

Manifold Learning Based Gait Feature Reduction and Recognition

Suyuan Wei

Research Inst. Of Computer Peripheral, Xidian Univ., Xi'an, China

Xi'an High Technology Inst., Xi'an, China

Email: wei_suyuan2002@163.com

Chao Ning,

Xi'an High Technology Inst., Xi'an, China

Youxing Gong,

Research Inst. Of Computer Peripheral, Xidian Univ., Xi'an, China

Xi'an High Technology Inst., Xi'an, China

Abstract—The moving objectives' images are in tensor format in reality. That using for reference the thought of tensor space dimension reduction to gain the optimal gait characters with low dimension inaugurate a new gait recognition way. A novel gait expression and recognition algorithm based on the tensor space is introduced here. It is a tensor space learning algorithm that could investigate the inherent geometrical structure of the data manifold. The within-class and the between-class similarity graphs are respectively defined so as to preserve the local structure of the manifold and the global data information. It improves the ability of gait data reconstructing and the recognizing efficiency. The optimization problem of finding the optimal tensor subspace is deduced to an iteratively computation problem about resolving the generalized eigenvectors. The optimal tensor is used to express the gait character and recognize the individual. And it reduced the gait character dimension, at the same time the storage and calculation cost were cut down. The experiments with the SOTON gait database demonstrated the validity of the proposed method. And the comparison among the tensor subspace analysis, the principal component analysis, the linear discriminate analysis and proposed method showed that the recognition performance of the our improved algorithm outperformed others.

Index Terms—gait character extraction, gait recognition, tensor space mapping, dimension reduction

I. INTRODUCTION

Gait recognition techniques identify people based on walking pattern of an individual. The outstanding advantage of gait recognition towards other biometrics is that it is inviolable, hard to conceal, and can be achieved using images taken from a long distance. As a valid biometric the gait could be used in the security surveillance to recognize the individual in the distance without the high resolution. Now the gait recognition is becoming a research focus in the computer vision and the pattern recognition.

Expressing the gait character based on the individual appearance is an important method [1][2][3][4][5][6][7].

This method projects the three-dimension gait data into a two-dimension space when extracting the gait characters. These two-dimension images are always expressed by the vectors. Those classical algorithms as the principle component analysis (PCA), the linear discriminant analysis (LDA) and the locality preserving projections (LPP), etc., adopt the vector expression to project the high dimension space into the low dimension space for reducing the dimension of the gait character data, then cutting down the computation cost. But the image exists in the form of tensor in fact. The vector expression is not enough for all image structure. The proposed tensor facial recognition algorithm brought a new way for dealing with the image pattern recognition. Such algorithms that the two-dimension principle component analysis (2DPCA), the two-dimension linear discriminant analysis (2DLDA) and the tensor subspace analysis (TSA) couldn't give attention to both the global and the local data structures. And they didn't sufficiently make use of the within-class information and the between-class information when classifying the samples.

Using for reference the thought of combining the global and the local characters for classification, a tensor space gait expression method is proposed here to recognize gait. This algorithm not only followed the tensor subspace analysis preserving the data manifold local geometrical structure, but also preserved the global structure sufficiently, which improved the data reconstructive ability and the recognition performance. Simultaneously obtaining optimal tensor subspace reduced the dimension of the gait characters, and cut down the storage and computation cost.

II. RELATED WORK

Vector expression of images is playing significant part in research of biometrical recognition. However, actually images exist in tension format, whose information can not be fully represented in vectors. The approaches, using vector based feature space as two-dimension principle component analysis (2DPCA) and two-dimension

linear discriminant analysis (2DLDA) , only focused on maintaining full dimension of manifold in Euclidean space, and always neglected particular architectures of manifold locally.

Manifold learning based dimensionality reduction for high dimensional data model was one of most active research fields in gait recognition. Locally Linear Embedding [8] was an improved algorithm than data embedding and nonlinear dimensionality reduction. And it targeted to keep intrinsic geometrical position, as well as global structure. It mapped data from high dimensional space to low dimensional space by a recessive mapping method. To solving the difficulty of processing new measured data-set, Neighborhood Preserving Embedded Projection [9] was proposed and applied to face recognition. And it achieved the better results. Tensorfaces presented by Vasilescu developed the face image expression from vector space to tensor space. It performed the Singular Value Decomposition in tensor space, and decomposed the high rank tensor of the face image into several factors as lights, face expressions and poses according to different orientations, which made the face recognition more convenient and accurate. Tensor Subspace Analysis (TSA) [10], used a directionless powered graph to model manifold in order to realize the constructed dimension reduction. Its disadvantage was that it ignored the global structure as holding the local relationship of the manifold, and no distinctive differences between the internal and dextral class information made it deficient for reconstruct data. Tensor Locality Discriminant Projection (TLDP) [11], constructed the directionless powered graph, and embedded the graph into the tensor subspace to find the low dimension feature forms. TLDP inherited the advantage of TSA to hold the manifold local relationship. At the same time, it paid more attention to hold the global discriminant structure. It improved successfully the data re-constructive ability and the recognition efficiency. With the same capability of holding the manifold local relationship, the difference between TNPE and TSA laid in that TNPE reconstructed the local linear coefficients to hold the manifold local structure, while TSA made up the neighborhood graph to model the manifold local relationship. TNPE had to estimate the linear coefficients of each sample that cost more calculations than TSA.

