ROBUST VIEW TRANSFORMATION MODEL FOR GAIT RECOGNITION

Shuai Zheng, Junge Zhang, Kaiqi Huang, Ran He, Tieniu Tan

National Laboratory of Pattern Recognition,
Institute of Automation, Chinese Academy of Sciences
{szheng,jgzhang,kqhuang,rhe,tnt}nlpr.ia.ac.cn

ABSTRACT

Recent gait recognition systems often suffer from the challenges including viewing angle variation and large intra-class variations. In order to address these challenges, this paper presents a robust View Transformation Model for gait recognition. Based on the gait energy image, the proposed method establishes a robust view transformation model via robust principal component analysis. Partial least square is used as feature selection method. Compared with the existing methods, the proposed method finds out a shared linear correlated low rank subspace, which brings the advantages that the view transformation model is robust to viewing angle variation, clothing and carrying condition changes. Conducted on the CASIA gait dataset, experimental results show that the proposed method outperforms the other existing methods.

Index Terms— Gait Recognition,View Transformation Model, Low-rank

1. INTRODUCTION

Gait is a potential behavior biometric used to recognize and analyze person at distance [1], but the deviation of gait representation in practical visual surveillance scenario often gets out of control due to the viewing angle variation, large intra-class variations. Gait features of the same subject differ from each other due to various degrees of intra-class variability. Recent work [2, 3, 4, 5] proved that the single gait recognition performance drops when computing similarity measurement of gait appearances in different viewing angles. In that scenario, the viewing angle of probe gait is not the same as that of gallery data. In this paper, we addressed multi-view gait recognition problem under the difficulties caused by wearing or carrying condition variations.

Previous work have been proposed to address the problem of multi-view gait recognition, but most of them do not handle well with associated variations (i.e. carrying or wearing condition change) in visual surveillance scenarios together. Recently, Makihara et.al [6] established a Singular Value Decomposition (SVD) based view transformation model to transform the gait features in probe viewing angle to that in gallery viewing angles. In addition, Kusakunniran et al.
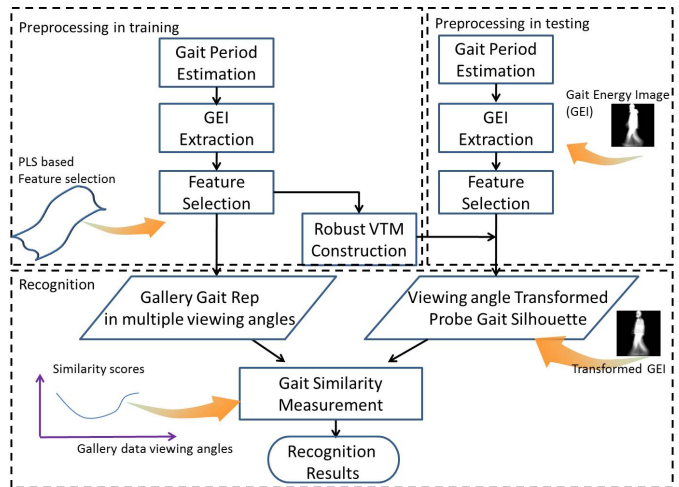
Fig. 1. Overview framework of Robust View Transformation Model based Gait Recognition.

[7, 8] recently demonstrated the advantages of hand craft rank reduced SVD approach and multiple layer perceptron regression models. In fact, gait representation is suitable for low-rank modeling since the correlation betweeen frames in gait image sequence. One of the most challenge tasks in gait recognition is to develop a good representation model which is invariant to the variations including viewing ange change, wearing different coats or taking a bag. Hence, the new robust view transformation model using robust low-rank method [9, 10] needs to be considered.

