Using a Generic Model for Codebook-based Gait Recognition Algorithms

Muhammad Hassan Khan[†], Muhammad Shahid Farid[‡], Marcin Grzegorzek[†,*]
[†]Research Group of Pattern Recognition, University of Siegen, Germany
[‡]College of Information Technology, University of the Punjab, Lahore, Pakistan
[*]Faculty of Informatics and Communication, University of Economics in Katowice, Poland
Email: hassan.khan@uni-siegen.de, shahid@pucit.edu.pk, marcin.grzegorzek@uni-siegen.de

***Abstract*—Gait has emerged as a distinguishable human biological trait. It refers to the walking style of an individual and is considered an important biometric feature for person identification. Codebook based gait recognition algorithms have demonstrated excellent performance by achieving high recognition rates. However, such methods construct a codebook for each database or scenario. In this paper, we investigate the idea of using a generic codebook for gait recognition. The proposed codebook is built by using spatiotemporal characteristics of gait and is based on a large diverse synthetic gait database. We also propose a gait recognition algorithm based on this generic codebook. The advantages of the proposed algorithm over the existing methods include its independency from generating a codebook for each database, rather the proposed generic codebook can be used to encode any gait scenario. Moreover, the proposed algorithm is model free and does not require human body segmentation or modeling. The performance of the proposed generic codebook-based gait recognition algorithm is evaluated on two large gait databases TUM GAID and CMU MoBo, and recognition rate reveals the effectiveness of the proposed algorithm.**

***Index Terms*—Gait recognition, generic codebook, spatiotemporal features, Fisher vector encoding**

I. Introduction

In the recent years, biometrics have attracted significant research efforts due to its various applications in access control, human surveillance and authentication. Different distinguishing biological traits have been proven to be effective means for person identification [1]. Biometrics refers to the use of biological and behavioral characteristics for individual identification. Biological characteristics, such as, iris, fingerprints, DNA, facial features, and earlobes structure have been proven to be unique for each individual. The behavioral characteristics include voice, gait and others. Gait refers to the unique manner or style of walking. Human gait is different from other biometric features as it can be captured from a distance and does not require human interaction with the imaging device. Moreover, the gait features can be computed from low-resolution videos. These characteristics make it an ideal choice for surveillance systems, particularly at security-sensitive and monitoring environments such as banks, military bases, etc.

The gait biometric indeed has advantages over the other physiological biometrics, particularly its acquisition from distance and without individual interaction. However, gait recognition is challenging in many aspects due to the variations that may occur in clothing, footwear, walking surface, walking speed, injuries and others. Other biometric modalities to identify the individuals, such as iris and fingerprints, might be more powerful than gait, nevertheless its ability to recognize human from distance and without any interaction with the system makes it irreplaceable in many applications such as visual surveillance.

Gait recognition has received significant research in the recent years and various approaches have been proposed which can be classified into two broad categories: model-based approaches and model-free approaches. Model-based techniques build a gait signature using the human body structure and motion models. Several human body parts and joint positions are tracked over time and used to identify the walkers [2]–[5]. In [6], the authors modeled the human silhouette structure using seven different ellipses representing the various human body regions. Several statistical measurements on these regions are computed for gait and gender classification. The algorithm proposed in [7] locates the joint locations and computes the joint angle trajectories at these locations to form a gait signature.

The gait recognition algorithm proposed in [8] builds a 3D voxel model using the ellipsoids fitting technique into four different lower limbs components and the derived features are represented using a Fourier based representation. Bouchrika et al. [9] exploits the elliptic Fourier descriptors to extract features from human joints which are used to form a model for person identification. The gait feature proposed in [10] is composed of the angular motion of the hip and thigh. The authors in [11] split the human's body region into three parts and the variance of these parts over time are combined to obtain a gait feature.

Recent studies [9], [12], [13] have shown that the model-based gait recognition algorithms are effective and can deal with occlusion to some extent. However, the performance of these algorithms largely depend on the localization of torso, which cannot be easily extracted from the underlying model in gait sequences. Moreover, the model-based approaches are computationally expensive and are sensitive to the video quality [12].

