individual recognition. Lee et al. [3] defined gait as a combination of both the appearance and the dynamics of human walking motion. It has been proved in [2] that the dynamic ambulation of gait involves a stance phase of 60 percent of gait cycle and a swing phase of 40 percent of gait cycle. While other biometric features like face, iris, fingerprint etc., require relatively high image resolution for person authentication, gait can be available even at low image resolution [3]. Moreover, gait is a distinctive feature and is difficult to hide. It has been also adapted to vast applications in large areas of the society involving security-sensitive video surveillance like public transportation, airport, banks, car park monitoring, and identifying perpetrators at a crime scene [4]. In addition, it can be used as clinical assessment to evaluate medical disorders [2], gender, ages and emotions

Gait recognition methods can be mainly divided into two categories: model-based and appearance-based. Model-based methods [6,7] use human model parameter that fit into the shape and motion of gait for feature extraction. The computational cost of the model-based methods is significantly high because of its complex matching, searching and model fitting errors. Appearance-based methods [1,8,9] are relatively insensitive in capturing the actual shape/appearance information of gait silhouettes at a distance from the camera. Moreover, it has a significant benefit of low computational costs comparing to model-based methods.

In this paper the cost effective appearance-based gait representation are considered to extract the shape-invariant gait features. First we extract the space time shapes (STS) silhouettes from a gait sequence including three important parameters space, time and viewing angle. To determine the body regions where the appearances change, we have adopted a fixed size rectangular window. We have developed a comprehensive rectangle representation that minimizes the variations between 2 rectangles: one is normal gait sequence (gallery) and another is targeting gait sequence (probe). Thus the extracted gait features are comprised of most effective body regions with minimizing the large intra-class variations and show much better performance.

2. RELATED WORKS

The appearance style of such more powerful gait can be changed by the different covariate conditions such as 1) viewing angles (i.e., different camera positions) 2) elapsed time (i.e., changing walking speed) 3) carrying objects (i.e., belonging with hand bag/side bag/backpack) 4) clothing combinations (i.e., wearing different challenging clothes like as skirt, down jacket, rain coat etc) 5) walking surface (i.e., grass and concrete surface) 6) ages (i.e, walking in different age times) 7) other factors, such as injury, disguise, image quality [38], observer' familiarity with the people under surveillance, lighting, background, and foot wear (shoe types). As an example, suppose a gallery subject wears standard clothing like regular pants and full shirt. If a probe subject wears a raincoat or long coat or carries objects or changes view angle then a large body parts of the gait is changed significantly [1,5]. In fact, it is true that the areas in changes of appearance by these challenging co-factors are variable around the whole gait signature.

Several methods have been proposed to solve the problem of different covariate conditions in gait recognition. All of the appearance based methods are mainly divided into two broad categories: whole-based (i.e., without division the whole body) and part-based (i.e. divide the whole body into individual part). Whole-based methods [1,8,9] are mainly focus on gait representation techniques. The gait signature is represented by extracting and selecting the most discriminate information to improve the performance. When the style of gait is affected by the challenging covariate conditions, it faces difficulties for individual identification. Moreover, it is faced in high computational cost for high dimensional spaces. It is observed from the part-based works [5,10] that the straight forward parts definition not only minimize the variations of different covariate conditions, but also individual differences which are crucial for person identification. In the previous work [11,12], we have developed a systematic approach to defined the effective and less effective body parts and a new gait representation technique to minimize the challenging variations in gait recognition.

3. HUMAN GAIT REPRESENTATION

To represent the most robust gait feature [8] proposed spatial-temporal average silhouettes over a complete gait cycle called Gait Energy Image (GEI). Most of the recent works are used GEI as for input gait feature which is introduced in below.

