Skeleton Approach Based Gait Recognition for Human Identification

Debasish Pradhan [1] Prakash Chandra Jena [2] Bikash Sethi [3]

[1]Asst. Professor, Einstein Academy of Technology & Management, Bhubaneswar
[2]Asst.Professor, Einstein Academy of Technology & Management, Bhubaneswar
[3]Student, Einstein Academy of Technology & Management, Bhubaneswar

Abstract

In this paper, firstly the human skeleton is extracted from the each silhouette of an individual. Then, skeleton energy image (SkGEI) is obtained by averaging the sequence of skeletons. Further, HOG features (Histogram of Oriented Gradients) are obtained from the skeleton energy image. Finally, a simple KNN (K – Nearest Neighbor) classifier is used for the classification. The extensive experiments are conducted on considerably largest, publicly available standard datasets (i.e. CASIA B&C) and the promising results have shown that the proposed work is robust to different gait covariate factors.

Keywords: Energy image; Gait; Histogram of oriented gradients; Robust; Skeleton

1. Introduction

Human gait recognition recently gained a wider interest from machine vision research community because of its rich amount merits. A gait biometric source has many advantages compared to traditional biometric sources such as face, iris, finger print, speech, signature and hand geometry, etc. Gaits consist of both physiological characteristics and psychological state of humans. Hence, it is also referred as a behavioral biometric source. The following are the merits of gait recognition technology. [15]

- Both gait and face can be captured from at a larger distance. Hence, both works well with low resolution images. But face recognition failed in some certain practical scenario"s such as when the person projects his/her face in back view to the recording device/camera and when a person"s face covered by masks.
- User cooperation not require for gait recognition. But fingerprint, iris, signature, hand geometry and speech necessarily require the user cooperation.
- Gait is view independent biometric trait. Since, the person can walk in any direction in front of the recording device/camera. However, gait can be captured in all viewing angles.

The current state of art methods have been thoroughly studied and reviewed in this section. The related works are clearly explained in the below paragraphs.

Kale et al [1] choose the width of the outer contour of the silhouette as the feature vector. Heesung Lee et al [3] have proposed a method for eliminating the effect of a carried backpack for efficient gait recognition. They have used recursive principal component analysis (PCA) reconstructions and error compensation to remove the backpack from the gait representation.

Hayder Ali et al [4] have incorporated PCA with Radon Transform (RT) and without RT. The Radon Transform is used to detect features within an image and PCA is used to reduce dimension.

Han and Bhanu et al [5] have represented the popular and widely used gait representation method (i.e.GEI). GEI is refereed as a mean of aligned and resized silhouettes. GEI has already proven its strength in the gait recognition research field. Since, it is insensitive to practical conditions such as lighting conditions and poor illumination problem, etc.

Jyoti Bharti and Gupta [7] have proposed an approach which is based on graph of all pair shortest path distance. They have chosen 4 points (i.e. palm, knee, ankle and toe) from each silhouette and generated the graph by connecting the frames using these four points. These four points were used as the node of the graph.

ISSN: 2278-4632
Vol-11 Issue-02 2021

Further, they computed the Euclidean distance between the points and consider the Euclidean distance as a weight between two nodes. Finally, all pair shortest path distances were calculated by using these weights. The Classification procedure was achieved by matching the shortest path distance of input images with the database.

Jinyan Chen and Jiansheng Liu [8] have explored the average gait differential image (AGDI) as gait representation method. The AGDI is generated by the accumulation of the silhouettes difference between adjacent frames. The Two-dimensional principal component analysis (2DPCA) was used to extract features from the AGDI.

Jing Luo et al [9] have explored a fusion of the moment invariants which were extracted from GEI and AFDEI. Then, gait recognition was accomplished using the nearest neighbor classifier based on the Euclidean distance.

Lili Liu et al [10] have proposed the outermost contour based approach. A combination of Principal Component Analysis (PCA) and Multiple Discriminant Analysis (MDA) was adopted to reduce the dimensionality and to optimize the separability. The three classifiers (Nearest Neighbor, Back propagation Neural Network and Support Vector Machine) were used for the classification procedure in their work.

Lishani et al [11] have extracted the Haralick‟s features from three regions of GEI and SVM is incorporated as the classifier in their work.

Sungjun et al [13] have proposed the mass vector based gait recognition approach. Dynamic time-warping (DTW) approach was used for the matching.Sungjun Hong et al [14] have proposed the width vector mean based gait recognition approach.

Xiaoxiang LI and Youbin CHEN [16] have proposed a new gait recognition method called structural gait energy image (SGEI). SGEI is generated by a fusion of foot energy image (FEI) and head energy image (HEI).