The model-free approaches do not use a structural model of human motion, instead they operate on the sequence of extracted binary human silhouettes. Perhaps, the most simple and effective technique is gait energy image (GEI) [14], [15], which has been extensively used in model-free gait recognition algorithms. A GEI is computed by extracting the human body silhouette using background modeling techniques and averaged them over time in a gait cycle. The algorithm proposed in [16] extracts the human skeleton information from the silhouette images and combines it with the motion information to obtain an effective gait feature. The algorithm described in [17] uses several statistical measurements computed from the human silhouette in a gait cycle. The method proposed in [18] exploits the height and width from normalized and scaled silhouette of human body, over a gait cycle to obtain a gait signature. Tan et al. [19] developed a normalized pseudo-height and width histogram using silhouette images to recognize the individual's walk. The algorithm proposed in [20] exploits the human motion characteristics to drive a gait feature.

In [22], the edges and depth gradient of the person's silhouette are extracted from depth images and are used to recognize the individual's gait. The authors in [23] obtained the principal components coefficients from the silhouette images and wavelet descriptors respectively, and are analyzed using Independent Component Analysis (ICA) to get the more independent gait features. Shape analysis of human's silhouette is also exploited in many gait recognition techniques. Wang et al. [24] used the Procrustes Shape Analysis (PSA) to obtain Procrustes Mean Shape (PMS) from a sequence of silhouettes as gait signature. PMS represents the both motion and body shape into a unified descriptor and similarity is measured using Procrustes Distance.

The silhouette based model-free approaches are computationally efficient compared to the model-based technqiues and they have demonstrated high recognition results on various public benchmark gait databases. However, their performance is somewhat sensitive to the variation in the silhouette shapes and thus depends on the precise silhouette segmentation. An inaccurate segmentation may lower the recognition accuracy [17]. In this paper, we present a model-free gait recognition algorithm based on a novel generic codebook. The proposed algorithm does not involve human body segmentation or gait cycle estimation.

There are two major contributions of this paper. First, a novel proposal of using a generic codebook to encode the motion descriptors of the gait sequences is presented. In existing codebook-based gait recognition algorithms, a codebook is generated for each dataset and used in feature encoding. The idea of generic codebook is to use a single codebook for gait recognition. We show that it can achieve a number of advantages over the conventional database specific codebooks, including: (1) Usually a codebook is generated for each dataset to encode the available gait scenarios and it can be used to recognize the individuals in that particular environment only. Whereas the generic codebook can be effectively used to build a gait signature for any type of walk sequence. (2) A generic codebook generated from a large and diverse set of walking styles can serve as a universal codebook and can be used to encode any kind of individual's walk. (3) Experimental evaluation on two benchmark gait datasets demonstrate that the generic codebook approach is efficient.

Second, a model-free, generic codebook-based gait recognition algorithm is proposed. The proposed algorithm is based on our previous work [25] in which a spatiotemporal gait representation is presented. In contrast to most existing gait recognition algorithms that require the extraction of human body silhouette, contour or other skeletal information; the proposed approach is model-free and it does not involve any kind of human body segmentation or the gait cycle computation. Therefore, the performance of the proposed algorithm is particularly better than other competing approaches under low resolution and low quality video sequences.

An extensive experimental evaluation on two large benchmark gait databases, TUM GAID and CMU MoBo, reveals the effectiveness of the proposed algorithm. The results are compared with the state-of-the-art gait recognition techniques.

The rest of the paper is organized as follows: Sect. II describes the proposed generic codebook for gait recognition. Experimental evaluation and results are presented in Sect. III and the conclusions are drawn in Sect. IV.

II. Proposed Generic Codebook for Gait Recognition

Usually the codebooks are application specific and are generated using a subset of the target dataset. The idea of 'generic codebook' advocates the concept of using a single codebook built on one dataset, and encode and test the sequences of other datasets. However, the selection of a dataset to obtain the generic codebook is very important as it directly deals with feature encoding and can significantly effect the performance of algorithm. A dataset with a large variety of walking styles is momentous for good performance of gait recognition algorithm. A codebook obtained from a dataset with limited walking styles would loose important gait cues in encoding, resulting in poor performance.

