Multi-scale Principal Component Analysis based Gait Recognition

S.Rohith Sai\textsuperscript{a,\,*}, Rohan Ravi\textsuperscript{b}

\textsuperscript{a}Indian Institute of Technology Ropar, Ropar, 140001, India.
\textsuperscript{b}Indian Institute of Technology Ropar, Ropar, 140001, India.

\*S.Rohith Sai: srohith@iitrpr.ac.in

Abstract

Gait recognition aims to identify individuals by the manner in which they walk. The prime innovation in this paper is a simple yet effective method of modelling the lower limbs using spline curves. To start with, we extract the binary silhouette from the monocular images. The gait cycle is extracted by exploiting the periodic variance of the width vector of the silhouette. Then, the novel gait feature area under the lower limbs, modelled using a cubic spline curve, is computed for each of the silhouettes in a cycle. Later Discrete Cosine Transform (DCT) is applied to the feature vector to create a feature matrix. The method of Multi-scale Principal Component Analysis (MSPCA) is adopted for dimensional reduction of feature matrix containing the area signals. Finally Neuro-fuzzy and the K-NN classifiers are used to classify the final feature vectors. Experimental results on the CASIA datasets A, B show that the best accuracy achieved using Neuro-fuzzy and K-NN classifier is 95\% and 97.1\% respectively.

Keywords: cubic spline curve, DCT, MSPCA, Neuro-fuzzy, K-NN.

1. Introduction

With the increasing demand for visual surveillance systems, human identification at a distance has gained a lot of attention recently. Gait, though a weak biometric, is an attractive feature for human identification at a distance and has gained a lot of interest from computer-vision researchers. Gait is a particular way of moving on foot. Compared to the conventional biometric features such as face, iris, palm print and finger print, Gait has unique features such as being non-contact, non-invasive and perceivable at a distance.

The genesis of the idea of person identification from gait can be traced back to Cutting and Kozlowski’s perception experiments based on light point displays\textsuperscript{(1)}. The first efforts towards recognition from gait in computer vision was attempted by Niyogi and Adelson\textsuperscript{(2)} in the early 1990’s.

In recent years this area of biometrics has become a hot research area in the field of computer vision. It finds its application in a diverse range of systems ranging from simple biometric access control, to demographic analysis of population.

Though Gait has some benefits over its conventional counterparts, it suffers from several drawbacks which make its implementation for recognition systems a tough challenge. For instance Gait analysis is very sensitive to deficient or incomplete segmentation of the subject silhouette. Variations in clothing and footwear, distortions in gait pattern produced by carrying objects or walking speed could make analysis an arduous task. These complexities lead to low recognition rates in the algorithms proposed so far.

Existing methods on Gait recognition can be classified into model based ones and holistic ones. Model based methods model the human body with appropriate geometric curves. Holistic methods extract spatio-temporal and statistical features.

A. Model based:

As stated earlier Model based approaches aim to model human body or motion with geometrical curves. An early attempt at modelling could be seen in\textsuperscript{(3)} and Cunado et al\textsuperscript{(4)}, in which the legs were considered as interlinked
pendulum. Then, phase weighted Fourier magnitude spectrum was used to recognize the Gait signatures, which were derived from frequency components of variations in human thigh inclination. Lee et al. fit ellipses to seven regions of human body and derived magnitude and phase of these moment based region features. Bobick et al. used a structural model to recover static body and stride parameters. Furthermore, statistical methods were used such as, Principal Component Analysis (PCA), and Multiple Discriminant Analysis (MDA) to analyze effective features.

B. Holistic/Model free methods:

The holistic methods attempt to characterize spatial variation of dynamic variables in gait cycle. They analyze the variations in shape and distance vectors in the sequence of images to characterize the gait features. Early efforts at gait recognition adopting holistic approach can be traced back to Niyogi and Adelson, who distinguished different subjects from their spatio-temporal gait patterns obtained from the curve fitted “snake”. Little and Boyd used frequency and phase features from optical flow information of walking figures to differentiate individuals. Chai et al, introduced perpetual shape descriptor to analyze human gait. R. Tanawongsuwan and A. Bobick used time-normalized joint angle trajectories to create gait signatures.

Though a lot of progress has been achieved using the above stated approaches, there is no foolproof method established, which is why the scope of research in this area is diverse.

Motivation

The gait recognition methods proposed so far are sensitive to variations in silhouette shape as well as clothing of the subject. This was one of the motivations for our work. Though a wide variety of model free methods were proposed in the literature, they were all sensitive to covariate features like subject carrying a bag or wearing a coat. This was the prime motivation for our proposed method.