As illustrated in Fig.1, we propose a Robust View Transformation Model (Robust VTM) using robust Principal Component Analysis (robust PCA) [9]. The proposed multi-view gait recognition system consists of two separated procedures, one is gait signature registration, and the other is gait recognition. For each process, the gait energy image is chosen as the original gait feature descriptor, while the Partial Least Square (PLS) based feature selection method is adopted. During the registration process, the roubst VTM is constructed while the view transformation projection and feature selection functions are learned. The gait features of probe viewing angle are transformed into that of gallery viewing angles. Then gait similarity measurement is conducted to produce recognition result.

The remainder of this paper is organized as follows. Section 2 describes the optimized gait representation. The detail of robust view transformation model is introduced in Section 3. Section 4 presents experimental results to demonstrate the effectiveness of the proposed algorithm as well as its advantages. Finally, Section 5 concludes this work.

2. OPTIMIZED GAIT REPRESENTATION

In this section, we propse a Partial Least Square feature selection based optimized gait representation. Normal human walking is a periodical action. To preserve the temporal information and reduced unnecessary computation cost, we need to detect period. We estimate the bounding box changes using the methods illustrated in [7, 2], since the aspect ratio of silhouette bounding box change periodically during person walk. Besides, Gait Energy Image (GEI) [11] has been constructed as gait feature descriptor based on the previous period estimation. In addition, Partial Least Square feature selection method [12] is employed to extract discriminative part of the gait feature descriptor.

Based on the results from the period estimation, GEI is used as gait representation for the gait information in spatial and temporal domain. The silhouettes extracted from background modeling to construct gait energy image. Suppose each $I_{n,t}(x,y)$ is a particular pixel located at position (x, y) of $t(t = 1, 2, ..., T)$ image from $n(n = 1, 2, ..., N)$ gait cycle. All the silhouettes are normalized along both horizontal and vertical directions to a fixed size. Assuming that width and height of GEI are W and H representatively, GEI is defined as

$$g(x,y) = \frac{1}{T} \sum_{n=1}^{N} \sum_{t=1}^{T} I_{n,t}(x,y), \qquad (1)$$

where T is the number of frames in gait sequence. I is a silhouette image at frame t, x and y are the image coordinates.

The original GEI feature representation is a 1-D vector, g_k^m, by concatenating the value of each position in $I_{n,t}(x,y)$ along all consecutive rows, where m represents the $m_{t}h$ subject and k represents the $k_{t}h$ viewing angle. Thus the dimension of the g_k^m is $W \times H$.

We employ the Partial Least Square (PLS) regression [12] as feature selection algorithm to learn optimal featur representation vectors. PLS is an efficient supervised dimension reduction approach used as a feature selection method. It also brings the advantages that the target reduced dimension does not limited by the class number of training dataset. In addition, by applying PLS on GEI, similar as that in [7], the optimized GEI is expected to be better factorized than the original spatial-domain GEI.

Given two sets of gait feature vectors from different subject under same viewing angle, i.e. g_k^m is obtained from the m_{th} subject under the k_{th} viewing angle while g_k^n is obtained from the n_{th} subject under the same viewing angle. PLS computes an optimal projection by searching the directions for

maximum following object function between two variables

$$\max_{w_k}[\text{cov}(g_k^m w_k, g_k^n)]^2, \qquad (2)$$

where w_k is the learned projection matrix of the k_{th} viewing angle. Cov operation means to compute the covariance between original GEI feature representation vectors from different individuals under the same viewing angle. Hence, given a new GEI feature representation vector g_k^m, we learn the optimal gait feature vectors o_k^m under the k_{th} viewing angle via

$$o_k^m = g_k^m w_k. \qquad (3)$$

3. ROBUST VIEW TRANSFROMATION MODEL

An optimized gait representation matrix o_k^m is created as the left hand side matrix in equation (4). Each row contains the gait information from different subjects under the same viewing angle while each column includes that from the same subjects under different viewing angles. In that case, there are total K viewing angles and M subjects for constructing VTM. As illustrated in [7], The factorization process by SVD is defined as:

