Applying Image Pre-processing Techniques for Appearance-Based Human Posture Recognition: An Experimental Analysis

M. Masudur Rahman and Seiji Ishikawa

Department of Control Engineering, Kyushu Institute of Technology,
Sensuicho 1-1, Tobata, Kitakyushu 804-8550, Japan
rahman@ss10.cntl.kyutech.ac.jp

Abstract. This paper investigates that loose clothing such as wearing dresses and human body shapes create individual eigenspaces and, as a result, conventional appearance-based method cannot be effective for recognizing human body postures. We introduce *a dress effect* due to loose clothing and *a figure effect* due to various human body shapes in this particular study. This study particularly proposes an image pre-processing by '*Laplacian of Gaussian (LoG)*' filter over input images and a '*mean posture matrix*' for creating an eigenspace in order to overcome the preceding effects. We have tested the proposed approach on various dress environments and body shapes, and robustness of the method has been demonstrated.

1 Introduction

Due to inexpensive mathematical computations, an appearance-based eigenspace (abbr. as ES hereafter) technique has many applications, e.g., human computer interaction, visually mediated interaction and automated visual surveillance. In the past years, we have seen an extensive use of the ES method on capturing the appearances of objects and human faces under various conditions [1-6, 13]. It should be mentioned that the appearance-based ES method was firstly proposed by Murase and Nayar [7] for recognizing 3D object's poses from 2D images. However, PCA algorithm was successfully implemented for human face recognition by Sirovich and Kirby [1]. Besides, several other models have been proposed for the underlying objectives, e.g., separating styles and contents using bilinear models [8], human face detection using one-example views [9], Bayesian rules [10], body pose detection using specialized mapping [11], and spatio-temporal correlation [12].

However, we concentrate on an unexplored area of research, i.e., human body posture recognition using an ES technique which has many expected practical applications such as security. In case of employing ES technique for human body posture recognition, we have found that loose clothing such as wearing clothes and various body shapes have some undesirable effects on shaping the pattern of eigenspaces and we need to overcome these problems for successful representation and recognition of human postures. We introduce notions of *a dress effect* due to loose

G.I. Webb and Xinghuo Yu (Eds.): AI 2004, LNAI 3339, pp. 152–159, 2004.

clothing and *a figure effect* due to various human body shapes in this particular study. We employ image pre-processing by Laplacian of Gaussian (LoG) filter for reducing the dress effect, and a mean posture matrix from some selected posture sets for avoiding the figure effect. This mean posture matrix is used for generating the basic ES. The basic ES is responsible for further recognizing unfamiliar but similar postures. We have succesfully overcome the preceding effects by the proposed methods for human body postures recognition and tentative experimental proofs have been demonstrated in this particular paper.

2 Proposed Approach

2.1 Image Pre-processing

In order to employ an eigenspace in human body posture recognition, posture representation should have generality, i.e., it should solely depend on the posture change and not on the person or dress change. If one employs the original eigenspace technique [7], however, respective eigenspace is inevitably generated each time the person changes his/her dress. This is because it employs gray images for the generation of an eigenspace. We employ blurred edge images of given images to obtain a solely apperance-dependent eigenspace.

A blurred edge image $E(x,y)$ of the original image I(x,y) is defined by

$$E(x, y) = D^2(G * I(x, y)) \tag{1}$$

G is the gaussian distribution for reducing the texture effect and the resultant image is differentiated by a Laplacian operator D^2. In the proposed technique, the blurred edge images provided by Eq.(1) are employed for generating an eigenspace.

2.2 Computing a Mean Posture Matrix and Proposed Eigenspace

Let us take a blurred edge image $E(x,y)$ and take P successive sampled images $\boldsymbol{x}_p$ (p=1,2,…,P). The sampled image $\boldsymbol{x}_p$ having $M_0 \times N_0$ pixel size is converted into a column vector of the form

$$\boldsymbol{x}_p = \left(x_{1p}, x_{2p}, ..., x_{N,p}\right)^{\mathrm{T}}. \tag{2}$$

by arranging pixels in a raster scan manner. Here $N \equiv M_0 \times N_0$. Superscript 'T' denotes transpose of a vector or a matrix.