In this paper we have extracted a novel feature adopting model based approach. The method proposed is based on spatial variations of subject’s limbs and width vector of silhouettes over a cycle of frames. This method is robust to changes in shape variation of the subject due to added objects such as a bag or a coat. The variances in the spatio-temporal features extracted are computed and a feature matrix is constructed to describe gait signature of the individual.

The rest of the paper is organized into 5 sections. Section 2 and 3 deal with Approach overview and preprocessing. Section 4 deals with gait feature extraction, MSPCA dimensional reduction. Section 5 deals with experimental results and comparison with recent methods. Section 6 concludes the paper.

2. Approach Overview

The proposed approach can be implemented in the following steps:
1) Silhouette is extracted using background subtraction technique, and preprocessed to remove noise components introduced.
2) The extracted silhouette is then resized by cropping, to create image template with fixed dimensions. A gait cycle is then extracted using width vector as a feature.
3) The proposed feature namely, area under the limbs of the subject, is extracted, DCT is applied on the feature matrix created and MSPCA is adopted for dimensional reduction of area signals extracted.
4) After dimensional reduction, the feature matrix is fed to Neuro-fuzzy and K-NN classifiers for evaluation.

3. Preprocessing

3.1 Silhouette Extraction

Silhouette extraction holds prime importance in gait analysis. It is essential to analyze the value of each pixel in every frame of the sequence. The method of background subtraction is used to acquire the subject of interest. Here the subject should be the only object in motion in the sequence of frames. Background subtraction usually introduces some noise to the binary images. We use algorithms of erosion, dilation, and component labeling in mathematical morphology to filter out image noises and fill small empties.

3.2 Image Template

After background subtraction we can see that the subject occupies a small area of image. To eliminate the redundant boundary around the object that occupies a larger
portion of image, we resize the image by cropping the extra portion and fit the subject into a smaller image template choosing appropriate width and height so that the image is not corrupted. Firstly, height of the human silhouette is chosen as the height of the image and secondly a fixed width is chosen which avoids most of the computational ambiguities. This type of scaling not only reduces computational complexity but also corrects the scale changes due to the variation of object distance from the camera. Similar work can be seen in [9].

4. Gait Feature Extraction

From the gait silhouette sequence obtained, the only cue to identify the gait signature depends on the temporal changes in the silhouette. We propose a novel silhouette modeling method which uses spline curves to model the limbs. The procedure involves finding the coordinates of coxa joint, two knee joints and two ankle joints of each silhouette. The five joints thus found, are used as interpolating points to construct a cubic spline curve. The procedure for finding the joints and constructing the spline curve is enumerated in the following sections.

4.1 Joint Positioning

The novel feature extracted in this paper, the area under the limbs, requires silhouette’s joints as interpolating points. The control points on the curve are the coxa, ankle and knee joints which are obtained by the process below:

a) Coxa Point- The y co-ordinate is at 0.72H from the top of the image. When horizontal scanning is done it leads to the following cases
One Region: The center of the region is taken as the coxa point.
Two regions: This happens if our scanning position is below the actual coxa hence we need to regulate the scanning width 0.165H to find the coxa point

b) Knee Point- A circle with radius 0.245H is drawn with the coxa point as the centre. Two cases arise here
Two Regions: This is the condition of left and right biped bracing. Centre of each region is the corresponding knee joint.
One Region: This is when the left or right knee standing. The human knee is about 0.1H wide, so we choose the point 0.05H left/right from the rightmost/leftmost point as the right/left knee joint.
c) Ankle joint: This is similar to the knee joint. The left and right knees are chosen as the centers of the circles and length of shank 0.246H as radius.

4.2 Area under the spline curve

As observed from a sequence of frames, the area under the limbs has a periodic temporal variance just like width vector of silhouette. This area is found by constructing a spline curve and finding the area under the limbs enclosed by the curve.

For constructing an interpolating curve given a set of points, there are three different possibilities namely, polynomial interpolation, Bézier curves and spline curves. All three methods produce polynomial curves as a linear combination of a set of basis polynomials. Our choice of spline curves is based on their properties which allow us to design complex shapes with lower degree polynomials as compared to the other two methods. In the Fig 3, B-spline curve of degree 3 and Bézier curve of degree 10 are constructed for the same set of control points and it is pretty evident that the Bézier curve still cannot follow the polyline.
Since the degree of the constructed interpolating curve is lower using splines the computational time which is \( O(n^2) \), for the same is reduced considerably, with \( n \) as its degree.