$$O = \begin{bmatrix} o_1^1, & \cdots, & o_1^M \\ \vdots & \vdots & \vdots \\ o_K^1 & \cdots & o_K^M \end{bmatrix} = P \begin{bmatrix} v^1 & \cdots & v^M \end{bmatrix}, = USV^T \quad (4)$$

where U is the $KN_g \times M$ orthogonal matrix. The dimension of o_k^m that is supposed to be $N_g \times 1$. V is the $M \times M$ orthogonal matrix. S is the $M \times M$ diagonal matrix contains the singular values. $P = [P_1, \cdots, P_K]^T = US$, where P_k is the $N_g \times M$ sub-matrix of US. v^m is M dimensional column vector.

The vector v^m is a shared gait feature of the m_{th} subject from any viewing angle. P_k is a transforming matrix which can project shared gait feature vector v to the gait feature vector under specific viewing angle k. P_k is independent of the subject. Based on equation (4), given an optimized gait feature vector o_j^m from the m_{th} subject under the j_{th} viewing angle, the learned VTM transform gait feature vectors from the j_{th} viewing angle to the i_{th} viewing angle, and obtain the transformed one as

$$to_i^m = P_i P_j^+ o_j^m \qquad (5)$$

Where P_j^+ is the pseudo inverse matrix of P_j.

Motivated by the improvement of truncated Singular Value Decomposition (TSVD) [7], the reduced rank approximation for SVD achieves better performance in gait recognition. In fact, gait representation is complicated by the appearance variation under different viewing angle for the same subject, as well as the variability in the same subject at different time, for example due to the appearance change of carrying or wearing conditions. In many cases, however, it is reasonable to assume that the shared components for the same subject under different viewing angles are low-rank,
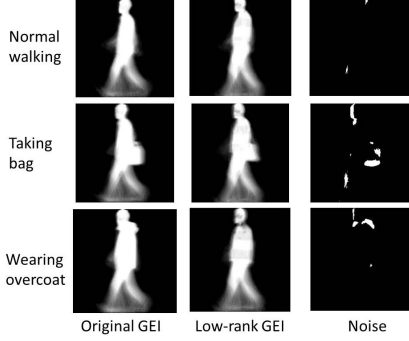
Fig. 2. Robust View Transformation Model for Gait Representation.

Gallery viewing angle	54°	72°	108°	126°	144°
FG+SVD[6]	0.29	0.40	0.45	0.30	0.20
GEI+LDA+TSVD[7]	0.50	0.70	0.72	0.40	0.20
Rectified method[5]	**0.72**	0.70	0.68	**0.66**	-
Yu's method[3]	0.16	0.81	0.87	0.22	0.03
Ours	0.42	**0.86**	**0.88**	0.50	**0.26**

Table 1. Recognition rates of different gait recognition algorithms under the condition that the probe gait data under viewing angle 90° and the gallery gait data under viewing angle from 54° to 144°.

while the noise caused by the appearance change of carrying a bag or wearing overcoats is sparse. Figure 2 illustrated the robust view transformation model for gait representation.

Given gait matrix O, A represents the low-rank shared components of gait features used in SVD, while E is a noise component matrix. In order to obtain an optimal shared subspace, the criteria of proposed robust transformation model is to minimize the following objective function as

$$\begin{cases} \min_{A,E} rank(A) + \lambda\|E\|_0, \\ s.t., O = A + E. \end{cases} \quad (6)$$

where $A = USV^T$, which can be considered as the low-rank SVD method. However, it is difficult to give the solution of the above optimization problem due to the discrete nature of the rank function. Assuming that noisy components are sufficiently sparse, we relax the equality constraint and get the following convex relaxation problem as

$$\min_{A,E} \|A\|_* + \lambda\|E\|_1, s.t. O = A + E, \quad (7)$$

where $\|\bullet\|_*$ denotes the matrix nuclear norm operation, i.e., the sum of its singular values. $\|\bullet\|_1$ denotes the sum of the absolute values of matrix entries, and λ is a positive regularization parameter.