If person h involves to make respective posture sets, a matrix X^h containing P columns and N rows can be denoted by

$$X^h = \left(\boldsymbol{x}_1^h,\ \boldsymbol{x}_2^h, ...,\ \boldsymbol{x}_P^h\right). \tag{3}$$

Here $h = 1, 2, ..., H$. Taking a particular posture set X^h, an ES can be produced and respective postures are represented in the produced eigenspace. For H humans, the posture curves (graphical representation of ES) corresponding to respective persons should ideally coincide with each other in the ES, which is not the case in practice. Therefore a mean expression of the postures is taken into account.

A mean posture matrix $\overline{X}$ is defined in the following way to obtain a basic ES;

$$\overline{X} = \frac{1}{H}\sum_{h=1}^{H} X^h = \left(\overline{\boldsymbol{x}}_1, \quad \overline{\boldsymbol{x}}_2, \quad \ldots, \quad \overline{\boldsymbol{x}}_P\right), \tag{4}$$

where

$$\overline{\boldsymbol{x}}_p = \frac{1}{H}\sum_{h=1}^{H} \boldsymbol{x}_p^h, \tag{5}$$

which is called a mean image. A mean posture set $\overline{X}$ is a set of average images.

We define a covariance matrix C as follows;

$$C = \overline{X}\overline{X}^{\mathrm{T}} \tag{6}$$

and determine eigenvalues λ_i with its corresponding eigenvectors $\boldsymbol{e}_i$ of the covariance matrix C using an eigen equation $C\boldsymbol{e} = \lambda\boldsymbol{e}$. The N dimensional space defined by all the eigenvectors of matrix C is compressed via PCA algorithm by choosing only k (k is an integer satisfying $k << N$) eigenvectors $\boldsymbol{e}_i$ (i=1,2,…,k) corresponding to the largest k eigenvalues to make an ES.

Once we determine the selected eigenvectors employing the mean posture matrix, the successive procedures, i.e., creating eigenspace and posture recognition technique can be found in the literature [7]. The developed ES having edge images and mean posture matrix is defined here as the proposed or basic ES.

3 Experimental Results

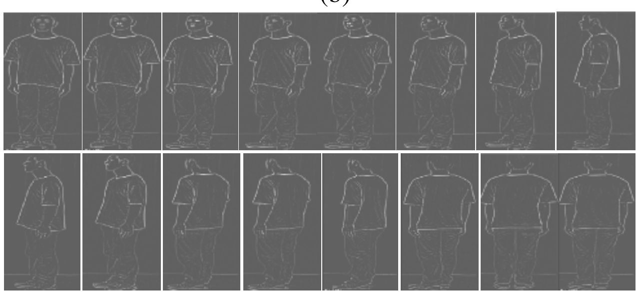
3.1 Effect of the Blurred Edge Images

In the performed experiment, a setup is conducted using a camera and 30 different human models (H=30) with their presently worn clothes including their different body shapes. A person is asked to stand in front of a fixed video camera and to make a slow turn. The video motion is sampled approximately every 10 degrees yielding 36 (=P) images of different postures. The original sampled image is reduced to a 32×32 pixels image for the sake of memory efficiency. Considerations of occlusion and background issues are out of scope in this particular study. Fig. 1a shows 12 human models out of 30 used in the experiment and 16 (out of 36) body postures of a particular person are also shown in Fig. 1b. These images are submitted to the proposed image processing as described earlier. The result of edge images is also shown in Fig. 1c.

We have generated 30 individual posture curves (single person) as shown in Fig. 2. These posture curves are obtained from individual models and placed in a same axis dimension in order to compare the effect of proposed image processing graphically. It should be noted that these posture curves (Fig. 2b) have generated using LoG images and this is just separate eigenspaces from the respective models. Therefore, we have avoided the averaging the posture sets for this issue, i.e., covariance matrix $C = X^h X^{h^{\mathrm{T}}}$. This mean posture matrix is applied for recognizing unknown postures.