The interpolating spline curve has the human body joints as its control points, namely coxa joint, a pair of knee and ankle coordinates. A polygon was first constructed with the body joints as its vertices, and then a cubic spline curve was constructed with the joints as control points.

A spline is a piecewise-polynomial real function

\[ S : [a,b] \rightarrow \mathbb{R} \]

on an interval \([a, b]\) composed of \( k \) ordered disjoint subintervals \([t_{i-1}, t_i]\) with

\[ a = t_0 < t_1 < \ldots < t_k < t_k = b \]

The restriction of \( S \) to an interval \( i \) is a polynomial

\[ P_i : [t_{i-1}, t_i] \rightarrow \mathbb{R} \]

So that,

\[ S(t) = P_1(t) , \quad t_0 \leq t < t_1 \]
\[ S(t) = P_2(t) , \quad t_1 \leq t < t_2 \]
\[ \ldots \]
\[ S(t) = P_k(t) , \quad t_{k-1} \leq t < t_k \]

The highest order of \( P_i \) is known as the order of spline curve, which in our case is 3. For a spline of order \( n \), \( S \) is required to be continuously differentiable to order \( n-1 \) at the points \( t_i \) for all \( i=1,2,\ldots,k-1 \) and all \( j \in [0,n-1] \)

\[ P_i^{(j)}(t_i) = P_{i+1}^{(j)}(t_{i+1}) \quad (5) \]

In our method of spline curve interpolation of knee joints, we use the \( B \)-form of spline curves which is a weighted sum with the weights as \( B \)-spline functions.

The spline \( f(t) \) is given by

\[ f(t) = \sum_{i=1}^{k} C_i B_{i,n}(t) \quad (6) \]

The function \( B_{i,n} \) is called a \( B \)-spline of degree \( n \) which is given by the recursive formula

\[ B_{i,1}(x) = \begin{cases} 1 & \text{if } t_i \leq x \leq t_{i+1} \\ 0 & \text{otherwise} \end{cases} \quad (7) \]

\[ B_{i,n}(x) = \frac{x-t_i}{t_{i+n-1}-t_i} B_{i,n-1}(x) + \frac{t_{i+n}-x}{t_{i+n}-t_{i+1}} B_{i+1,n-1}(x) \quad (8) \]

Thus for each silhouette image, we obtain the area under the spline curve and for given \( N \) training samples and \( M \) images in each, we create a feature matrix

\[ A = [A_{1,1}, A_{1,2}, \ldots, A_{1,M}; \ldots; A_N, A_{N,2}, \ldots, A_{N,M}] \]

This matrix is considered for further processing, using Discrete Cosine Transform (DCT) to describe the area feature better, followed by dimensional reduction using MSPCA.

4.3 Multi-scale principal component analysis

The dimensionality of the feature matrix containing the area signals is very large and contains redundant information so, we adopt the method of Multi scale principal component analysis (MSPCA) to find transformation for dimensionality reduction. MSPCA was first proposed by Bakshi \(^{(10)} \), for statistical process monitoring. Multi scale principal component analysis (MSPCA) combines the ability of PCA to de-correlate the variables by extracting a linear relationship, with that of wavelet analysis to extract deterministic features. MSPCA implements PCA to wavelet coefficients at each scale to filter the unwanted components. The essence of MSPCA is enumerated in Fig 5 and 6.
Fig 5: Multi-Scale Principal Component Analysis Algorithm.

Where $\gamma = WY^T$ is the wavelet transform coefficient matrix of $Y$.

$$\gamma^T = [Y_D^{(1)}^T, Y_D^{(2)}^T, \ldots, Y_D^{(L)}^T, Y_V^{(L)}^T], \quad (9)$$

$W$ is the Discrete wavelet transform (DWT) operator,

$$Y_D^{(j)} = YG_j^T, \quad j=1,2,\ldots,L \quad (11)$$

Implementing Inverse DWT (IDWT), $\hat{Y}$ can be reconstructed via (14).

Denote $\{j_1,j_2, \ldots, j_S\} = \{ j | \tau_j \neq 0, j=1,2,\ldots,L \}$, \quad (12)

$$\Omega = \{1, 2, \ldots, L+1\} \setminus \{j_1,j_2, \ldots, j_S\}, \quad \text{then} \quad (13)$$

$$\hat{Y} = Y - \sum_{i \in \Omega} (YG_i^T G_i) \quad (14)$$

Where $\tau_j$ is defined as in (10). Traditional PCA is then applied on $\hat{Y}$, the wavelet coefficients matrix, to acquire the final feature matrix which is fed to the classifiers for recognition in the following subsection.