In order to solve the optimization problem in equation (7), we employed the inexact Augmented Lagrange Multiplier [13] method to recover the missing entries of a matrix and achieve matrix completion.

4. GAIT SIMILARITY IN GAIT RECOGNITION

The purpose of this paper is to present a view transformation model via factorization process. Hence we simplify the gait similarity measurement. We adopted L1-norm distance, then the gait similarity measurement is defined as

$$d(o_k^i, o_k^j) = \left\| o_k^i - o_k^j \right\|. \quad (8)$$

where d is a distance between gait signatures. N is the dimension of gait feature. The smaller value of d means the larger similarity between gait feature representation o_k^i and o_k^j.

5. EXPERIMENTS

Experiments are conducted on CASIA gait dataset[3]. The dataset consists of the data from 124 subjects under 11 viewing angles. Besides, there are 6 sequences from each normal walking subject under one of the 11 viewing angles. In addition, the dataset also contains two sequences from each normal walking subject wearing overcoat or carrying a bag under one of the 11 viewing angle.

Since LDA based optimized Gait Energy Image with Trucated Singular Vector Decomposition(GEI+TSVD) [7] and Frequency domain of Gait features with Singular Vector Decomposition (FG+SVD) [6] are all the view transformation models via factorization process, they are comparable with the proposed method. We also consider some results from [3, 5], although they are not belong to the factorization based VTM. We take into account the multi-view gait recognition with additional challenges caused by the change of carrying or wearing conditions. In order to prove the effectiveness of propose robust VTM, we only considered the (GEI+TSVD) approach. The difference is that Frequency-domain representation of Gait features (FG) are used rather than GEI in [6]. While, based on the LDA optimized GEI, [7] employed the TSVD based VTM. In addition, [3] can be considered as a baseline of CASIA gait dataset and [5] provides a self-calibrating gait feature representation.

During experiments, we compared several related view transformation models under the same circumstance. Original features are extracted from gait energy images. Variations among different methods are the feature selection algorithm and the view transformation model. Given a probe gait feature vector from a specific view, the view transformation model and the feature selection algorithm are employed to transform the given probe gait feature vector into the shared subspace under the gallery viewing angles. The L1-norm distance is employed to measure the similarity of different gait feature representation vectors.

To evaluate the performance of different view transformation models with its associated feature selection algorithm, gait recognition rate is employed to evaluate the correct matching numbers of gait feature vectors. Besides, three tasks are designed to evaluate the algorithms. One is to evaluate the scenario that the gait recognition performance in the condition that person registered in normal walking and tested

Gallery viewing angle	54°	72°	90°	144°	162°
FG+SVD[6]	0.20	0.29	0.48	0.60	0.40
GEI+LDA+TSVD[7]	0.30	0.43	0.72	0.89	0.48
Rectified method[5]	**0.71**	0.59	0.60	0.70	-
Yu's method[3]	0.21	**0.68**	0.82	0.48	0.03
Ours	0.31	0.60	**0.89**	**0.89**	**0.60**

Table 2. Recognition rates of different gait recognition algorithms under the condition that the probe gait data under viewing angle 126° and the gallery gait data under viewing angle from 54° to 162°.

Gallery viewing angle	72°	90°	108°	144°	162°
Ours	**0.24**	**0.26**	**0.53**	**0.88**	**0.31**
GEI+LDA+TSVD[7]	0.09	0.10	0.20	0.30	0.13
Yu's method [3]	0.14	0.09	0.06	0.18	0.02

Table 3. Recognition rates of different gait recognition algorithms under the condition that the probe gait data under viewing angle 126° and the gallery gait data under viewing angle from 72° to 162^o(with overcoat).

in normal walking. The second is to evaluate the gait recognition performance where person registered with a bag and tested with a bag. And the third is to evaluate the gait recognition performance where person registered wearing overcoat and tested wearing overcoat. The dataset collected from 24 subjects is used to train VTM. The remaining datasets from 100 subjects are separated into probe and gallery group to simulate the registration and test process.