3.1 Silhouettes Extraction

To compute the Gait Energy Image (GEI), first the binary silhouettes of a gait sequence are extracted from the background subtraction of a video frames. The main thing is to separate the target object from a predefined reference background template. The predefined template is modelled by using the pixel statistics that represent the background scene. The pixel statistics that's mean the RGB values of the background template involve two parameters: mean and covariance. So a pixel is modelled by two tuple <E^i ,S^i>. Where E^i is the expected color value and S^i is the covariance of color value. The covariance color value of a pixel i is given by

$$S^i = [\sum R(x,y), \sum G(x,y), \sum B(x,y)] \qquad (1)$$

And the expected color value of a pixel i is given by

$$E^i = [\mu_R^i, \mu_G^i, \mu_B^i] \qquad (2)$$

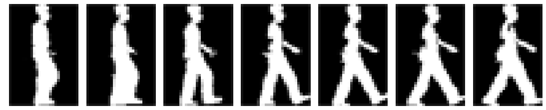


Figure 1. Background subtraction normalized and aligned silhouettes of a gait cycles from a gait sequence.

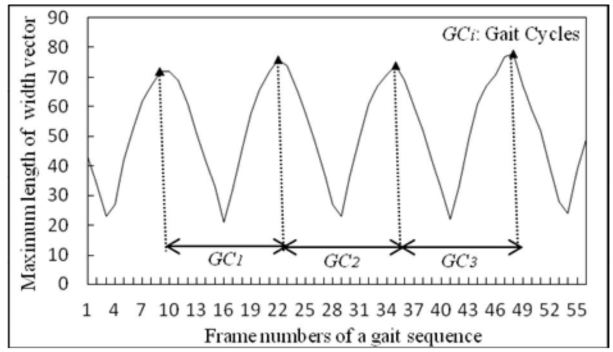
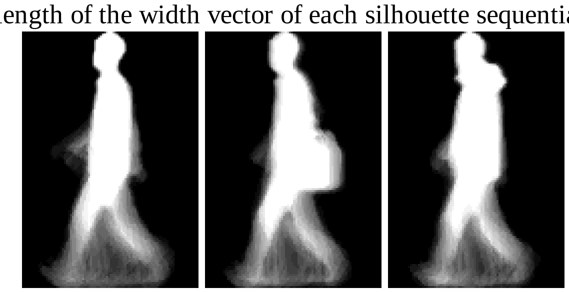


Figure 2. Gait cycles detection of a video sequence

The difference between the background image and the target object is calculated using the Euclidian distance. Then the pixels are classified into foreground or background using Expectation Maximization (EM) with a Gaussian mixture model. After the extraction of the target object then the noise free silhouette of a video frame is registered with the normalization and aligned into a fixed size image given in Fig 1. In this work the silhouette size is 128x88.

3.1 Gait Cycles Detection

A gait sequence is a series of repeated cycles where each cycle is called a gait cycle or gait period. To detect the gait cycles first we have computed the width vector of each silhouette in each frame over time. The length of the width vector will reach a maximum when the two legs are farthest apart (full stride stance) and drop to a minimum when the legs over-lap (heels together stance). To increase the sensitivity, we consider the width vector from the legs region, which are selected simply by considering only the bottom half of the silhouette. Fig. 2 shows an illustration of the variation of the length of the width vector of each silhouette sequentially.



(a) (b) (c)

Figure 3. Gait Energy Images (GEIs) of a subject from CASIA dataset: (a) normal walking sequence (b) walking sequence with carrying bag and (c) walking sequence with bulky-coat

Notice that two consecutive full strides stance constitute a gait cycle. We compute the two successive frame numbers with the two successive minima and then the maxima of the frame numbers between two minima is computed. Finally the two frame numbers with the two successive maxima (marked by a marker symbol in Fig. 2) are selected as starting and ending points of a gait cycle. Applying this technique, we can get all the gait cycles from a whole gait sequence. It may be more accurate technique for gait cycles detection in the presence of outlier's effect. We can see that from the Fig 2 there are three gait cycles (GC) are detected from a walking sequence of a subject.