Manifold learning based dimensionality reduction for high dimensional data model is one of most active research fields in gait recognition. Locally Linear Embedding was an improved algorithm than embedded data and nonlinear dimensionality reduction. And it targeted to keep intrinsic geometrical position, as well as global architecture. It mapped data from high dimensional space to low dimensional space by a recessive mapping method.

Before gait recognizing a normal step is to abstract the gait features and reduce their dimensions to gain the efficient discriminate sample characters with low dimension. Yan and Xu proposed an intuitionistic embedded graphic frame which expatiated those data dimension reduction algorithms, such as LDA, PCA, etc.

Using one uniform model, Marginal Fisher Analysis [12] overcame some shortcomings of LDA, completed dimension reduction of matrix and manifold, and had the capability to directly process the 2D gray images. Orthogonal Marginal Fisher Analysis, OMFA [13] combined orthogonal criterion and cost function to calculate the transformation matrix under given constraints.

TLDP and TMFA had some similarities that they all kept the discriminant information. When modeling local configuration, TLDP needed only one critical parameter, but TMFA needed two. The relatively higher intricacy of preferences resulted in lesser precision when depicting local structure of manifold, comparing to TLDP. Locality Preserving Projections, LPP [14][15], could do Laplacian Eigenmap under linear constraint, then figure out the mapped points for data in high dimension space to low dimension space. One strong-point is that LPP preserved local structure, and maintained relativity of data. Comparing to algorithms as PCA and LDAP, which were used to keep global architectures, LPP could get projections efficiently carrying the information of manifold structure. Another advantage of LPP is the linearity of projection. It was applied not only to find sample data's counter-parts in low dimension space, but also to find out their correspondent mapping matrix. As a result, when new sample points were added, their expression in low dimension could be easily conducted too.

Although LPP itself was one of the linear algorithms, which was common to PCA and LDA, PCA and LDA prefer to maintain the global characteristic, while LPP and LPPCA (Locality Preserving Canonical Correlation Analysis) [16] paid more attention to local data structure. In realistic cases, most clarification issues were more related to the local pattern, so LPP had more powerful recognition capabilities than other approaches. In calculation efficiency's point of view, nonlinear algorithm was comparatively slower than linear method. Exploring linear mapping, LPP was quite effective. Meanwhile, the representative database for local structure was stored in adjacency matrix in LPP, and this gained some good straits of LLE. Moreover, when a new sample point was need to be recognized, LPP could quickly classify it to sub-group it belonged to, but LEE had to recalculate all data to get new clarification for all sub-groups. It was very rational and could save large amount of time when processing huge sample data in high dimension space.

III. GAIT EXPRESSION

A gait sequence sample is comprised of several gait images sampled in a stretch. The difference among frames shows the gait difference among the walkers on the time axes. If the 3rd dimension, the time field, gait features of the samples could be expressed in the two dimension image, the gait data would be reduced greatly, and efficiency of the post-processing would be correspondingly improved.

Along with time variety the figure change in a gait image directly reflects the gait variety. If the difference between frames is taken as the weight of image pixel energy, a Spatial-Temporal Energy image (STE) could be defined to express the body outline of an individual. As well as it could contain the walking orientate and other information about the motion feature relative to time.

When taking the gait recognition with the low resolving capability equipment in the distance, it is not advisable by all appearance that the facial features and clothing detail etc. are seen as the main characters. The binary contour line image, which ignored the inside detail, is used here to extract the gait characters.

The Spatial-Temporal Energy (STE) could be calculated with the contour frame series which contains frames number more than that of a gait cycle.

$$STE(x, y) = \frac{1}{nFrame} \sum_{t=1}^{nFrame} C(x, y, t) \quad (1)$$

Whereas $C(x, y, t)$ is a contour image at moment t , and there are $nFrame$ normalized frames in a sequence. The pixel brightness denoted the cumulative energy at this point when walking through, which is the function of spatial variety by the time. Since the difference of the individual gait comprised mainly the posture and the walking habit, the STE combined these two. Thereby the gait characters could be extracted from the STE and identified the unique individual.