As illustrated in Tab.2,Tab.1, Tab.3, and Tab.4, the proposed approach outperforms GEI+LDA+TSVD, FG+SVD, Yu's method. Since the Robust VTM helps to preserve more global structures in intrinsic low-rank dimensionality and detach view-independent gait features from the appearance variations caused by wearing overcoats or carrying a bag, It is not surprise that it performs better. Besides, the proposed method also beats the method used in [5] in terms of the best recognition rates. And, the proposed method outperforms than GEI+TSVD methods, when they both used the same feature selection method PLS.

6. CONCLUSIONS

In this paper, we present Robust View Transformation Model and Partial Least Square feature selection algorithm for multi-view gait recognition under different wearing and carrying conditions. Assumed that shared gait appearance features in different viewing angles are low rank and the distribution of additional variations are sparse, the proposed methods bring the significant satisfied performance on those multi-view gait dataset with additional variation in wearing or carrying conditions. Besides, the proposed model adopts the inexact Augmented Lagrange Multiplier (ALM) method as a solver to ensure the efficiency. In the future, we plan to evaluate the proposed method in more difficult datasets such as Human ID dataset or the dataset from visual surveillance systems.

Gallery view angle	54°	72°	108°	126°	144°
Ours	**0.19**	**0.59**	**0.55**	**0.23**	**0.10**
GEI+LDA+TSVD[7]	0.10	0.31	0.23	0.13	0.10
Yu's method [3]	0.13	0.31	0.44	0.15	0.02

Table 4. Recognition rates of different gait recognition algorithms under the condition that the probe gait data under viewing angle 90° and the gallery gait data under viewing angle from 54° to 144^o(With a bag).

7. REFERENCES

[1] S.Sarkar, P.Janathon Phillips, Z.Liu, I.Robledo, P.Grother, and K.W.Bowyer, "The human id gait challenge problem: data sets, performance, and analysis," *IEEE Trans. on PAMI*, 2005.

[2] L.Wang, T.Tan, H.Ning, and W.Hu, "Silhouette analysis based gait recognition for human identification," *IEEE Trans. on PAMI*, 2003.

[3] S.Yu, D.Tan, and T.Tan, "A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition," in *ICPR*, 2006.

[4] K.Bashir, T.Xiang, and S.Gong, "Cross-view gait recognition using correlation strength," in *BMVC*, 2010.

[5] M.Goffredo, I.Bouchrika, J.N.Carter, and M.S.Nixon, "Self-calibrating view-invariant gait biometrics," *IEEE Trans. on SMC- part B*, 2010.

[6] J.Makihara, R.Sagawa, and Y.Mukaigawa, "Gait recognition using a view transformation model in the frequency domain," in *ECCV*, 2006.

[7] W.Kusakunniran, Q.Wu, H.Li, and J.Zhang, "Multiple view gait recognition using view transformation model based on optimized gait energy images," in *ICCV workshops*, 2009.

[8] W.Kusakunniran, Q.Wu, J.Zhang, and H.Li, "Multi-view gait recognition based on motion regression using multiplayer perceptron," in *ICPR*, 2010.

[9] E.Candes, X.Li, Y.Ma, and J.Wright, "Robust principal component analysis?," *Journal of the ACM*, 2010.

[10] R.He, B.Hu, W.Zheng, and Y.Guo, "Two-stage sparse representation for robust recognition on large-scale database," in *AAAI*, 2010.

[11] J.Hua and B.Bhanu., "Individual recognition using gait energy image," *IEEE Trans. on PAMI*, 2006.

[12] R.Roman and N.Kramer, "Overview and recent advances in partial least squares in subspace," in *Latent Structure and Feature Selection Techniques*, 2006.

[13] Z.Lin, M.Chen, L.Wu, and Y.Ma, "The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices," in *UIUC Technical Report*, 2009.