Jurnal Teknologi, 46(D) Jun 2007: 35–44

PCA–BASED HUMAN POSTURE CLASSIFICATION

NOORITAWATI MD TAHIR[1*], AINI HUSSAIN[2],
SALINA ABDUL SAMAD[3] & HAFIZAH HUSAIN[4]

Abstract. This paper outlines a mechanism for human body posture classification based on various combination of eigenspace transform, which we named as eigenposture, and using Multilayer Perceptron (MLP) as classifier. We apply principal component transformation to extract the features from human shape silhouettes. Combinations of the extracted features were used to classify the posture of standing and non-standing based on the human shape obtained from the segmentation process. We experiment by using various combinations of eigenvectors as input representations for classification purpose. Results showed that the second and fourth eigenpostures combination gives reasonably good result with 98% correct classification of human posture and can be adopted as the optimal choice of input for classification using a minimal combination.

Keywords: Principal component analysis (PCA), eigenvectors, classification, artificial neural network, human posture

Abstrak. Kertas kerja ini membentangkan suatu mekanisme untuk pengelasan susuk tubuh manusia berdasarkan kombinasi pelbagai jelmaan ruang eigen yang dinamakan sebagai *eigenposture* dan *Multilayer Perceptron* (MLP) sebagai pengelas. Penjelmaan komponen utama telah digunakan untuk menyari sifat pada bayang-bayang bentuk badan manusia. Gabungan sarian sifat ini digunakan untuk pengelasan susuk tubuh manusia sebagai berdiri atau sebaliknya berasaskan bentuk badan yang diperoleh selepas proses peruasan. Uji kaji telah dijalankan dengan mengubah bilangan vektor eigen yang dijadikan perwakilan untuk tujuan pengelasan. Keputusan yang diperoleh menunjukkan gabungan *eigenposture* kedua dan keempat memberi keputusan pengelasan bentuk badan manusia yang agak baik iaitu 98% dan boleh dijadikan sebagai pilihan optimal masukan bagi tujuan pengelasan menggunakan bilangan input minima.

Kata kunci: Analisa komponen utama, vektor eigen, pengelasan, rangkaian neural tiruan, susuk tubuh manusia

1.0 INTRODUCTION

One common machine vision application is to teach a computer to discriminate some dataset automatically, to save the man-hours or boredom attributed to these tasks. Classification of human posture is a very challenging problem. The importance of human posture classification is evident by the increasing requirement of machines that are able to interact intelligently and effortlessly with a human inhabited environment.

[1,2,3&4] Signal Processing Research Group, Dept. of Electrical, Electronic and Systems, Faculty of Engineering, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia
[*] Corresponding author: Email: norita@vlsi.eng.ukm.my

Past attempts reported in literature to recognise human actions by machines require intrusive devices that limit the scope of their applications to situations where people specifically intend to communicate with computers [1, 2]. Some major early works involved the use of Moving Light Display (MLD) on subjects in a darkened room [3] and Structure from Motion (SFM) techniques [4] where a 3 dimensional model of the person is reconstructed to recognise human action. Although useful, MLD was a very intrusive experiment while SFM technique was more complex and computationally expensive.

This paper reports the preliminary research conducted to identify simple human posture using computationally simple techniques that does not require the use of intrusive devices. This is the first step towards developing applications in real time surveillance, pedestrian detection and gait recognition. This will take the capability of machines into 'understanding human' domain. To improve machine capabilities in real time applications such as surveillance, activity recognition, human interaction with machines, it is desirable to represent human shapes from very high dimensional space to low dimensional space without loss of true shape characteristics. In this work, we restricted the scope of study to discriminate between standing and non-standing postures only. The eigenspace transform was chosen as the desired method for discrimination. To demonstrate that we have created a good statistical model of human posture derived from low-level features, we reconstruct the original image. Utilization of the eigenspace, which we named 'eigenposture', required preprocessing of the images. In this work, we rigorously defined the eigenposture used and provide a brief review of previous implementation adopting similar method. These are given in Section 2, followed by a discussion on the methodology, which is given in Section 3. Results are presented and discussed in Section 4 and finally, in Section 5, we conclude our findings.