STE could be normalized in $[0, 255]$ as per following equation:

$$STE(x, y) = \frac{1}{nFrame} \sum_{t=1}^{nFrame} C(x, y, t) \times \frac{255}{I_{max} - I_{min}} \quad (2)$$

Where I_{max} and I_{min} are the maximum and the minimum brightness of STE respectively.

Relative to those time series-based gait representation, STE is more robust for the noisy binary image. And it doesn't depend on the choice of the start stance in a gait cycle, thus the gait alignment is unnecessary. It is established that the complex gait recognition involved many factors as surface, time variation and shoes etc. And the low performance can not be explained by poor silhouette quality [17]. After converted the silhouette into the contour line image, the store space is reduced to 20%. When a grey-level STE replaced a gait sequence, the required space reduced more, much less than 1% that of the original silhouette sequence.

IV. DISTINGUISHABILITY OF GAIT CHARACTERISTICS IN TENSOR EXPRESSION

A Manifold and Tensor

A manifold is a mathematical space that on a small enough scale resembles the Euclidean space of a specific dimension, called the dimension of the manifold. Every point of an n -dimensional manifold has neighborhood homeomorphisms to the n -dimension space \mathbf{R}^n . In mathematics manifold is used to describe the geometric body. It provides a natural platform to research the geometric differentiability. If the topological structure of

a geometric body is looked upon as completely soft, since all distortions could maintenance the topological structure invariability, and the resolution cluster is seen as hard, since the total structure is fixed, the lubricity manifold could be taken as a body between these two, namely its infinite small structure is hard while the total body is soft. The manifold hardness make it contain the differential structure and that its softness make it act as the mathematics or physical model including many independent local perturbations. It is not ordinarily using one reference frame to describe the total manifold that the manifold is different from the constructed model, generally a simple space, on the global structure. When several reference frames are used to model the manifold, the overlap regions are important and difficult to settle because of their containing the global information.

There are several ways to construct the manifold. Each method emphasized one aspect of the manifold that different treatment brings the different view. When using the manifold to research the geometric bodies some adjunctive structures are generally needed to enrich these spaces. In mathematics tensor is a geometric body, or the generalized "magnitude". Tensor could be denoted by the reference frame. But it is independent to the selective reference frame. In geometrics, tensor does be a geometric quantity indeed. It is namely that it doesn't vary with the coordinates transform in the reference frame.

The classical method to define a tensor is regarding it as a multi-dimension array, and the weights of the tensor are those values in the array. So as to the tensor field the element of the tensor is function even differential coefficient. The modern weightless method is treating the tensor as an abstract object which expresses certain fixed style for the linear conception.

Perceive science found that the human countenance, hand posture and gait could be expressed using the low dimension manifold that embedded into the high dimension space. The classical vector subspace learning algorithms is based on the vector form of the data manifold. While in reality the gait manifold is in existence in the tensor form essentially. The tensor space analysis would play better role to uncover the inherent geometric structure of the gait manifold.

B DISTINGUISHABILITY OF GAIT DATA IN TENSOR EXPRESSION

Gradient information correlates to the structure of the objects in an image. It marks the discriminative characters of the object, or provides the important clues for the father vision analysis. Accordingly gradient information is often used in the image target recognition and classification. Sampling every individual gait sequence several times, and calculating the Spatial-Temporal Energy images for all sequences as per the equation (1), then the structure tensor for an individual gait data $I(x, y, z)$ as per following equation:

$$S = \begin{bmatrix} I_x^2 & I_x I_y & I_x I_z \\ I_x I_y & I_y^2 & I_y I_z \\ I_x I_z & I_y I_z & I_z^2 \end{bmatrix} \quad (3)$$

Where x and y are the horizontal and the vertical coordinate respectively in the STE for a gait sequence; $z = i$ denotes the i th gait sample for the individual; I_x, I_y and I_z are the partial derivatives respectively along the x, y and z direction. This is the gradient information representation in the form of tensor for image I .

The eigen-decomposition of the tensor of rank two results in the eigen-values $(\lambda_1, \lambda_2, \lambda_3)$ and the eigenvectors (e_1, e_2, e_3) . The interpretation of these components can be visualized as 3D ellipses where the radii are equal to the eigen-values in descending order and directed related to their corresponding eigenvectors.