To this end, we investigated various existing gait datasets and found that a subset of CMU Motion Capture (mocap) dataset [26] is the most appropriate dataset for generic codebook generation. The synthetic video sequences of walk from mocap dataset are used to generate the generic codebook. These sequences covers a large range of walk types e.g. normal, slow, fast, exaggerated stride, brisk walk, wander, etc. A total of 80 video sequences are used to generate the generic codebook. Fig. 1 shows few sample images of a normal walk sequence from the selected dataset. To build the generic codebook, the spatiotemporal motion information of each mocap sequence is used.

A. Feature selection

Several features have been investigated and proposed to efficiently recognize individuals' gait. Lately, the dense trajectories have shown good performance in action recognition [21], [27]. Wang et al. [27] proposed the computation of Histogram

(a) 65th frame (b) 70th frame (c) 75th frame (d) 80th frame

Fig. 1. Sample images from a synthetic gait sequence of mocap dataset used to generate the generic codebook.

of Oriented Gradient (HOG) and Histogram of Optical Flow (HOF) features along the dense trajectories. Moreover, to encode the relative motion information between the pixels along the horizontal and vertical axis, the derivatives along the respective component of optical flow are computed, and their orientation information is quantized into histograms. These features are known as Motion Boundary Histogram (MBH), and represented as MBH_x and MBH_y, respectively. In [25], the performance of several motion descriptors including, HOG, HOF, MBH, and their various combinations is evaluated for gait recognition. The results of this investigation showed that the combination of HOG and MBH achieves the best results. Therefore, we chose HOG and MBH (i.e., MBH_x and MBH_y) descriptors to construct the generic codebook. The superior performance for the combination of HOG and MBH descriptors over the rest is because the HOG captures the static appearance of the person and the MBH incorporates the changes in optical flow field. Therefore, they collectively perform better in identifying a person using his/her appearance and local motion characteristics.

B. Feature extraction

To extract trajectories from a video sequence, a sample of dense points is chosen from each frame and tracked in successive frames using the displacement information from a dense optical flow field. Specifically, each point P_i in frame i is tracked in frame $i+1$ using median filtering in a dense optical flow field. The set of tracked points in successive frames are concatenated to construct a trajectory (i.e., $P_i, P_{i+1}, P_{i+2}, \ldots$). Let us assume that L is representing the length of a given trajectory, S is the sequence of displacement vector $\Delta P_i = (P_{i+1} - P_i) = (x_{i+1} - x_i, y_{i+1} - y_i)$ and can be formed as follows,

$$S = (\Delta P_i, ..., \Delta P_{i+L-1}), \tag{1}$$

The sequence vector S is then normalized by the sum of the magnitudes of the displacement vector ΔP. That is,

$$\mathcal{D} = \frac{(\Delta P_i, ..., \Delta P_{i+L-1})}{\sum_{j=i}^{i+L-1} ||\Delta P_j||} \tag{2}$$

The descriptor $\mathcal{D}$ define the shape of trajectory (i.e., the local motion pattern).

C. Generic codebook generation

The local descriptors are used to build a signature (i.e., feature encoding), to characterize an image or video sequence. Feature encoding is a process to transform the local descriptors into a fixed length vector, usually using the vector quantization and building a histogram. In the frame of this study, our local descriptors (i.e., HOG+MBH) are encoded using Fisher vector (FV) and a generic codebook based on GMM. FV is derived from Fisher kernel [28] and comprising the description of local descriptors by its deviation from the generative model (i.e., GMM). The deviation is computed using the gradient of the descriptor log-likelihood with respect to the model parameters.

We randomly selected one million features from each local descriptor of CMU mocap sequences, to build a generic codebook. The GMM define the distribution over feature space and can be expressed as [29]:

$$p(X;\theta) = \textstyle\sum_{i=1}^{K} w_i \mathcal{N}(X; \mu_i, \sum_i) \tag{3}$$

where X is representing the local descriptors, $i = 1, 2, ..., K$ is the mixture component (i.e., cluster number), w_i is the weight, μ_i is the mean vector and $\sum_i$ is the covariance matrix of ith component. Moreover, $\mathcal{N}(X; \mu_i, \sum_i)$ is describing the D-dimensional Gaussian distribution and $\theta = \{w_i, \mu_i, \sum_i, i = 1, 2, ..., K\}$ is the set of model parameters which can be estimated using the expectation-maximization (EM) algorithm.