2.0 PCA-BASED HUMAN POSTURE CLASSIFICATION SYSTEM

Over the last several years, numerous algorithms have been developed based on Principal Component Analysis (PCA). PCA is popular because it is a natural dimensionality reduction method, easy to implement and achieves reasonable performance levels. It is the bases for algorithms and serves as a computational model [5-10]. The PCA algorithm has been applied in broad spectrum of studies including face detection [5], face recognition [5-6], and gesture classification [9]. PCA is a statistical method for reducing the dimensionality of high dimensional data, where the data are represented as a vector. In designing an algorithm relating to PCA, a number of critical design issues have to be addressed. Each of these design decisions has an impact on the overall performance of the algorithms.

2.1 Previous Work Using Principal Component Analysis (PCA)

The Principal Component Analysis (PCA) is a useful statistical technique that has found applications in fields such as recognition, classification and image data compression. It is also a common technique in extracting features from data in a high dimensional space. This quality makes it an interesting tool for our study. It is a systematic method to reduce data dimensionality of the input space by projecting the data from a correlated high-dimensional space to an uncorrelated low-dimensional space. Turk and Pentland [5] applied PCA for faces recognition. Ji and Yang [7] built an 'eigeneye' feature space using PCA that captured relationship between 3D face pose and the geometric properties of the pupils. The 'eigeneye' space is then used for 3D face pose classification. Results [7] showed that the technique could estimate face pose in real time and produce good results for subjects closer to the camera. Ohba *et al.* [8] applied PCA to classify several facial expressions such as anger, normal, surprise and smile. Original image of each category is projected onto a facial expression space and only the first three eigenvectors are used for classification of facial expression. The ability of PCA is also employed by Algorri and Escobar [9] for facial gesture recognition. They used the eigenspace method to build the facial gesture space and later used it for image reconstruction during video conferencing. Ozer *et al.* [10] applied PCA to detect human activities, namely walking, running and kicking.

2.2 Eigenposture Approach

PCA uses the eigenvalues and eigenvectors generated by the correlation matrix to rotate the original data coordinates along the direction of maximum variance. When ordering the eigenvalues and their corresponding eigenvectors, or principal components, in decreasing order of magnitude, the first principal component accounts for the largest variance in the original training data set, the second orthogonal principal component for the largest remaining variance and so forth. The full set of principal components is as large as the original set of variables. However, it is common for the sum of the variances of the first few principal components also known as eigenvalues to exceed two third of the total variance of the original data. Several techniques from numerical analysis have been suggested to efficiently compute principal components [5-10]. This study is based on results from the matrix theory, the Singular Value decomposition (SVD), which is closely related to PCA [11]. For the training set matrix X, of dimension $N \times p$, and rank r, it can be rewritten using SVD as:

$$X = U * S * V^T \tag{1}$$

where U is an orthogonal $N \times r$ matrix, V^T is an orthogonal $p \times r$ matrix with the eigenvectors $(e_1, e_2, \dots e_r)$, and S is $r \times r$ diagonal matrix containing the square roots of the eigenvalues of the correlation matrix X^TX, and hence the variances of the principal components.

The r eigenvectors, i.e. principal components of matrix V, form an orthogonal basis that spans a new vector space, called the feature-space. Thus, each vector $x_{i,j}$ can be projected to a single point in this r-dimensional feature space. However, according to the theory of PCA for highly correlated data, each training set vector can be approximated by taking only the first few k, where $k \leq r$, Principal Components $(e_1, e_2, \ldots e_k)$.

By linearly transforming the images into eigenspace, we project the images into a new N dimensional space, which exhibits the properties of the samples most clearly along the coordinate axes. The most significant features information of the images will be in the first few principal components. Principal component analysis has been applied in many fields [11]. Pentland *et al.* [5 - 6] took the same approach to extract features from faces. Here, we apply it to human shape silhouette images.