After transformed these gait sequences images into STE images, it is difficulty to distinguish the walking habit model or other dynamic characters captured by naked eyes, only the appearance differences among part individuals could be divided constrainedly. While figure 1 shows that after abstracting the gradient information in the form of structure tensor for gait data, the differences among individuals could be determined evidently.

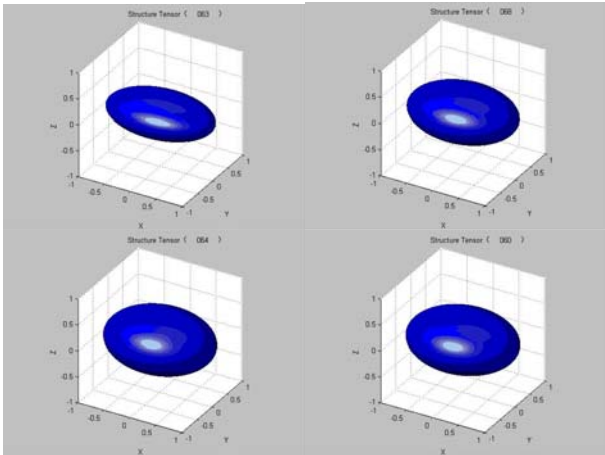


Figure 1. Structure tensor of the gait features

V. GAIT CHARACTER IN THE TENSOR SPACE

A $(m \times n)$ -dimension STE calculated from someone's gait sequence could be seen as a second order tensor (matrix) in the $R^m \otimes R^n$ tensor space, R^m and R^n are m -dimension and n -dimension first order vector spaces respectively. Thus the gait space is a subspace embedded into the $R^m \otimes R^n$ tensor space. It is a nonlinear sub-manifold. It is expected that the geometrical and the topological structure of this unknown sub-manifold could be reconstructed with its random samples without the manifold for the gait subspace of an individual. The geometrical structure of the sub-manifold could be modeled through the neighbor graph constructed by the gait sequence samples. Therefore the obtained projective matrix made the tensor character space

approximate linearly the gait manifold. And the tensor space would be transformed from $R^m \otimes R^n$ to $R^l \otimes R^k$ ($l < m, k < n$), namely the optimal tensor space would be the gait character after the dimension reduction.

If the column vector $\{p_i\}_{i=1}^u$ of matrix P and $\{q_j\}_{j=1}^v$ of matrix Q are respectively the perpendicular basis of R^m and R^n . Then the basis of the tensor space $R^m \otimes R^n$ is $\{p_i \otimes q_j\}$. Projecting any tensor $X \in R^m \otimes R^n$ in $p_i \otimes q_j$ could be calculated by inner product $\{X, p_i \otimes q_j\} = p_i^T X q_j$. $p_i q_j^T$ is the matrix form of the basis tensor $p_i \otimes q_j$. Therefore $X \in R^m \otimes R^n$ could be seen as a $m \times n$ matrix as well as a second order tensor $\sum (p_i^T X q_j) p_i q_j^T$. Suppose Y is the mapping of X in $R^l \otimes R^k$. The tensor reduction could be realized through searching a $(m \times k)$ -dimension transformation matrix P and a $(n \times l)$ -dimension transformation matrix Q , ($l < m, k < n$), which satisfy $Y = P^T X Q$.

If two samples are near enough, they may be from the same individual. The distance between these two points is calculated through a defined function which is taken as the weight in the within-class or between-class similarity matrix.

Given n STE images $X_i, i = 1, 2, \dots, n$, the geometrical structure of the gait sub-manifold are modeled by constructed the samples neighbor graph. The within-class and the between-class similarity weighted matrixes $S_{W,ij}$ and $S_{B,ij}$ are defined as per following equation to figure the structure character of the neighbor graph:

$$S_{W,ij} = \begin{cases} e^{-\frac{\|X_i - X_j\|^2}{\lambda}}, & X_i \in class_k \wedge X_j \in class_k \\ 0, & \text{ot her wi se} \end{cases} \quad (4)$$

$$S_{B,ij} = \begin{cases} e^{-\frac{\|X_i - X_j\|^2}{\lambda}}, & X_i \in class_k \wedge X_j \notin class_k \\ 0, & \text{ot her wi se} \end{cases} \quad (5)$$

Whereas λ is the adjusting coefficient. And the class labels $class_k$ of the sample X_i are available.