(a)

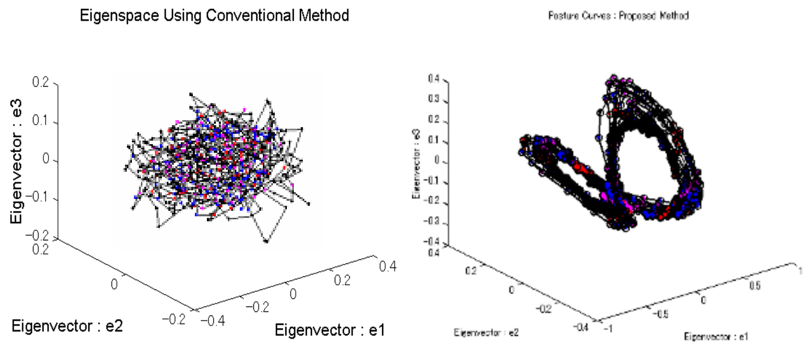
(b)



(c)

Fig. 1. Human models, postures and processed images: (a) 12 human models out of 30 used in the experiments, (b) 16 postures out of 36 of a model, and (c) blurred edge or LoG images

This performance is just for highlighting the effect of proposed image processing, i.e., LoG images. The posture curves obtained from the conventional method [7] are illustrated in Fig.2a, whereas those derived from the proposed method are depicted in Fig.2b. It is obvious that the dress effect has made the posture curves completely different with each other by the conventional method, though the models' postures are similar. On the other hand, the dress effect has been successfully overcome in the proposed approach as shown in Fig.2b. It is noted that we differentiate between the conventional and proposed method by their input images in creating eigenspace.



(a) (b)

Fig. 2. Individual posture curves obtained from (a) the conventional method and (b) the proposed method

3.2 Recognizing Human Postures Employing the Proposed Eigenspace

We divide all data sets randomly (taking equal number) into three sets, i.e., LS, TS_A and TS_B where LS denotes learning samples and TS denotes testing samples. We employ a leave-and-out or k-fold cross validation method for choosing the learning samples for making the proposed ES so that all data sets can be used either for training or testing. Therefore, when we use LS for generating an ES, TS_A and TS_B data sets remain for testing. Similarly, if we choose TS_A as learning samples, LS and TS_B are used for testing purposes, and *vice versa*. Table 1 shows the distribution of data set for generating the proposed ES and the samples for testing. The average recognition results of each test are also shown in this table. Employing a learning sample such as LS, we calculate a mean data set by Eq.(5) and generate the proposed eigenspace from it. Fig. 3 shows a proposed ES of the data set LS. This proposed ES is used for recognizing image data contained in TS_A and TS_B. Moreover, the leave-and-out method can also arrange the data sets such a way that each time 29 data sets could be used for learning and single data set could be employed for testing. In this case, we have obtained an average of 88.6%. However, if we use more data set, the CPU time will be a bit higher. The recognition rate is calculated dividing the ratio of successful hits and the total postures projected. It is mentioned that we have projected unknown but similar postures onto the basic eigenspace and decided the successful hits based on the minimum description length.

We have obtained an average recognition rate of 71.66% using only the edge images, mean posture matrix are not taken into consideration here, where only one set was used for learning and 29 sets were for testing. This result only highlights the progress of our proposed method. It is noted that the proposed method works only with

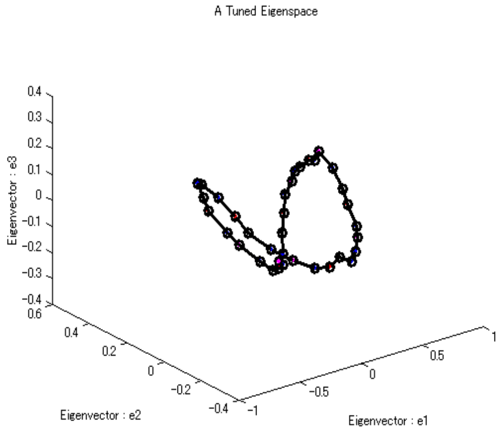


Fig. 3. A proposed eigenspace created from S_A data set

the combined application of LoG image processing and mean posture matrix for creating eigenspace. Classification results between two conventional and the proposed methods are also shown in Table 2. According to the papers of Murase and Nayar [7] and Murase and Sakai [12], we have used the original gray images and silhouette images, respectively for their input images. In these cases, a best-search method has applied for selecting an appropriate learning sample while other data sets are used for testing (i.e., total of 29 data sets) purposes. According to the conventional methods, we need to search a based model for the learning samples and the others will be remaining for testing. Therefore, we have used only one sample for learning and rest of the samples have used for testing. On the other hand, the proposed method calculates a mean posture matrix and we have taken 10 image set for obtaining this mean. Therefore, the rest of 20 samples have used for testing. Obtained higher recognition rates have proved the robustness of the proposed method. In case of time efficiency, we need only 1.5 Sec. (CPU time) using 1 G Htz PC and Matlab implementation software.