3.0 METHODOLOGY

Computational models of human posture recognition, in particular, are interesting because they can contribute to not only theoretical insights but also practical applications. Computers that recognise human shapes could be applied to a wide variety of problems, including intruders alert, security systems, sport-science studies and medical gait applications. Our approach treats the human shape problem as a 2-D classification problem. The system functions by projecting human posture images onto a feature space that spans the significant variations among known human posture images. The significant features of the postures set, which we called as 'eigenposture' because they are the eigenvectors (i.e. principal components), do not necessarily correspond to feature such as head, body, or torso. The projection operation characterised a human shape by a weighted sum of the eigenposture features, and to recognise a particular posture it is only necessary to compare these weights to those of known postures. Some particular advantages of our approach are that it provides the ability to learn and later classify new postures in an unsupervised manner, and that it is easy to implement using neural network architecture. In this paper, we present a PCA-based posture classification algorithm. In the experiment, we select and form various combinations of eigenvectors and use them as inputs to the MLP for classification purpose.

3.1 System Description

Figure 1 depicts an overview of the overall system where the basic structure is outlined. It consists of the following steps: Segmentation, feature extraction and finally, classification.

The segmentation stage extracts the silhouette of a person using the binary image extraction process, which consists of background differencing followed by thresholding to obtain a binary mask of the foreground region. Since the camera is stationary, the detection of object can be achieved by comparing each new frame with the representation of the background scene. This process is called background differencing and the scene representation is called the background model. In order to remove noise, median filtering and morphological operations are used. Next, the feature extraction component function is applied by projecting the training images onto a feature space that spans the significant variations among known images. The significant features, which we termed as 'eigenposture', are the eigenvectors (principal components) of the set of images. It then serves as inputs to the MLP classifiers that we have chosen in this study.

In a simple term, relevant information is extracted from human silhouettes, encoded it as efficiently as possible and compared it to a database of models encoded similarly. In order to extract the information contained in an image of a human shape is to somehow capture the variation in a collection of human shapes, independent of any judgement of features, and use this information to encode and compare new human shape and classify it. In mathematical terms, we find the principal components of the distribution of human shapes, or the eigenvectors of the covariance matrix of the set of human shape images, treating an image as a vector in a very high dimensional space. These eigenvectors, which we named as eigenposture, can be thought of as a set of features that together characterize the variation between human shape images. Each pixel contributes more or less to each eigenvector. Figure 2 depicts selected sample image of human postures used in this study.

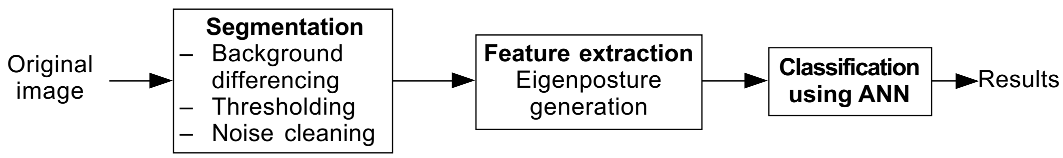


Figure 1 Overview of the overall system

Figure 2 Some human shape silhouettes used as the training data

The approach used in this work includes the following steps:

(i) Build feature space: Calculate the eigenvectors from the covariance matrix of the training set, keeping only the first k eigenvectors that correspond to the highest eigenvalues. These k eigenvectors define the feature space, or the eigen space.

(ii) Project a new human shape silhouette image into feature space: Calculate a set of weights based on the k-dimensional feature space and the new image by projecting its image onto the eigen space. This is done by subtracting the mean of the training images and some simple projection. Assume P is the new human shape image to be projected into the eigen space, we need to first center the image by subtracting the mean of the training images (X).

$$w = V^T (P\text{-mean } (X)) \tag{2}$$

The vector w, which is also an N dimensional vector, can be seen as the new encoding of the image in the eigen space. V is the corresponding orthogonal $p \times r$ matrix with the eigenvectors (e_1, e_2, ... e_r) from the training images (X).

(iii) To test our method, we use the eigenvectors as training inputs to the classifier.