According to Laplacian eigenmap[18], the objective function for dimension reduction from $R^m \otimes R^n$ to $R^l \otimes R^k$ is

$$\min_{P,Q} \sum_{i,j} \|Y_i - Y_j\|^2 S_{W,ij} \quad (6)$$

$$\max_{P,Q} \sum_{i,j} \|Y_i - Y_j\|^2 S_{B,ij} \quad (7)$$

If sampled from the same individual, they are more close to each other in the new projected space; contrarily samples from different individuals are farther apart from each other in the new space.

Define a diagonal matrix $D_{ii} = \sum_j S_{ij}$.

Since $\|A\|^2 = tr(AA^T)$, then

$$\begin{aligned} & \sum_{i,j} \|Y_i - Y_j\|^2 S_{ij} \\ &= \sum_{i,j} tr((Y_i - Y_j)(Y_i - Y_j)^T) S_{ij} \\ &= \sum_{i,j} tr(Y_i Y_i^T + Y_j Y_j^T - Y_i Y_j^T - Y_j Y_i^T) S_{ij} \\ &= 2tr(\sum_i Y_i Y_i^T D_{ii} - \sum_{i,j} Y_i Y_j^T S_{ij}) \\ &= 2tr(\sum_i P^T X_i Q Q^T X_i^T P D_{ii} - \sum_{i,j} P^T X_i Q Q^T X_j^T P S_{ij}) \\ &= 2tr(P^T (\sum_i X_i Q Q^T X_i^T D_{ii} - \sum_{i,j} X_i Q Q^T X_j^T S_{ij}) P) \\ &= 2tr(P^T (D_Q - S_Q) P) \\ &= 2tr(P^T D_Q P) - 2tr(P^T S_Q P) \end{aligned} \tag{8}$$

Whereas

$$\begin{cases} D_Q = \sum_i X_i Q Q^T X_i^T D_{ii} \\ S_Q = \sum_{i,j} X_i Q Q^T X_j^T S_{ij} \end{cases} \tag{9}$$

Since $\|A\|^2 = tr(A^T A)$, similar to (8) it has

$$\begin{aligned} \sum_{i,j} \|Y_i - Y_j\|^2 S_{ij} &= 2tr(Q^T (D_P - S_P) Q) \\ &= 2tr(Q^T D_P Q) - 2tr(Q^T S_P Q) \end{aligned} \tag{10}$$

Whereas

$$\begin{cases} D_P = \sum_i X_i P P^T X_i^T D_{ii} \\ S_P = \sum_{i,j} X_i P P^T X_j^T S_{ij} \end{cases} \tag{11}$$

Since $tr(AB) = tr(BA)$, then

$$\begin{aligned} & tr(Q^T D_P Q) \\ &= tr(Q^T \sum_i X_i P P^T X_i^T D_{ii} Q) \\ &= \sum_i D_{ii} tr(Q^T X_i P P^T X_i^T Q) \\ &= \sum_i D_{ii} tr(P^T X_i Q Q^T X_i^T P) \\ &= tr(P^T D_Q P) \end{aligned} \tag{12}$$

Replaced S_{ij} by $S_{W,ij}$ or $S_{B,ij}$ respectively, and based on (8) and (10), the objective function (6) and (7) can be transformed into

$$\min_{P,Q} tr(P^T (D_{WQ} - S_{WQ}) P) \tag{13}$$

$$\max_{P,Q} tr(Q^T (D_{BP} - S_{BP}) Q) \tag{14}$$

Based on [19], $tr(P^T D_Q P)$ is the global variance on the manifold which is a constant for the samples embedded tensor subspace. Thus (13) and (14) equivalent to

$$\max_{P,Q} tr(P^T S_{WQ} P) \tag{15}$$

$$\min_{P,Q} tr(Q^T S_{BP} Q) \tag{16}$$

And the objective function would be rewritten to

$$\min_{P,Q} \frac{tr(P^T (D_{WQ} - S_{WQ}) P)}{tr(P^T S_{BQ} P)} \tag{17}$$

$$\max_{P,Q} \frac{tr(Q^T (D_{BP} - S_{BP}) Q)}{tr(Q^T S_{WP} Q)} \tag{18}$$

According to (17) and (18), P is the generalized eigenvectors matrix of $(D_{WQ} - S_{WQ}, S_{BQ})$, and Q is the generalized eigenvectors matrix of $(D_{BP} - S_{BP}, S_{WP})$. P and Q are difficult to calculate since they are all unknown. Then we can initialize one, for instance P , then calculate Q , and the generalized eigenvector matrix of (19) and (20) are gained by interactive computation.

$$(D_{WQ} - S_{WQ})\mathbf{p} = \lambda S_{BQ}\mathbf{p} \tag{19}$$

$$(D_{BP} - S_{BP})\mathbf{q} = \lambda S_{WP}\mathbf{q} \tag{20}$$

That lists the main steps in the table I for the tensor space based gait recognition algorithm. It is an interactive algorithm. If the size of the image is n , the computational complexity for each step is $O(n^3)$ which is so low that total computational complexity is low too.