3.1 Gait Energy Images (GEIs)

Finally Gait Energy Image (GEI) is computed based on computing the average silhouettes of a complete gait cycle. Suppose there are N frames in a gait sequence $F = \{F(1), F(2),...F(N)\}$. Let the total number of gait cycles is N_{gait}, denoted by $F_{gc} = \{F_{gc}(1), F_{gc}(2), ..., F_{gc}(N_{gait})\}$. Gait Energy Image GEI_i for a gait cycle $F_{gc}(i)$ $\{i=1,...,N_{gait}\}$ is computed as:

$$GEI_i = F_{gc}(i) = \frac{1}{M}\sum_{f=1}^{M} I(x, y, f) \qquad (3)$$

where M is the total number of frames in a complete gait cycle F_{gc}, I is a silhouette image whose pixel coordinates are given by x and y, and f is the frame number in the gait cycle. Example of Gait Energy Images GEIs for individual recognition is shown in Fig. 3. The image represents the normal walking sequence without carrying a bag or wearing a bulky-coat. The second and third images represent the GEIs where the subjects carrying a bag and wearing a bulky-coat respectively. GEI actually represent the gait feature by a single template from a sequence of walking posture for a gait cycle. Although GEI losses the style of walking sequences but it can represent the gait with the more robust posture of gait and noise free. The high intensity values in GEI hold the static (e.g. head, torso) information which is move little during a walking time period. The low intensity values in GEI hold the dynamic information which is move constantly during a walking time period.

4. PROPOSED SHAPE-INVARIANT GAIT FEATURES

In the light from the above study of gait, the question is which body regions should be effective under the change of appearance of gait by the different covariate conditions, are still an open problem. It is difficult to cop all the covariate conditions in training session and also in real world applications. So we need a gait recognition engine that it can tackle these problems smartly. The proposed method can be used as a recognition engine so that it can minimize the variations of a sensitive body region where the appearance is changed more due to the large scale intra-class variations.

Spatial representation has a good technical sound for analysing a video sequence under the variations of different poses or actions of a targeting object [13]. Here we have used the rectangular window based video analysis technique [15] which is used in the field of Qualitative Spatial and Temporal Representation (QSTR). Then we have used another most usable space time shapes (STS) silhouettes based representation technique [14] which is used in Human action recognition. To use the useful properties of these techniques, first we have considered a fixed size rectangular window (W). Suppose the window size is (DxD). We adopt the non-overlapping fixed size square shape window. Each window W has a fixed co-ordinate position. So each silhouette at t time frame in a gait cycle of a sequence is now composed of a set of foreground and background windows. The co-ordinate positions of foreground and background windows are fixed and we can easily calculate it by using an auxiliary training set. Then a sequence of information by mapping a particular foreground window which provides a volume of consistent qualitative explanation of the walking nature during a gait cycle, or detect if the sequence of maps is inconsistent. In this case, the qualitative of walking posture of gait for each window W has three descriptors space (s), time (t), and viewing angle. In this process, we additionally memorize only first two descriptors because the viewing angle is fixed. Suppose a gallery sequence window W_g and a probe sequence window W_p consist the shape variations Vg *and* Vw respectively in t time silhouettes and space s is the window size (DxD). Actually the shape variations of a window indicate the volume of intensity values of this particular window area. Then we calculate the intersection between two variations in t time silhouettes which is used as an inference rules:

$$V(t,s)= \{R \mid R \in W : V_g \cap V_p\} \qquad (4)$$

These variations of two descriptors time t and space s are thus facilitating in reasoning inference rules to generate the final cognitive maps of each window. The descriptors of each window will be used in order to extract and to track the behaviour of a subject as well as removing the affected windows which are changed significantly in spatial relations by the different covariate conditions. Finally, we have integrated all the information of each window which is consistency to determine a walking posture into a concrete representation as an input gait features for individual classification.

4.1 Classification

Let a probe sequence P with m subsequences P_r $\{r=1,..,m\}$ and a gallery sequence G with n subsequences G_s $\{s=1,...,n\}$.

The matching measure for the subsequences is simply chosen as the Euclidean distance between P_r and G_s (let, $d^{sub}(P_r, G_s)$).