In a given set of descriptors $X = \{x_1, ..., x_t\}$, the optimal parameters of GMM are learned using maximum likelihood estimation. We used an iterative EM algorithm [30] to solve this problem. The soft assignment of descriptor x_t to cluster i, also known as posterior probability is defined as,

$$q_t(i) = \frac{w_i \mathcal{N}(x_t; \mu_i, \sum_i)}{\sum_{j=1}^{K} w_j \mathcal{N}(x_t; \mu_j, \sum_j)} \tag{4}$$

We consider that each model describe a specific motion pattern shared by the local descriptors in the codebook. The EM algorithm of GMM performs soft assignments of local descriptor to each mixture component therefore, the local descriptors are assigned to multiple components (i.e., clusters) in a weighted manner using the posterior component probability given by the descriptor. Unlike k-means clustering, the encoded features provides not only the mean information of codewords, but also the shape of their distribution. The generic codebook clustering size (K) is empirically chosen and is fixed to 256.

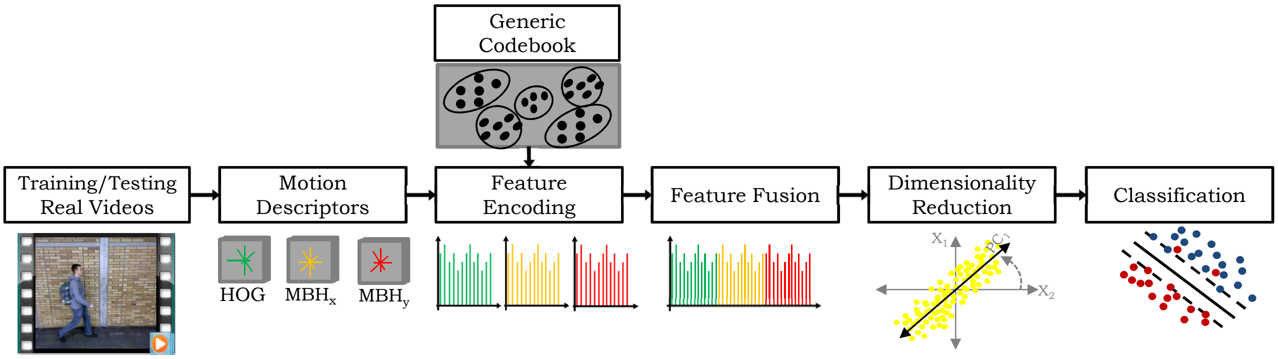

Fig. 2. Proposed gait recognition model using generic codebook.

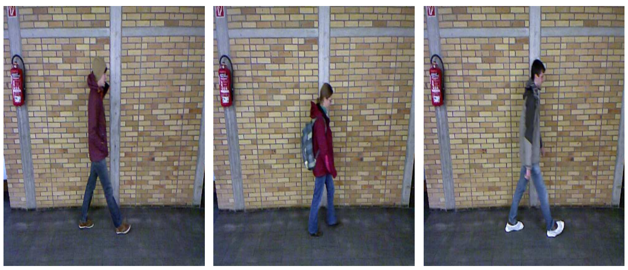

Fig. 3. Sample images from TUM GAID database. Normal walking (left), walking with backpack (middle), and walking with coating shoes (right).

TABLE I
COMPARISON OF RECOGNITION RESULTS (%) ON TUM GAID GAIT DATABASE. EACH COLUMN N, B, S, TN, TB AND TS CORRESPONDS TO A DIFFERENT EXPERIMENT AND AVERAGE IS COMPUTED AS SUM OF THE WEIGHTED MEAN SCORES. BEST RESULTS ARE MARKED IN BOLD.