TABLE I. TENSOR GAIT CHARACTER MAPPING ALGORITHM

Input:	STE X_1, X_2, \dots, X_n of gait sample, class label and dimension l, k of expected low dimension tensor space
Output:	Transformation matrixes P, Q and gait character Y
Steps(1)	construct the neighbor graph of the samples and the similarity weighted matrix S_W, S_B ;
Steps(2)	set $P = I$, identity matrix, interactive time =M;
Steps(3)	For m = 1 to M
Steps(4)	$D_{B,ii} = \sum_j S_{B,ij}$; $D_{W,ii} = \sum_j S_{W,ij}$;
Steps(5)	$D_{BP} = \sum_i X_i P P^T X_i^T D_{B,ii}$; $S_{BP} = \sum_{i,j} X_i P P^T X_j^T S_{B,ij}$; $S_{WP} = \sum_{i,j} X_i P P^T X_j^T S_{W,ij}$;
Steps(6)	calculate the generalized eigenvector matrix Q of $(D_{BP} - S_{BP}, S_{WP})$;
Steps(7)	$D_{BQ} = \sum_i X_i Q Q^T X_i^T D_{W,ii}$; $S_{WQ} = \sum_{i,j} X_i Q Q^T X_j^T S_{W,ij}$; $S_{BQ} = \sum_{i,j} X_i Q Q^T X_j^T S_{B,ij}$;
Steps(8)	calculate the generalized eigenvector matrix P of $(D_{WQ} - S_{WQ}, S_{BQ})$
Steps(9)	End
Steps(10)	return $P = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_l\}$ and $Q = \{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_k\}$
Steps(11)	return $Y = P^T X Q$

VI. EXPERIMENTS AND ANALYSIS

The SOTON gait database is used in the experiments, which was set up by university of Southampton etc. institutions in British staked by Defense Advanced Research Projects Agency (DARPA).The digital video is taken from the 90° orientation to the walking path. There are 8 to 34 sequences each individual and 30 frame proximately per sequence. And there are 2128 sequences for 115 individuals total. That silhouette images in the SOTON are used directly. 20 individuals and 20 sequences each one are selected randomly from SOTON as the experiment data set as the close-universe model. The recognition performance of the proposed algorithm, PCA, LDA and TSA are tested based the leave-one-out cross-validation with the Probe set and the Gallery set.

Those silhouette area images from SOTON gait database are adopted directly. And all sequences must be preprocessed to convert the silhouette area images into the contour line images, extract the object and normalize to 50 × 32 pixels. Fig.1 shows examples of contour extraction at the left 8 columns. STE for each sequence in the galley set is calculated as per equation (1). Fig.2 shows STE examples at the right column.

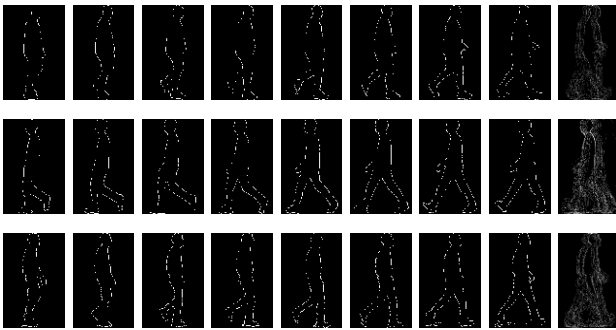


Figure 2.Examples of contour extraction and STE

The optimal gait character subspace including the transformation matrix and character tensor could be gained as per the algorithm listed in table I. And all samples from the Probe set and Galley set are projected into the character tensor subspace. Calculate the Euclidian distances between the character of the test gait and one from the Gallery set. Then the test gait is identified through the nearest-neighbor classifier. The identification are repeated 100 times with different Probe set and Gallery set that the mean are taken as the result.

The experiment results list in the table II, which compared the efficiency of four algorithms, two based on the vector space and another two on the tensor space.