Yumi Iwashita et al [17] have partition the human body image into multiple areas, and features for each area are extracted. Next, a matching weight for each area is estimated based on the similarity between the extracted features and those in the database for standard clothes. Finally, the subject is identified by weighted integration of similarities in all areas.

2. Proposed Methodology

The proposed work is divided into three major sections. The section 2.2 clearly shows the skeleton representation and section 2.3 describes the feature extraction procedure. Finally, section 2.3 shows the simple classification procedure.

Preprocessing

Morphological operators are used in order to remove the isolated noise, to fill the small gaps and to fill the hole in the silhouette. Then, bounding rectangle is used to extract the silhouette region. The same procedure is applied to all silhouettes in a sequence.

Further, we perform (1) size normalization to ensure constant height silhouettes and (2) horizontal alignment to centre silhouettes with the centroid of top 10 %figure height as a reference. [19]

Skeleton Energy Image Representation (SkGEI)

Skeleton is a lower dimensional shape description of an object. The requirements of a skeleton differ with applications. For example, object recognition requires skeletons with primitive shape features to make similarity comparison. In shape analysis, skeleton (or topological skeleton) of a shape is a thin version of that shape that is equidistant to its boundaries.

In this paper, the skeletonization of human silhouette is generated by the classical medial axis method [18]. The medial axis of an object is closely connected to the distance function from the boundary of the object (the medial axis can be defined as the set of singularities of the distance function) [19].

After the computation of the skeleton, the pruning method was used to remove the spurious branches in order to get robust skeleton. Over a complete gait sequence of an individual, skeleton motion can be extracted by considering average pixel intensity values.

The below equation (1) and (2) shows the gait representation computation. In which, the symbol „Sk‟ represents the sequence of extracted skeletons and „SkGEI‟ is the average of the skeletons.

The SkGEI, which places emphasis on body motion as opposed to covariate factor motion compared to GEI (Gait Energy Image) representation. In GEI, a rucksack (backpack) undergoes motion due to natural gait motion. But, the skeleton represents the rucksack as a mere bend in the skeleton compared to a mass of static and dynamic pixel values for silhouette representations (i.e. GEI). [19]

The figure 1 shows the representation of the skeleton energy image.

$$Sk = \{S_1, S_2, ..., S_N\} \tag{1}$$

$$SkGEI = \frac{1}{N}\sum_{i=1}^{N} S_i \tag{2}$$

ISSN: 2278-4632
Vol-11 Issue-02 2021

Figure 1. Skeleton Energy Image (SkGEI)

Histogram of Oriented Gradients (HOG)

This paper has incorporated the HOG [20] as the powerful feature extraction method. The histogram of oriented gradients is a feature descriptor used in computer vision and image processing for the purpose of object detection/recognition. The merits and properties of HOG are as follows.

- This technique counts occurrences of gradient orientation in localized portions of an image.
- The essential thought behind the histogram of oriented gradients descriptor is that local object appearance and shape within an image can be described by the distribution of intensity gradients or edge directions.
- The HOG descriptor has a few key advantages over other descriptors since it operates on local cells; it is invariant to geometric and photometric transformations.

The detailed algorithmic procedure of HOG is depicted in the below Algorithm 1. [20]

Algorithm 1: HOG (SkGEI)
//Input: Skeleton Energy Image (SkGEI)
//Output: HOG feature vector

Step 1: Compute the Gradients in X and Y direction. G_X is the gradient in X direction and G_Y is the gradient in Y direction

$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

Step 2: Compute the magnitude and angle of the gradients respectively

$$\nabla f = \sqrt{G_x^{\,2} + G_y^{\,2}}$$

$$\alpha(x, y) = \tan^{-1}\left(\frac{G_y}{G_x}\right)$$

Step 3: Partition the image into the number of cells. The size of each cell is 8X8. Next, the four neighboring cells are put together to form a block of size 16X16.

Step 4: Obtain the 9-bin histogram for each cell. Further, concatenate the four histograms to form one block feature vector. The length of the block feature vector is 36 i.e. 9 bin x 4 cells. Then, normalize the values in the block using L1-Norm.

Step 5: The HOG descriptor is the concatenated vector of the components of the normalized cell histograms from all of the block regions. These blocks typically overlap, meaning that each cell contributes more than once to the final descriptor. Total feature length is 5400 (i.e. 36 x 10 horizontal positions X 15 vertical positions). Lastly, normalize the resultant feature vector using L2-Norm.

2.4. Classifier

At last, we use a simple classifier (i.e. the K-nearest neighbor classifier) to identify the subjects with the help of SkGEI+HOG. To measure the dissimilarity between the probe and gallery samples, city block distance measure is used .Further, the value for „K‟ is chosen based on the experimental trials in order to consider „K‟ nearest neighbors for the probe sample.