3.2 Classification Task

Classification is a pattern recognition (PR) problem of assigning an object to a class. Thus, the output of the PR system is an integer label. The task of the classifier is to partition the feature space into class-labeled decision regions. Multilayer Perceptron (MLP) is chosen as our classifier in this study. A Multilayer Perceptron (MLP) neural network is an extremely popular and widely documented architecture. It is a good tool for classification purposes. It can approximate almost any regularity between its input and output. The neural network weights are adjusted by supervised training procedure called back propagation. Back propagation is a kind of gradient descent method, which searches an acceptable local minimum in the neural network weight space in order to achieve minimal error. MLP comprises an input, output and one or more hidden layers. With the exception of the input layer, all layers compute their output with a weighted output formula, an optional bias and an activation function. For this study, only a single hidden layer was used.

4.0 EXPERIMENTS AND RESULTS

4.1 Variation of Eigenpostures in the Classification Task

In order to develop a classification system, we first collect a data set of 100 images of various image size of human postures (50 each for standing and non-standing) as our training data to generate the eigenpostures and another 200 images of human postures (100 each for standing and non-standing) for testing. We need to resize all the images in the training and testing database to a standard size. In this work, each image has $m \times n$ pixels, and we reshaped them into a $1 \times mn$ vector. Then, the eigenvectors and eigenvalues are computed according to (1). In the end, we kept the first six eigenvectors and let k =1, 2, ...6 which also represents the most significant eigenvectors and obtained matrix V of size $mn \times k$. In our experiment, we used different values of k for classification. MLP was trained to classify the human posture of standing and non-standing. A combination of these most six significant eigenpostures of the training images serves as inputs to each classifier. For each experiment, a combination of six eigenpostures to minimum of two are selected as in Tables 1 and 2 respectively.

Table 1 Classification results using three or more eigenpostures combination

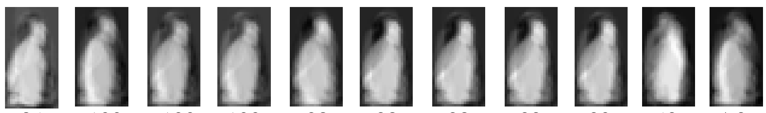
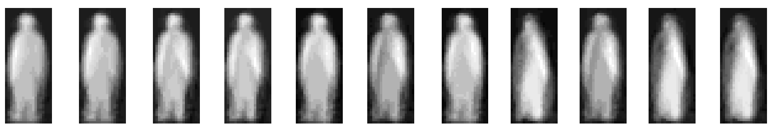
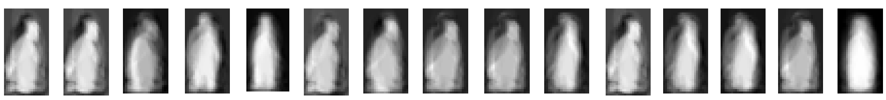
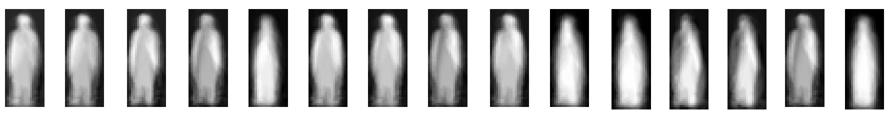
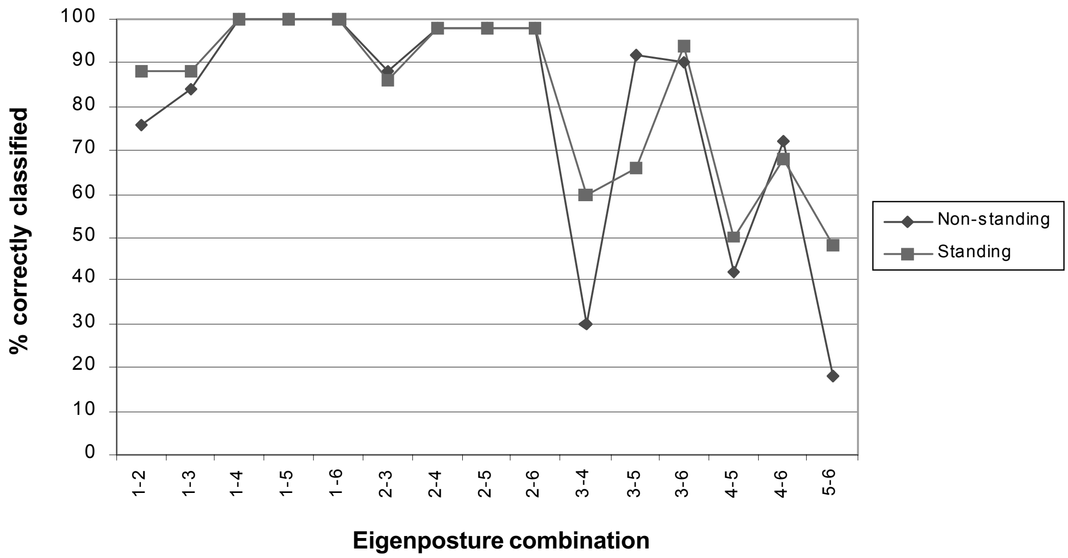
Case no.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
Eigenpostures combination	1-3	1-4	1-5	1-6	2-4	2-5	2-6	3-5	3-6	4-5	4-6
Reconstructed non-standing posture											
Classification accuracy	84	100	100	100	98	98	98	92	90	42	72
Reconstructed standing posture											
Classification accuracy (%)	88	100	100	100	98	98	98	66	94	50	68