5. Recognition

After the extraction of gait features, followed by dimensional reduction classification is done using two different classifiers namely, KNN and Neuro-fuzzy. First we evaluate the proposed method using Neuro-Fuzzy classifier as it is the main method of classification adopted. Then we compare the achieved results with the results of K-NN classifier. The gait feature matrix extracted using the proposed method is used to train the classifiers.

5.1 Gait database

In our experiments, we use the CASIA Gait Database which is one of the largest gait databases in gait-research community currently. We have tested the algorithm on the CASIA Gait database due to its completeness and wide availability.

CASIA Dataset – A

This set consists of 20 subjects, each subject having 12 samples captured at 3 different angles (0°, 45°, 90°). 4 image sequences captured at 0° are chosen from each subject. Out of the 4 sequences, 2 of them have subject walking towards the left and the other 2 have subject waking towards the right.

CASIA Dataset - B

The database consists of 124 subjects (93 males and 31 females) captured from 11 view angles (ranging from 0 to 180 degrees, with view angle interval of 18). The frame size is 320*240 pixels, and the frame rate is 25 fps. There are 10 walking sequences for each subject per view. We use gait sequences numbered from 001 to 124 (subject ID, i.e., 124 subjects) of view angle 90 degrees in Dataset B to carry out our experiments.

Out of the 10 samples chosen from each subject, 2 samples have images with subject carrying a bag, and 2 have subject wearing a coat.

5.2 Experimental Results

The proposed feature matrix is transformed by applying the Discrete Cosine Transform (DCT) and then reduced in dimensionality using the proposed MSPCA method. The first implementation of this method is on CASIA dataset-A, considering 4 samples for each of the 20
subjects. Of the 4 samples we chose, 3 samples are fed to the classifier for training and 1 sample is put aside for testing. Cumulative match scores (CMS) are used to assess the performance quantitatively. The CMS value \( \delta \) corresponding to rank \( r \) indicates a fraction \( 100 \times \delta \% \) of probes whose top \( r \) matches must include the real identity matches.

### Table 1: Accuracies of proposed methods on CASIA A database using KNN

<table>
<thead>
<tr>
<th>Recognition method</th>
<th>CMS (%)</th>
<th>CMS (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rank 1</td>
<td>Rank 5</td>
</tr>
<tr>
<td>K-NN</td>
<td>50</td>
<td>70</td>
</tr>
<tr>
<td>DCT+ K-NN</td>
<td>65</td>
<td>85</td>
</tr>
<tr>
<td>MSPCA+ K-NN</td>
<td>90</td>
<td>100</td>
</tr>
<tr>
<td>DCT+MSPCA+ K-NN (Proposed method)</td>
<td>95</td>
<td>100</td>
</tr>
</tbody>
</table>

Unlike the Neuro-fuzzy classifier that uses membership functions extracted from the data set describing the system, K-NN applies Euclidean distances as the measurement parameter in classifying the data. The test results of K-NN classifier are enumerated in the table 1.

### Table 2: Accuracies of proposed methods on CASIA A database using Neuro-Fuzzy classifier

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neuro-fuzzy</td>
<td>62.3</td>
</tr>
<tr>
<td>DCT+Neuro-fuzzy</td>
<td>70</td>
</tr>
<tr>
<td>MSPCA+Neuro-fuzzy</td>
<td>81.9</td>
</tr>
<tr>
<td>DCT+MSPCA+Neuro-fuzzy (Proposed method)</td>
<td>95</td>
</tr>
</tbody>
</table>

Four different methods of testing are adopted to compute the accuracies: directly using Neuro-fuzzy classifier on the feature matrix and the cosine transform coefficient matrix; using Neuro-fuzzy classifier on the feature matrix and the cosine transform coefficient matrix after dimensional reduction using the proposed MSPCA method. The results are as shown in table 2.

In order to test the robustness of our proposed feature extraction method, we also test our algorithm’s performance on K-NN classifier. We adopt similar strategies of testing as in the case of Neuro-fuzzy classifier.

The above results demonstrate the robustness of our method to changes in direction of motion of the subject in dataset-A. Our method of modeling spline lower limbs using spline curves was found to be very effective, even though the direction of subject’s motion changed in two samples. The best accuracy of 95% retained is promising and the method itself is quite feasible for recognition.