3. Experimental Results

This paper has utilized two considerably largest, publicly freely available standard datasets [2] [12]. Among the two, CASIA B [12] consists of multi-view, carrying and wearing condition sequences. CASIA C [2] consist of backpack and different walking speed variation sequences.

CASIA B Dataset

This dataset consists of 124 person‟s gait sequences. In this, each person walked in 11 view directions in front of the recording device i.e. 0^0, 18^0, 36^0, 54^0, 72^0, 90^0, 108^0, 126^0, 144^0, 162^0 and 180^0. Each view has 10 sequences i.e. 2 carrying bag (C_B), 2 wearing cloth (W_C) and 6 normal walk sequences (N_W). However, the dataset consist of total 13,640 gait sequences (i.e. 11 views X 10 sequences per view X 124 people).

As indicated in the above, CASIA B dataset consist of three gait covariates conditions (i.e. carrying, wearing

ISSN: 2278-4632
Vol-11 Issue-02 2021

coat and normal walk). However, the whole dataset is split into three sets (i.e. K = 3) by considering each covariate as a separate class. In each turn, one of the „K" subsets is used for testing and remaining the K-1 subsets are used for training. The same is repeated for three times. The below Table 1 shows the CCR (Correct Classification Rate) for the prominent covariate conditions.

Table 1. Results on CASIA B Dataset

Covariates	CCR	
	K = 1	K = 5
Carrying Bag	93.47%	98.97%
Cloth	95.30%	99.04%
Normal Walk	95.44%	99.04%

3.2 CASIA C Dataset

This dataset consists of 153 person"s gait sequences. This dataset has total 1,530 lateral view sequences with backpack and different walking speed conditions (i.e. slow, medium and fast walk). Each subject consists of 10 sequences i.e. 2 backpack (B_P), 4 normal walk (N_W), 2 slow walk (S_W) and 2 fast walk (F_W).

As indicated in the above, CASIA C dataset consist of four gait covariates conditions (i.e. backpack, slow, fast and normal walk). However, the whole dataset is split into four sets (i.e. K = 4) by considering each covariate as a separate class. In each turn, one class is used for testing and remaining the three classes are used for training. The same is repeated for four times. The below Table 2 shows the CCR rates for the prominent gait covariate conditions.

Table 2. Results on CASIA C Dataset

Covariates	CCR	
	K = 1	K = 3
Backpack	94.77%	98.03%
Slow walk	96.40%	100%
Fast Walk	97.38%	100%
Normal Walk	98.36%	100%

4. Comparative Analysis

Xiaoxiang LI and Youbin CHEN et al [16] have conducted the experiments on CASIA B dataset. They have achieved the 89.29% of average CCR for 90^0 view angle sequences. In their work, three normal walking sequences per subject were used as the gallery set and the seven sequences per subject were used as the probe set (i.e. two carrying bag + two wearing coat + three normal walking sequences). Singh & S. Jain, et al [6] reported the 79.01% of average CCR for the same experimental setting which is used in the above literature. For the same mentioned experimental setting, the proposed work achieved the 90.72% of average CCR.

Jinyan Chen and Jiansheng Liu et al [8] have used CASIA B dataset. They have reported 61.09% of average CCR for eleven viewing angle sequences. They have used six normal walking sequences per viewing angle. For the same mentioned experimental settings, the proposed work achieved the 96.04% of average CCR.

Heesung Lee et al [3] and Yumi Iwashita et al [17] have included the CASIA C dataset for the experiments and have reported 82.68% and 94% of average CCR respectively. Heesung Lee et al [3]have used four normal walking sequences as the gallery set and the six sequences as the probe set (i.e. 2 slow walk + 2 backpack + 2 fast walk). For this setup, our proposed work achieved 93.68% of CCR.

Yumi Iwashita et al [17] have used CASIA C dataset. They have incorporated 4-fold cross validation procedure and have utilized the four sequences per subject for the experiments. In each turn of testing procedure, 153 sequences were used as the probe set and 153 X 3 number of sequences were used as the gallery set. For the same mentioned experimental settings, the proposed work achieved 98.03% of average CCR.

The direct silhouettes based literatures [1][7][10][14] were extracted the features from the sequence of binarized silhouettes. These approaches may suffer from the practical challenges such as noise, poor illumination, different lighting conditions and lack of frames etc.

Lili Liu et al [10] have used CASIA (B) dataset and have reported 97.67% CCR for 90^0 viewing angle sequences. For each subject, six side views with normal walking condition sequences were selected. They have incorporated hold out type procedure by considering three sequences as the training set and the remaining three sequences as the testing set. For the same, the proposed work achieved 98.66% of CCR.