Method	N	B	S	TN	TB	TS	Avg
GEI [22]	99.4	27.1	56.2	44.0	6.0	9.0	56.0
GEV [22]	94.2	13.9	87.7	41.0	0.0	31.0	61.4
SEIM [16]	99.0	18.4	96.1	15.6	3.1	28.1	66.0
GVI [16]	99.0	47.7	94.5	62.5	15.6	62.5	77.3
SVIM [16]	98.4	64.2	91.6	65.6	31.3	50.0	81.4
DGHEI [22]	99.0	40.3	96.1	50.0	0.0	44.0	87.3
CNN-SVM [32]	99.7	97.1	97.1	59.4	50.0	**62.5**	94.2
CNN-NN128 [32]	99.7	98.1	95.8	62.5	56.3	59.4	94.2
Proposed	**99.7**	**99.0**	**98.4**	**68.8**	**56.3**	53.1	**95.3**

D. Generic Codebook based Gait Recognition Algorithm

The idea of generic codebook can be used in the encoding of any gait descriptor. In this section, we extend our previous algorithm proposed in [25] and couple it with generic codebook. The algorithm exploits the spatiotemporal characteristics of human motion to recognize the individuals. It extracts the motion information of a walker from the gait video sequence and s them into local descriptors, as described in Sec. II-A. These descriptors are encoded using a generic codebook and Fisher vector encoding [28], and they are fused in a representation level fusion [29]. In all experiments, the generic codebook is used to encode the local descriptors of real gait sequences in the gallery and the probe sets. Finally, the computed features are classified using Linear Support Vector Machine (SVM). A block diagram of the proposed gait recognition algorithm is shown in Fig. 2.

III. EXPERIMENTS AND RESULTS

The performance of the proposed method is evaluated on two large benchmark gait databases: CMU MoBo [31] and TUM GAID [22]. The synthetic video sequences of walk from CMU Motion Capture (mocap) database [26] are used to build a generic codebook. The proposed method is implemented using Matlab R2016 software and experiments are carried out an Intel corei5 2.6GHz machine with 8GB RAM.

A. Performance on TUM GAID database

The TUM gait audio image and depth (GAID) database is one of the largest gait database, contains $3,370$ walk sequences of 305 subjects. The database was recorded at outdoor environment of Technical University of Munich, Germany in two different sessions. The first session of recording was held in January 2012, which is the winter season in the region and temperature was around -5°C, therefore, the subjects were wearing heavy jackets and winter boots. The gait sequences of 176 subjects were recorded in this session. The second session of recording was held in April 2012 and the temperature was around +15°C in the region thus, the subjects were wearing significantly different clothes and shoes. The gait sequences of 161 subjects were recorded in the second session. There is a subset of 32 subjects, who participated in both recording sessions therefore, the database contains the gait sequences of 305 subjects. The variation in clothing, shoes, lighting and other captured properties make this database extremely challenging in the field of gait recognition. Fig. 3 display few sample images from the TUM GAID database.

Each subject in the database has ten walk sequences from left-to-right and right-to-left, with three different variations namely: normal walking (N), walking with backpack (B) and walking with coating shoes (S). The subset of 32 subjects have ten additional gait sequences (i.e., in total 20) and can be represented such as: normal walking after time (TN), walking

TABLE II

COMPARISON OF RECOGNITION RESULTS (%) ON CMU MOBO GAIT DATABASE. EACH COLUMN CORRESPONDS TO A DIFFERENT EXPERIMENT AND AVERAGE (AVG) IS COMPUTED AS THE MEAN SCORE OF ALL THE EXPERIMENTS. BEST RESULTS ARE MARKED IN BOLD.