TABLE II. COMPARISON OF ALGORITHM EFFICIENCY ON SOTON

Method	CCR	Dimension	Time(s)
LDA	82.0%	1600	2772.11
PCA	83.5%	160	541.81
TSA	86.3%	12 × 8	171.62
Ours	87.5%	12 × 8	173.46

Singular value problem and the high dimension problem: If the tensor was transformed into the vector as per LDA, the similarity matrix $S_{B,ij}$ would has the size of

$\prod_{i=1}^l m_i \times \prod_{i=1}^k m_i$. It is common that the class number of the gait samples in the gait recognition is less than it. Thus LDA method would meet the singular value problem. While in our method here the dimension of $S_{B,ij}$ is $m_i \times m_i$ for each step. And the number increased

per step for the gait samples is bigger than m_i . The singular value problem could be overcome. Moreover since the objective dimension is such low the calamity for data having high dimension could be settled at certain degree.

Rank Order Statistic (ROS) is another evaluation measurement for algorithm performance, which is widely applied on facial recognition [20] and gait recognition. It denotes the accumulating probability $p(k)$ that the true class in the top n matches corresponding to the testing sequence. The performance statistics are reported as the cumulative math score, which are plotted on a graph. The horizontal axis of the graph is rank and the vertical axis is the probability of identification, namely the cumulative match score. Fig. 3 reported the identification performance of these four methods.

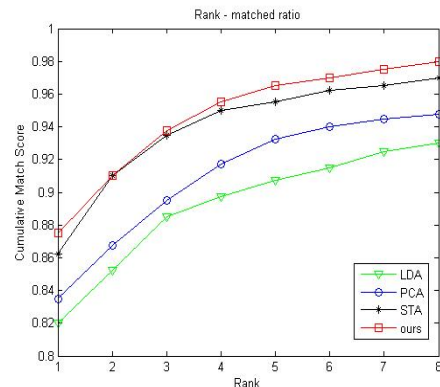


Figure 3.Recognition results with the FERET

Figure 3 indicates that our method, manifold learning based gait feature reduction and recognition, has got ahead of other three algorithms, LDA, PCA and STA, on ROS scale. Those tensor space based methods exceed that vector space based algorithms since the reconstructed the objective manifold through the tensor analysis is evidently more close the reality.

Table II also shows that the recognition performances of the tensor algorithms outperform that of the vector algorithms evidently, no matter CCR, feature dimension scale or the computational time. The faster computational speed profit from the gait character space reduction directly. Notwithstanding the dimension of the result character space from PCA has the same quantitative scale as that from TSA and the manifold learning based gait recognition. The transformation matrix is more complex.

Thus the processing time is longer than those other algorithms.

In the manifold learning base gait recognition a manifold is described by a non-direction weighted graph which is a similar thought followed the tensor subspace analysis. After set up the neighbor graph the local geometrical structure of the manifold would be modeled. Then STE of these gait sequences are embedded into the tensor space. The dimension reduced form, namely the optimal tensor subspace, would be taken as the gait character space. Constructing the neighbor graph is a simple way for reconstructing the local structure of a manifold. Here the within-class similarity and the between-class similarity are discriminated when defining the weighted matrix for the neighbor graph.

Relative to the tensor subspace analysis, our method utilized the class information more sufficiently that preserved not only the local structure but the global structure. Moreover the experiments demonstrate that these two tensor algorithms are insensitive to the noisy silhouette images and other irrelevance interfering.

VII. CONCLUSION

The traditional vector dimension reduction algorithms as PCA and LDA are seeking the special transformation from \mathbf{R}^{n_1} to \mathbf{R}^{n_2} ($n_1 > n_2$). they are using the vector space to express the gait character. The proposed manifold learning based gait recognition method is seeking the transformation from $\mathbf{R}^m \otimes \mathbf{R}^n$ to $\mathbf{R}^l \otimes \mathbf{R}^k$, ($l < m, k < n$) thus the dimension is reduced in the tensor space. This method definitely considers the manifold structure of the image space where the gait data is a manifold in nature.

Construing the neighbor graph modeled the manifold local structure. And the preserved adjacent relation made the algorithm insensitive to noise. The computation is simple since the optimal gait tensor subspace could be obtained by calculating eigenvector. And the matrix for calculating the eigen values is much smaller than that in PCA and LDA. Thus the operation efficiency has more preponderance on the time and space cost.

It focused on the second tensor primarily; for all that the algorithm proposed can also be applied to the high order tensors. This would be seen as an improved method for the tensor subspace analysis besides make the most of the class information so as to gain the gait characters from STE for every individual.

Nevertheless the gait sequences are the nonlinear video manifold essentially which is different to the image manifold. This method proposed here maybe ignores some important inherent structure for classification. The further effort would settle this problem.

ACKNOWLEDGMENT

The authors would like to thank Prof. Y. M. Tian, Dr. T. J. Ma, Dr. S. Chen and Dr. B. Wan, research Inst. of computer peripheral, Xidian University, for the constructive comments and suggestions, which have resulted in a number of improvements in this paper.