Table 1. Data distribution for the proposed ES and recognition

Test	S1-10	S11-20	S21-30	Recognition Rates (avg.)	Average (all data)
(i)	*LS*	TS_A	TS_B	86.2%	85.6%
(ii)	TS_A	*LS*	TS_A	83.4%	
(iii)	TS_B	TS_B	*LS*	87.2%	

Table 2. Classification results between conventional and proposed methods

Methods	Recognition Rates	Test Sets	Eigen Dimension	Mean Square Error
Murase 95[7]	40.05%	29	8	0.0345
Murase 96[12]	71.29%	29	8	0.0345
Proposed ES	85.6%	20	4	0.0345

4 Discussion and Conclusions

This paper has proposed an updated eigenspace technique for human body posture recognition, which allows some image pre-processing techniques. The proposed eigenspace technique has also considered a mean posture matrix for developing the ES. It has successfully reduced the effects of human posture-change due to loose clothing and various body shapes and this result was experimentally shown. We have also compared the proposed method with conventional methods and the obtained results have indicated that the proposed ES method is better than the conventional ones in terms of recognition rates for human body posture recognition. Since our method has employed the image contour and their internal edges, it has more demand in the field of gesture recognition. We have also reduced the processing time by compressing image size and requiring less eigen dimension.

The classification of dresses and human body shapes normally has a large variety. We have employed 30 person's body postures with their presently worn dresses. Since they are mostly from the same society (however there are two models from different nationalities), the body shapes were not so different. Employment of more different types of clothes such as lady's dresses and textured clothes and various body shapes such as Japanese traditional Osumo people are expected to the future study.

Some possible extensions of the proposed method can also be proposed in this section. In particular, the applications of image processing techniques for reducing dress texture effect, generalization (mean) of postures, eigenspace selection may develop some new algorithms in this area for flexible objects recognition where individual eigenspace generating may not be essential. This proposed approach can be of interest to the researchers of various fields and the scope of this study can be extended for recognizing human behavior, activities, motions, gestures, etc.

Bibliography

1. Sirovich, L. and Kirby M.: "Low dimensional procedure for the characterization of human faces", J. Optical Society of America, Vol. 4, No. 3, pp. 519-524(1987).
2. Turk, M.A. and Pentland, A.P.: "Face recognition using eigenfaces, Proc. of the Computer Vision and Pattern recognition", pp. 586-591(1991).
3. Leonardis, A., Bischof, H.: "Robust recognition using eigenimages", *Computer Vision and Image Understanding*, 78, pp.99-118(2000).

4. Ohba, K., Ikeuchi, K.: “Detectability, uniqueness, and reliability of eigen windows for stable verification of partially occluded objects”, *IEEE Tran. on Pattern Analysis and Machine Intelligence*, PAMI-9, pp. 1043-1047 (1997).
5. Borotschnig, H., *et. al.*: “Appearance-based active object recognition”, *Image and Vision Computing*, Vol. 18(9), pp. 715-727(2000).
6. Black, M. J., Jepson, A. D.: “Eigen tracking: robust matching and tracking of articulated objects using view-based representation”, *Int. Journal of Computer Vision*, Vol. 26(1), pp. 63-84(1998).
7. Murase, H., Nayar, S. K.: “Visual learning and recognition of 3-D objects from appearance”, *Int. J. Computer Vision*, 14, 5, 39-50(1995).
8. Tenenbaum, B., Freeman, W. T.: “Separating style and content with bilinear models”, *Neural Computation*, 12 (6), pp. 1247-1283(2000).
9. Beymer, D., Poggio, T.: “Face Recognition from one example view”, *Proc. of Int. Conf. on Computer Vision*, pp.500-507, 1995.
10. Moghaddam, B., *et al.*: “Bayesian face recognition”, *Pattern Recognition*, Vol.33 pp.1771-1782 (2000).
11. Romer R. and Stan S., “Specialized mapping and the estimation of human body pose from a single image”, *IEEE Workshop on Human Motion* (2002).
12. Murase, R. Sakai: "Moving object recognition in eigenspace representation: Gait Analysis and lip reading", *Pattern Recognition Letters*, Vol. 17, pp. 155-162, (1996).
13. Yilmaz, A., Gokmen, M.: “Eigenhill vs. eigenface and eigenedge”, *Pattern Recognition*, 34, pp.181-184 (2001).