Then we compute the median of the minimum distances for each of the probe subsequence P_r as the distance between a probe P and a gallery G sequence defined as:

$$D_i(P,G) = \underset{r=1}{\overset{n}{med}}\, d_i^{sub}(P_r, G) \qquad (5)$$

the minimum distance is computed for each of the probe subsequence P_r to a gallery sequence G which is defined as:

$$d^{sub}(P_r, G) = \min_{s=1}^{m} d^{sub}(P_r, G_s) \qquad (6)$$

The median of the minimum distances is then derived for each of the probe subsequence P_r as the distance between a probe P and a gallery G sequence defined as:

$$D(P,G) = \underset{r=1}{\overset{n}{med}}\, d^{sub}(P_r, G) \qquad (7)$$

The final distance $D(P,G)$ is used for individual classification.

5. Results and Discussions

All the experiments are conducted on the benchmarking CASIA database [16]. It consists of three variations namely view angle, clothing and carrying condition changes, and comprises of 124 subjects. Each subject has 10 walking sequences consisting of six normal walking sequences where the subject does not carry a bag or wears a bulky coat, two carrying-bag sequences and two wearing-coat sequences. First four sequences of six normal sequences were used as the gallery set. The probe set included the rest of the normal sequences. We have compared the proposed method with three most usable gait recognition algorithms [~~11~~16,~~7~~8,~~8~~5]. In [16] they have used the average silhouettes over a complete gait cycle namely called Gait Energy Image GEI (i.e., direct template matching) given in Fig 3. without any statistical data reduction techniques. In [8] they have used GEI as input gait feature and first the dimensions are reduced by PCA then

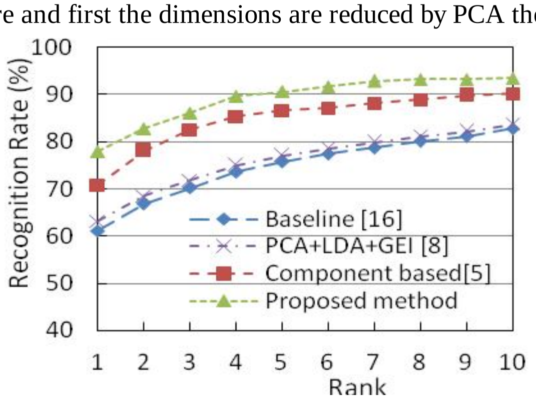


Figure 4. Performance evaluation

LDA is applied in lower dimensional PCA space. In [5], they divide the human body into seven components and the combinations of these components are used for classification.

The performances of all the methods are demonstrated in Fig. 5 using Cumulative Matching Curve (CMC) up to rank 5. It is observed from the Fig. 5, the proposed method shows much better result in a wide rank wise evaluation. The gait features are comprised of large scale intra-class variations due to different cofactors which are wrongly learned in training session in the previous works. The proposed method discards the most redundant windows effectively in early stage so that it can overcome the over fitting errors smartly. It is also effective to solve the under sampling problem. It is well known problem that the dimensions of gait features are very high than the training samples. Our proposed method minimizes the huge dimensions of gait features by using only the consistent windows and discarded the background windows in early stage. As it minimizes the huge dimensional gait features so the response time is also increased smoothly. Therefore the proposed novel method can be use as an to setup a faster recognition engine.

Actually the proposed system which generates a sequence of information by mapping a particular window region and provides a consistent qualitative explanation of the walking nature during a gait cycle, or detect if the sequence of maps is inconsistent. It extracts and tracks the targeting objects as well as removing the affected windows which are changed significantly in spatial relations by the different covariate conditions.

6. Conclusions

The main novelty of this proposed method is that the qualitative spatial and temporal reasoning based shape-invariant gait features. The idea is completely new and did not use in this major research areas. As gait features are comprised of large-intra class variations due to different covariate conditions and when it is wrongly learned in training session then the performance goes to down dramatically. The proposed method holds the local spatio-temporal qualitative consistency windows only and discards all the inconsistency windows which are affected by the different covariate conditions. Therefore, each window carries a volume of descriptors (i.e., space, time and shape) which are the most important discrimination power for individual classification. As a result, we have got much better performances than the existing quantitative based gait representation techniques.