Methods	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Avg
SC [34]	100	100	92	92	80	48	48	84	28	48	68	48	28	32	44	0	60
ICA [23]	100	100	100	100	-	-	79.2	64	-	-	-	-	-	-	-	-	90.5
Shape kinematics [35]	100	100	92	92	80	48	48	84	28	48	68	48	12	32	44	0	59
STM-SPP [36]	100	100	100	-	94	-	93	91	-	84	82	82	-	-	-	-	91.8
3D ellipsoid [8]	100	100	100	-	78.6	-	70.5	-	-	-	-	61	-	-	-	-	85.1
WBP [37]	100	100	100	98.7	92	-	72.67	92	-	60.7	74.7	63.3	-	-	-	-	85.4
Uniprojective [38]	100	100	96	-	72	-	-	60	-	-	-	-	-	-	-	-	85.6
NDDP [39]	100	100	96	-	88	-	-	80	-	-	-	-	-	-	-	-	92.8
HSD [33]	100	100	**100**	-	92	-	-	-	-	-	88	84	-	-	-	-	94
SDL [17]	100	100	98.7	-	96	-	86.7	92	-	88	86.7	88	-	-	-	-	92.9
Proposed	**100**	**100**	96	**100**	**100**	**96**	**92**	**100**	**100**	**96**	**92**	**92**	**88**	**100**	**100**	**88**	**96.3**

(a) (b) (c) (d)

Fig. 4. Sample images from CMU MoBo gait database. (a) Slow walking, (b) fast walking, (c) slow walking with ball and (d) walking at certain slope (i.e., incline).

with backpack after time (TB) and walking with coating shoes after time (TS). For classification, the same division of database into gallery and probe set is used, defined in [22]. That is, the first four recordings of normal walk are assigned to gallery set and the remaining two sequence of normal walk, walk with backpack and walk with coating shoes are assigned to probe set separately, in three different experiments namely: N, B and S. For the rest of experiments, the gallery set was same but the gait sequences of normal walk after time, walk with backpack after time and walk with coating shoes after time are used in probe set separately, namely: TN, TB and TS, respectively. The performance comparison of the proposed method with state-of-the-art techniques is outlined in Tab. I. Our method outperformed the state-of-the-art in all experiments except in TS where the CNN-SVM [32] performs better than our method. The proposed method obtained the best average recognition accuracy 95.3%.

B. Performance on CMU MoBo gait database

The CMU motion of body (MoBo) database comprising the gait sequences of 25 subjects walking on a treadmill. Each subject in the database has four variations of walk namely: slow walking (S), fast walking (F), slow walking with a ball in hands (B), and slow walking at certain slope (i.e., incline (I)). The proposed method is evaluated on the sequences recorded in lateral view to demonstrate its robustness in terms of walking surface, walking speed and carrying condition. Few sample images representing the different walking scenarios in database, are shown in Fig. 4. We conduct two different types of experiments: (1) within the same condition, where the same walking scenarios are used in probe and gallery set, and (2) across the condition, where different walking scenarios are used in probe and gallery set. Tab. III summarized the description of sixteen experiments on this database. The performance comparison of proposed method with state-of-the-art techniques is described in Tab. II. The recognition results reveals that the proposed method achieved excellent results in all the experiments except C, where HSD [33] perform better than our method. It can also be noted from the results that the performance of the proposed method is consistently better than other competing algorithms on almost all experiments. In particular, in the difficult set of experiments E to P, the compared techniques performed rather poor and our method consistently achieved more than 90% recognition accuracy in most of the experiments. The proposed method obtained the highest average recognition rate 96.3%.

TABLE III

LIST OF SIXTEEN EXPERIMENTS ON CMU MOBO GAIT DATABASE

Exp.	Gallery set	Probe set	Type
A	Slow walk	Slow walk	Same condition
B	Fast walk	Fast walk	Same condition
C	Walk with ball	Walk with ball	Same condition
D	Incline walk	Incline walk	Same condition
E	Slow walk	Fast walk	Across condition
F	Slow walk	Incline walk	Across condition
G	Slow walk	Walk with ball	Across condition
H	Fast walk	Slow walk	Across condition
I	Fast walk	Incline walk	Across condition
J	Fast walk	Walk with ball	Across condition
K	Walk with ball	Slow walk	Across condition
L	Walk with ball	Fast walk	Across condition
M	Walk with ball	Incline walk	Across condition
N	Incline walk	Slow walk	Across condition
O	Incline walk	Fast walk	Across condition
P	Incline walk	Walk with ball	Across condition

The performance evaluation of proposed method, presented in Tab. I and II confirm its effectiveness on two benchmark gait databases which contain diverse variety of gait sequences.