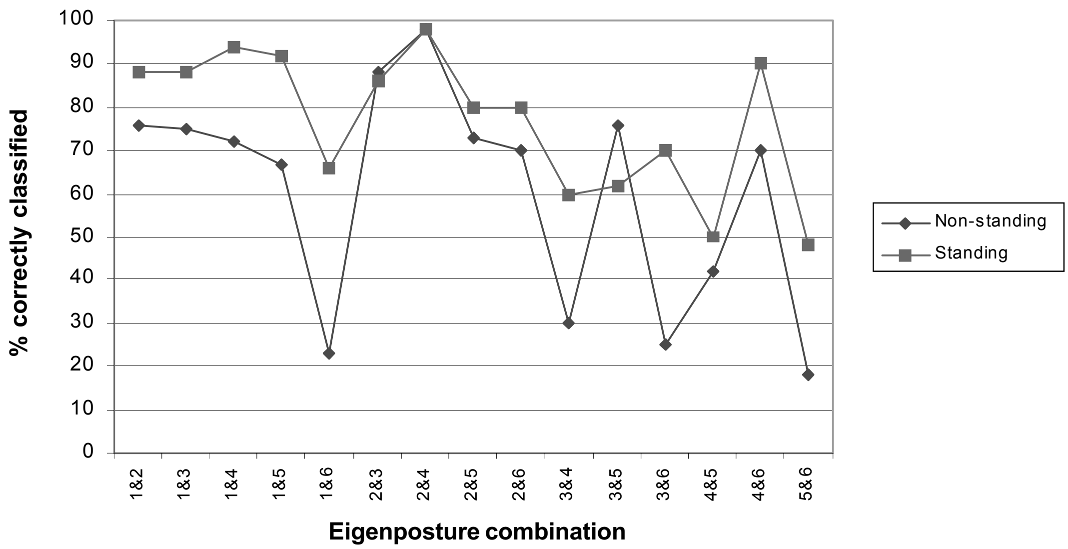
Table 2 Classification results using eigenpostures pair combination

Case no.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
Eigenpostures combination	1&2	1&3	1&4	1&5	1&6	2&3	2&4	2&5	2&6	3&4	3&5	3&6	4&5	4&6	5&6
Reconstructed non-standing posture															
Classification accuracy	78	75	72	67	23	88	98	73	70	30	76	25	42	70	18
Reconstructed standing posture															
Classification accuracy (%)	88	88	94	92	66	86	98	80	80	60	62	70	50	90	48