ACKNOWLEDGEMENT

M. S. Azam thanks to his great contribution to draw out the results and some effective ideas to complete this work. And also the Software Lab, Pabna University of Science and Technology, Pabna, Bangladesh gives us enormous facilities for all our experimental works.

REFERENCES

1. K. Bashir, T. Xiang and S Gong, **Gait recognition without subject cooperation**, Pattern Recognition Letters 31 (2010) 2052-2060.
2. Pathokinesiology Service & Physical Therapy Department: **Observational gait analysis handbook**. Downey (CA): Professional staff association of Rancho Los Amigos Medical Center, ISBN 0-9676335-1-6 (1989), pp. 1-55.
3. L. Lee, W. Grimson, **Gait analysis for recognition and classification,** in proceedings of the Fifth IEEE Conference on Face and Gesture Recognition, vol. 1, 2002, pp. 155–161.
4. D. Hatzinakos and K. Plataniotis, Gait Recognition, Bell Canada Chair in Multimedia. **IPSI: Identity, Privacy and Security initiative**,The Edward S. Rogers Sr. Dept. of Electr.&Comput. Eng, UofT, M5S 3G4, Canada.
5. X. Li, S. Maybank, S. Yan, D. Tao, D. Xu, **Gait components and their application to gender recognition**, IEEE Transactions on Systems, Man, and Cybernetics Part C 38 (2) (2008), pp. 145–155.
6. D.K. Wagg, and M.S Nixon, , An **automated model-based extraction and analysis of gait**, Proc. of the 6th IEEE Int. Conf. on Automatic Face and Gesture Recognition, 2004, pp.11–16.
7. G. Ariyanto and M.S. Nixon, **Marionette mass-spring model for 3d gait biometrics**, In Proc. of the 5th IAPR Int. Conf. on Biometrics, 2012, pp. 354–359.
8. J. Han, B. Bhanu, **Individual recognition using gait energy image**, Transactionson Pattern Analysis and Machine Intelligence 28 (2) (2006) 316–322.
9. K. Bashir, T. Xiang, S. Gong, **Gait recognition using gait entropy image**, 3rd International Conference on Crime Detection and Prevention, 2009, pp.1-6.
10. R. Martin-Felez and Tao Xiang, **Gait Recognition by Ranking**, ECCV 2012, Springer-Verlag Berlin Heidelberg, pp. 328–341.(add our paper)
11. M. Rokanujjaman, M. S. Islam, M. A. Hossain, M. R. Islam, **Effective Part Definition for Gait Identification Using Gait Entropy Image,** Int. Conf. on Informatics, Electronics & Vision (ICIEV), 2013, ISBN-978-1-4799-0397-9.
12. M. Rokanujjaman, M. S. Islam, M. A. Hossain, M. R. Islam, Y. Makihara, Y. Yagi, **Effective part-based gait identification using frequency-domain gait entropy features**, Multimedia Tools and Applications, Springer Nov 2013, Vol 67, No 3.
13. A. G. Cohn, and J. Renz, **Qualitative spatial reasoning**. In van Harmelen, F.; Lifschitz, V.; and Porter, B., eds., Handbook of Knowledge Representation. Elsevier, 2007.
14. L. Gorelick, M. Blank, **Actions as Space-Time Shapes**, IEEE transactions on pattern analysis and machine intelligence, 2007, vol. 29, no. 12.
15. A. G. Cohn, J. Renz, and M. Sridhar, Thinking Inside the Box: **A Comprehensive Spatial Representation for Video Analysis**, 13th Int. Conf. on the Principles of Knowledge Representation and Reasoning, 2012.
16. S. Yu, D. Tan, and T. Tan. (2006) **A framework for evaluating the effect of view angle, clothing and carrying conditionon gait recognition**. In ICPR, pages 441–444.
17. Z. Liu, and S. Sarkar, **Simplest representation yet for gait recognition: averaged silhouette**, International Conference on Pattern Recognition (ICPR), vol.4, pp. 211- 214, 2004.