The proposed algorithms showed very convincing results outperforming the state-of-the-art in most experiments. High recognition accuracy also confirm that the proposed generic codebook is capable to effectively encode gait sequences. In particular, the average recognition performance of the proposed algorithm is the highest on both of the gait databases.

IV. Conclusion

This paper proposed an idea of generic codebook for gait recognition. The contribution of the paper is twofold; first, a generic codebook is proposed to encode the gait signatures which presents a number of advantages over the conventional database specific codebooks. The proposed codebook is built using the motion descriptors, computed on a set of large, diverse synthetic gait sequences with variety of walking styles. Second, a gait recognition algorithm based upon the generic codebook is presented. It exploits the spatiotemporal motion characteristics of human's walk and encodes the motion descriptors using generic codebook. In contrast to existing gait based person identification techniques, our method neither requires the gait-cycle estimation nor the extraction of human body from the video sequences. The recognition results on two large benchmark gait databases confirm the effectiveness of generic codebook and its application with proposed method.

References

[1] F. Loula, S. Prasad, K. Harber, and M. Shiffrar, "Recognizing people from their movement.," *J. Exp. Psychol.-Hum. Percept.*, vol. 31, no. 1, pp. 210, 2005.

[2] M. Nixon et al., "Model-based gait recognition," 2009.

[3] A. Kleinsmith and N. Bianchi-Berthouze, "Affective Body Expression Perception and Recognition: A Survey," *IEEE Transactions on Affective Computing*, vol. 4, no. 1, pp. 15–33, 2013.

[4] A. Bedagkar-Gala and S. K. Shah, "A survey of approaches and trends in person re-identification," *Image and Vision Computing*, vol. 32, no. 4, pp. 270–286, 2014.

[5] M. H. Khan, K. Shirahama, M. S. Farid, and M. Grzegorzek, "Multiple Human Detection in Depth Images," in *Proc. Int. Workshop Multimed. Signal Process. (MMSP)*, pp. 1–6, 2016.

[6] L. Lee and W Eric L Grimson, "Gait analysis for recognition and classification," in *Proc. Int. Conf. Automatic Face and Gesture Recognit.* IEEE, 2002, pp. 155–162.

[7] L. Wang, H. Ning, T. Tan, and W. Hu, "Fusion of static and dynamic body biometrics for gait recognition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 2, pp. 149–158, 2004.

[8] S. Sivapalan et al., "3D ellipsoid fitting for multi-view gait recognition," in *IEEE Int. Conf. Adv. Video Signal Based Surveillance (AVSS)*. IEEE, 2011, pp. 355–360.

[9] I. Bouchrika and M.S. Nixon, "Model-based feature extraction for gait analysis and recognition," in *ICCV*. Springer, 2007, pp. 150–160.

[10] D. Cunado, M. Nixon, and J. Carter, "Automatic extraction and description of human gait models for recognition purposes," *Comput. Vis. Image Underst.*, vol. 90, no. 1, pp. 1–41, 2003.

[11] Y. Chai et al., "A novel human gait recognition method by segmenting and extracting the region variance feature," in *Proc. Int. Conf. Pattern Recognit. (ICPR)*, 2006, vol. 4, pp. 425–428.

[12] Y. Yang, D. Tu, and G. Li, "Gait recognition using flow histogram energy image," in *Proc. Int. Conf. Pattern Recognit. (ICPR)*, 2014, pp. 444–449.

[13] R. Lun and W. Zhao, "A survey of applications and human motion recognition with microsoft kinect," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 29, no. 05, pp. 1555008, 2015.

[14] J. Man and B. Bhanu, "Individual recognition using gait energy image," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 2, pp. 316–322, 2006.

[15] K. Bashir, T. Xiang, and S. Gong, "Feature selection for gait recognition without subject cooperation.," in *BMVC*, 2008, pp. 1–10.

[16] T. Whytock, A. Belyaev, and N. Robertson, "Dynamic distance-based shape features for gait recognition," *J. Math. Imaging Vis.*, vol. 50, no. 3, pp. 314–326, 2014.