Two different strategies are used to test the proposed algorithm on CASIA dataset-B: Discrete Cosine Transform is applied on the feature matrix and the coefficient matrix acquired undergoes training and testing using Neuro-fuzzy classifier; Dimensional reduction of cosine transform coefficient matrix is done by MSPCA and then fed to the Neuro-fuzzy classifier.

The results presented on dataset-B in table 3, show convincing results even after considering covariate features, in which subject is either carrying a bag or wearing a bulky coat.

### Table 3: Accuracies of proposed methods on CASIA B Dataset

<table>
<thead>
<tr>
<th>Recognition method</th>
<th>Accuracy(%) -NW</th>
<th>Accuracy(%) -CB</th>
<th>Accuracy(%) -WC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Set 5</td>
<td>Set 6</td>
<td>Set 7</td>
</tr>
<tr>
<td>DCT+MSPCA+Neuro-fuzzy (Proposed method)</td>
<td>83.33</td>
<td>87.25</td>
<td>91.2</td>
</tr>
<tr>
<td>DCT+MSPCA+K-NN CCR(%) - Rank 1</td>
<td>95.1</td>
<td>96.07</td>
<td>97.1</td>
</tr>
<tr>
<td>DCT+MSPCA+K-NN CCR(%) - Rank 5</td>
<td>98.04</td>
<td>98.04</td>
<td>100</td>
</tr>
</tbody>
</table>

(NW-normal walk; CB-carrying bag; WC-wearing coat)
This shows that our method is robust to these covariate features and the best accuracy of 91.2% acquired is in itself, quite feasible for recognition, considering the fact that covariate features are taken into account.

The consistent CMS for all the 6 sets of a subject shows that our method is robust to covariate features of CASIA dataset-B. The best accuracy of 97.1% CMS was obtained for set 7 in which the subject carries a bag, strengthening the claim of our algorithm’s robustness to covariate features. The feature, area under the limbs, chosen is clearly insensitive to subject wearing a bulky coat or carrying a bag which makes it much more effective and reliable for recognition.

Su-li et al [11], proposed a feature extraction method based on Fuzzy principal component analysis. They use the CASIA database-A with 20 subjects under consideration. In [12], chen et al propose a method based on Frame difference energy image. They performed experiments on CMU Mobo gait database and the CASIA dataset B with 100 subjects under consideration. Note that the numerical accuracies from these two techniques are obtained from CMS curves.

It is pretty evident from the table that the proposed Algorithm outperforms the other Algorithms in terms of Cumulative match scores (CMS). The best CMS of 97.1% was obtained over CASIA data set B, with 124 subjects under consideration. Our experimental results show that the method of MSPCA performs fairly good even with complications like carrying of covariate objects involved which is our key interest.

### Table 4: Experimental Results Compared with Other Algorithms

<table>
<thead>
<tr>
<th>Recognition method</th>
<th>Best CCR (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Su-li[11]</td>
<td>89.7%</td>
</tr>
<tr>
<td>Chen [12]</td>
<td>95.2%</td>
</tr>
<tr>
<td>Proposed method</td>
<td>97.1%</td>
</tr>
</tbody>
</table>

Su-li et al [11], proposed a feature extraction method based on Fuzzy principal component analysis. They use the CASIA database-A with 20 subjects under consideration. In [12], chen et al propose a method based on Frame difference energy image. They performed experiments on CMU Mobo gait database and the CASIA dataset B with 100 subjects under consideration. Note that the numerical accuracies from these two techniques are obtained from CMS curves.

6. Conclusion

In this paper, we propose a novel and simple method for gait recognition based on modeling of the limbs using spline curves. The Area signals are extracted after preprocessing. By MSPCA the components of the feature matrix are projected into a lower dimension space. Neuro-Fuzzy and K-NN classifiers are used for classification. MSPCA retains the information of original data better as compared to the traditional PCA even when data sequence changes over time or frequency.

The results demonstrate that our method is insensitive to covariate features like subject’s walking direction, subject carrying a bag or wearing a thick coat. Reducing the sensitivity of recognition to covariate features was the prime concern of our method. Reasonable results acquired on a large database like CASIA dataset B with covariate features ascertain the feasibility of our method. We wish to investigate the complications involved with gait recognition.
when the walking speed of the subject is variable in our future work.

**References**


(9) Yanmei Chai, Jinchang Ren, Rongchun Zhao, Jingpin Jia, Automatic gait recognition using dynamic variance features, Proceedings of 7th International Conference on Face and Gesture Recognition (FGR’06).


(11) Su-li XU, Qian-jin ZHANG. Gait Recognition using Fuzzy Principal Component Analysis, 2010 IEEE.