These various combinations of eigenpostures verify the performance of classification as in Tables 1 and 2. The classification performance using MLP is as tabulated in Tables 1 and 2 and the plots performance as in Figure 3. As shown in Figure 3(a), for



(a) Classification accuracy according to eigenposture combination of three or more



(b) Classification accuracy according to eigenposture pair combination

Figure 3 Classification performances both postures, using only the first four eigenpostures (1- 4) gives perfect classification results for unseen images. Omitting the first eigenposture and using eigenpostures 2, 3 and 4 yielded almost perfect classification (98%). As in Figure 3(b), another interesting performance is the combination of second and forth eigenpostures gave equal classification rate as using combination of eigenpostures 2, 3 and 4 for both standing and non-standing.

5.0 CONCLUSION

A method for human body posture classification based on eigenvectors analysis is presented. Its potential use for a simple task of classifying two main human postures has been demonstrated but its ability to classify more complex tasks that include other postures such as bending, laying and carrying an object has yet to be quantified. As can be seen from the experimental results, eigenspace technique can be used for human posture classification with high degree of accuracy. Results showed that 98% of the unseen data has been correctly classified by using only two eigenposture which are the second and forth eigenpostures as inputs to the classifier. These two eigenpostures can be adopted as the optimal choice of inputs to the MLP classifier. Thus, this suggests that the eigenspace technique can be put into practice for posture recognition which can lead to a wide variety of applications such as security systems, intruder's alertness, gait analysis, action recognition, human computer interaction, action recognition for surveillance applications and tracking techniques for video coding and image displays.

ACKNOWLEDGEMENTS

This work was supported by MOSTI under the IRPA Grant No: 03-02-02-0017-SR0003/ 07-03. The authors also acknowledge UiTM for the UiTM-JPA SLAB scholarship awards.

REFERENCES

[1] Granum, E. and T. B. Moeslund. 2001. A Survey of Computer Vision-Based Human Motion Capture. *Computer Vision and Image Understanding*. 81(3): 231-268.

[2] Aggarwal, J. K. and Q. Cai. 1999. Human Motion Analysis: A Review. *Computer Vision and Image Understanding*. 73(3): 428-440.

[3] Lakany, H. M., A. Birbilis, and G. M. Hayes. 1996. Recognising Walkers Using Moving Light Displays. Proceedings of AVBPA. UK. 111-118.

[4] Akita, K. 1984. Image Sequence Analysis of Real World Human Motion. *Pattern Recognition*. 17(1): 73-83.

[5] Turk, M. and A. Pentland. 1991. Eigenfaces for Recognition. *Journal of Cognitive Neuroscience*. 3(1): 71-86.

[6] Turk, M. 2001. A Random Walk through Eigenspace. Special Issue on Machine Vision Applications. *IEICE Transaction Information and System*. E84-D(12): 1586-1595.

[7] Ji, Q. and X. Yang. 2002. Real Time 3D Face Pose Discrimination Based on Active IR Illumination. International Conference on Pattern Recognition. Japan. 4: 310-313.

[8] Ohba, K., G. Clary, and T. Tsukada. 1998. Facial Expression Communication with FES. International Conference on Pattern Recognition. Japan. 1378-1381.

[9] Algorri, M. E. and A. Escobar. 2004. Facial Gesture Recognition for Interactive Applications. IEEE Proceedings of Fifth Mexican International Conference in Computer Science. Mexico. 185-195.

[10] Ozer, B. and W. Wolf. 2002. Real Time Posture and Activity Recognition. IEEE Proceedings on Motion and Video Computing. 133-138.

[11] Smith, L. 2002. A Tutorial on Principal Components Analysis. Cornell University. USA. http://kybele.psych.cornell.edu/%7Eedelman/Psych-465-Spring-2003/PCA-tutorial.pdf (Accessed on 06/01/2005).

[12] Zhang, G. P. 2000. Neural Networks for Classifications: A Survey. *IEEE Transaction on Systems, Man and Cybernetics*. 30(4): 451-462.