[17] W. Zeng, C. Wang, and F. Yang, "Silhouette-based gait recognition via deterministic learning," *Pattern Recognit.*, vol. 47, no. 11, pp. 3568–3584, 2014.

[18] M. Goffredo, J. Carter, and M. Nixon, "Front-view gait recognition," in *IEEE Int. Conf. Biometrics: Theory, Appl. and Systems (BTAS)*. IEEE, 2008, pp. 1–6.

[19] D. Tan, K. Huang, S. Yu, and T. Tan, "Recognizing night walkers based on one pseudoshape representation of gait," in *CVPR*. IEEE, 2007, pp. 1–8.

[20] M. H. Khan, M. S. Farid, and M. Grzegorzek, "Person Identification Using Spatiotemporal Motion Characteristics," in *Proc. Int. Conf. Image Process. (ICIP)*, pp. 166–170, 2017.

[21] H. Wang et al., "A robust and efficient video representation for action recognition," in *Int. J. Comput. Vis.*, Springer, 2016, vol. 119, no.3, pp. 219–238.

[22] M. Hofmann, S. Bachmann, and G. Rigoll, "2.5D gait biometrics using the depth gradient histogram energy image," in *IEEE BATS Conf.*, 2012, pp. 399–403.

[23] J. Liang et al., "Appearance-based gait recognition using independent component analysis," in *Int. Conf. on Natural Computation*. Springer, 2006, pp. 371–380.

[24] L. Wang et al., "Automatic gait recognition based on statistical shape analysis," *IEEE Trans. Image Process.*, vol. 12, no. 9, pp. 1120–1131, 2003.

[25] M. H. Khan, F. Li, M. S. Farid, and M. Grzegorzek, "Gait recognition using motion trajectory analysis," in *Proc. Int. Conf. Comput. Recognit. Systems (CORES)* pp. 73–82, 2017.

[26] "CMU motion capture database," http://mocap.cs.cmu.edu/.

[27] H. Wang and C. Schmid, "Action recognition with improved trajectories," in *IEEE ICCV*, 2013, pp. 3551–3558.

[28] J. Sánchez et al., "Image classification with the fisher vector: Theory and practice," *Int. J. Comput. Vis.*, vol. 105, no. 3, pp. 222–245, 2013.

[29] X. Peng, L. Wang, X. Wang, and Y. Qiao, "Bag of visual words and fusion methods for action recognition: Comprehensive study and good practice," *Comput. Vis. Image Underst.*, vol. 150, pp. 109 – 125, 2016.

[30] Yuichiro Anzai, Pattern Recognition and Machine Learning, *Elsevier*, 2012.

[31] Ralph Gross and Jianbo Shi, "The cmu motion of body (mobo) database," 2001.

[32] F.M. Castro et al., "Automatic learning of gait signatures for people identification," *arXiv preprint arXiv:1603.01006*, 2016.

[33] W. Kusakunniran, "Attribute-based learning for gait recognition using spatio-temporal interest points," *Image Vis. Comput.*, vol. 32, no. 12, pp. 1117–1126, 2014.

[34] A. Veeraraghavan et al., "Matching shape sequences in video with applications in human movement analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 12, pp. 1896–1909, 2005.

[35] A. Veeraraghavan et al., "Role of shape and kinematics in human movement analysis," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*. IEEE, 2004, vol. 1, pp. I–730.

[36] S. Das Choudhury and T. Tjahjadi, "Silhouette-based gait recognition using procrustes shape analysis and elliptic fourier descriptors," *Pattern Recognit.*, vol. 45, no. 9, pp. 3414–3426, 2012.

[37] W. Kusakunniran et al. "Automatic gait recognition using weighted binary pattern on video," in *IEEE Int. Conf. Adv. Video Signal Based Surveillance (AVSS)*. IEEE, 2009, pp. 49–54.

[38] D. Tan et al., "Uniprojective features for gait recognition," in *Int. Conf. Biometrics (ICB)*. Springer, 2007, pp. 673–682.

[39] D. Tan, S. Yu, K. Huang, and T. Tan, "Walker recognition without gait cycle estimation," in *Int. Conf. on Biometrics*, 2007, pp. 222–231.