5. Conclusion

This paper has demonstrated the skeleton approach based gait recognition method in order to achieve the high recognition rate under different gait covariate factors. A combination of skeleton energy image (SkGEI) and histogram of oriented gradients (HOG) has given the promising results on considerably large, publicly available datasets. The comparative results have shown that the proposed method is robust compared to the current state of art methods.

6. Acknowledgements

The authors would like to thank the creators of CASIA Dataset (B and C) for providing the publicly available gait datasets.

REFERENCES

[1] Kale., N. Cuntoor., B. Yegnanarayana., AN. Rajagopalan., & R. Chellappa, (2003). Gait analysis for human identification. In AVBPA. pp 706-14.
[2] Daoliang Tan, Kaiqi Huang, Shiqi Yu, & Tieniu Tan, (2006). Efficient Night Gait Recognition Based on Template Matching. In Proc. of the 18′th International Conference on Pattern Recognition (ICPR06). Hong Kong.
[3] Heesung Lee., Sungjun Hong., & Euntai Kim, (2009). An Efficient Gait Recognition with Backpack Removal. EURASIP Journal on Advances in Signal Processing. Volume 2009. Article ID 384384. 7 pages. doi:10.1155/2009/384384. Hindawi Publishing Corporation.
[4] Hayder Ali., Jamal Dargham.,Chekima Ali., & Ervin GobinMoung, (2011). Gait Recognition using Gait Energy Image. International Journal of Signal Processing, Image Processing and Pattern Recognition. Vol. 4. No. 3.
[5] Ju Han., &BirBhanu, (2006). Individual Recognition using Gait Energy Image. IEEE Transactions on Pattern analysis and Machine Intelligence. Vol 28. No 2.
[6] J. Singh., & S. Jain, (2010). Person identification based on gait using dynamic body parameters. In Trends in Information Sciences Computing (TISC). pp. 248 – 252.
[7] Jyoti Bharti., & M.K Gupta, (2011). Human Gait Recognition using All Pair Shortest Path. International Conference on Software and Computer Applications. IPCSIT. Vol 9. © IACSIT Press. Singapore.
[8] Jinyan Chen., &Jiansheng Liu, (2014). Average Gait Differential Image Based Human Recognition., Scientific World Journal. Volume 2014. Article ID 262398. 8 pages. Hindawi Publishing Corporation.
[9] Jing Luo.,Jianliang Zhang.,Chunyuan.,Ying Niu., HuixinTian, and ChunboXiu (2015).Gait Recognition Using GEI and AFDEI. International Journal of Optics. Volume 2015. Article ID 763908. 5 pages. Hindawi Publishing Corporation.
[10] Lili Liu., Yilong Yin., Wei Qin., & Ying Li, (2011). Gait Recognition Based on Outermost Contour. International Journal of Computational Intelligence Systems. pp 1090-1099.
[11] Lishani, Ait, Boubchir, Larbi, and Bouridane, Ahmed, (2014). Haralick features for GEI- based human gait recognition. In 26thInternational Conference on Microelectronics (ICM). Piscataway. IEEE.
[12] Shiqi Yu., Daoliang Tan., & Tieniu Tan, (2006). 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). Hong Kong.
[13] Sungjun Hong., Heesung Lee., Imran FareedNizami., & Euntai Kim, (2007).A New Gait Representation for Human Identification: Mass Vector. 1-4244-0737-0/07/$20.00. © IEEE.
[14] Sungjun Hong., Heesung Lee., & Euntai Kim, (2009).Automatic Gait Recognition using Width Vector Mean. ICIEA. 978-1-4244-2800-7/09/$25.00 © IEEE.
[15] V. G. Manjunatha Guru., & V. N. Kamalesh, (2015). Vision based Human Gait Recognition System: Observations, Pragmatic Conditions and Datasets. Indian Journal of Science and Technology. ISSN (Print): 0974-6846. ISSN (Online): 0974-5645. Vol. 8(15).
[16] Xiaoxiang LI., & Youbin CHEN, (2013). Gait Recognition Based on Structural Gait Energy Image. Journal of Computational Information Systems. pp 121–126. Available at http://www.Jofcis.com.
[17] Yumi Iwashita., Koji Uchino., & Ryo Kurazume, (2013). Gait-Based Person Identification Robust to Changes in Appearance. Sensors. doi:10.3390/s130607884. pp 7884-7901.
[18] Blum, H.: Transformation for extracting new descriptors of shape.In: Wathen-Dunn, W. (ed.) Models for the Perception of Speech and Visual Form. MIT Press, Cambridge (1967)
[19] Tenika Whytock · Alexander Belyaev · Neil M. Robertson,(2014). Dynamic Distance-Based Shape Features for Gait Recognition, J Math Imaging Vis, DOI 10.1007/s10851-014-0501-8
[20] Navneet Dalal and Bill Trigs, “Histograms of oriented Gradients for Human Detection”,CVPR,2005.