REFERENCES

- [1] Aaron F. Bobick and James W. Davis, "the Recognition for human movement using temporal templates", IEEE Trans on PAMI, 2001, vol.23, No.3, pp. 257-267.
- [2] Ju Han and Bir Bhanu, "Individual recognition using gait energy image", IEEE Trans on PAMI, 2006, Vol.28, No.2, pp. 316-322.
- [3] Toby H.W. Lam and Raymond S.T. Lee, "a New representation for human gait recognition: motion silhouettes image (MSI)", ICB 2006, LNCS 3832, pp. 612 - 618.
- [4] Wang Liang, Hu Wei-ming, Tan Tie-niu. "Gait-based human identification". Chinese Journal of Computers, 2003, vol.26, No.3, pp.353-360 (in chinese).
- [5] Heung-II Suk and Bong-Kee Sin, "HMM-based gait recognition with human profiles", SSPR&SPR 2006, LNCS 4109, pp. 596-603.
- [6] Ma Qin-Yong, Nie Dong-Dong, Wang Shen-Kang. "Gait expression and recognition based on primary motion contours". Acta Automatica Sinica, 2008, 0019, pp. 1-7 (in chinese).
- [7] Changhong Chen, Jinmin Liang, Heng Zhao and Haihong Hu. "Gait recognition using Hidden Markov Model". ICNC 2006, Part I, LNCS 4221. Berlin Heidelberg: Springer-Verlag, 2006, pp.399-407.
- [8] S T Roweis, L K Saul. "Nonlinear dimensionality reduction by Locally Linear embedding". Science. 2002, 290(5500): 2323-2326.
- [9] X F He, S C Yan, Y X Hu, et al. "Learning a Locality Preserving Subspace for Visual Recognition". ICCV, 2003.
- [10] X F He, D Cai, P Niyogi. "Tensor Subspace Analysis". Advances in Neural Information Processing System, 2005.
- [11] Yong-zhou Li, Da-yong Luo, Shao-qiang Shao, "Face Recognition Using Tensor Locality Discriminant Projection", ACTA Electronica Sinica, 2008, vol.36, No.10, pp. 2070-2075,
- [12] D Xu, S C Yan, D C Tao, et al. "Marginal Fisher Analysis and Its Variants for Human Gait Recognition and Content-Based Image Retrieval". IEEE Transactions on image Processing, 2007, vol.16, No.11, pp.2811-2821.
- [13] Zi-rong Li, Ming-hui Du, "Orthogonal Marginal Fisher Analysis for Face Recognition", Science Technology and Engineering, 2008, vol.8, No.20, pp. 5589-5592.
- [14] X He, P Niyogi. "Locality Preserving Projections". Advance in Neural Information Processing Systems, Vancouver, Canada, December 2003, Vol.16.
- [15] X He. "Locality Preserving Projections". Chicago, Illinois: University of Chicago of USA, 2005, 12.
- [16] T K Sun, S C Chen. "Locality preserving CCA with applications to data visualization and pose estimation". Image and Vision Computing. 2007, vol.25, No.5, pp. 531-543.
- [17] Zongyi Liu and Sarkar S. "Effect of silhouette quality on hard problem in gait recognition", IEEE Trans on System, Man, and Cybernetics, 2005, vol.35, No.2, pp. 170-183
- [18] M. Belkin and P. Niyogi, "Laplacian eigenmaps and spectral techniques for embedding and clustering", Advances in Neural Information Processing Systems, Vol. 14, 2001.
- [19] Xiaofei He, Deng cai and Partha Niyogi. "Tensor subspace analysis", Advances in Neural Information Processing System18, 2005.
- [20] P Joathon Phillips, Hyeonjoon Moon, Syed A. Rizvi and Patrick J. Rauss. "the FERET evaluation methodology for face-recognition algorithms". IEEE Trans on PAMI, 2000, vol. 22, No.10, pp. 1090-1104



Suyuan Wei is currently working toward the Ph.D. Degree at the research Inst. of computer peripheral, Xidian University, Xian, and China.

Her current research interests include computer vision, pattern recognition and motion image analysis.



Chao Ning is currently a Ph. D. candidate at Xian High Technology Inst., Xian, China.

His current research interests include virtual reality and flow fields analysis.



Youxing Gao, born in 1938. Professor and Ph. D. supervisor of the college of computer at Xidian University, China.

His main research interests include computer graphics, digital image processing, computer vision and pattern recognition.