EigenBody: Analysis of body shape for gender from noisy images

Matthew Collins, Jianguo Zhang, Paul Miller, Hongbin Wang and Huiyu Zhou
Institute of Electronics Communications and Information Technology (ECIT)
Queens University Belfast

Abstract

We present an analysis of full body images for gender classification using Principal Component Analysis (PCA). This has been widely used in the past for gender classification of faces and we show that similar techniques can be applied to the full body domain. Using Linear Discriminate Analysis (LDA) we are able to identify the key PCA components which encode information related to gender. The paper confirms that intuitive thoughts about what properties of the human body are important for gender classification, can be effectively represented in a small number of key components and that eigenpeople reproduced using just these key components can be visually classified by gender to a large extent.

1 Introduction

Within the field of video surveillance, the ability to profile people based on age and gender is of particular interest. Such profiling would allow the development of intelligent CCTV systems which could identify the intrinsic threat posed by an individual to others. This work is concerned specifically with gender.

A substantial body of work exists classifying gender from facial images, as face has been identified as being of key importance in humans determining gender [11]. For example Mäkinen and Raisamo [8] presented a systematic study on gender classification with automatically detected and aligned faces. They experimented with 120 combinations of automatic face detection, face alignment and gender classification and found that the best gender classification rates of around 86% were achieved with a Support Vector Machine (SVM).

Buchala *et al.* [1], investigated principal component analysis (PCA) for face classification with regard to gender, ethnicity, age and identity. They found that these characteristics could be encoded in a relatively few number of PCA components, and that these were predominantly to be found amongst the first few. With respect to gender, they found that the third and fourth PCA components were key and that information related to complexion, length of nose, the presence or absence of hair on the forehead, eyebrow thickness, and the presence or absence of a smile was useful.

Despite relative success in face based gender classification, there are weaknesses in this approach, especially in a video surveillance context where the high resolution face images required by these methods are not available due to the much wider field of view of the camera. Generally in a CCTV context, the subject is non cooperative and if the face is visible at all, it will be of low resolution or could be partially occluded. For this reason there has been recent investigation into gender classification from features extracted from full body static images of pedestrians [3, 2].

Our previous work investigated the suitability of histogram of gradient (HOG) [4] features for the full body gender classification problem. HOG has been extremely successful in pedestrian detection, so it was felt that these features should be particularly suitable in describing the human body shape. They achieved 80% accuracy scores using a combination of HOG shape information and colour features.

This work showed that human body shape could be captured and used for effective gender classification. It did not however go into detail about specific shape differences between the genders and what information in particular was useful for the classification task.

In this paper we perform principal component analysis (PCA) [7] on the full body image dataset [3]. We use Fishers linear discriminant analysis (LDA) to identify which components are important with regards to encoding the gender information and attempt to identify what properties are being encoded by these components. To the best of our knowledge, this is the first time that this kind of analysis has been performed on this type of data.

PCA has also been used in analysis of 3D human body shape with respect to gender from captured full body 3D

M. Collins, J. Zhang, P. Miller, H. Wang and H. Zhou are with the School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, United Kingdom.
Email: {mcollins17;jianguo.zhang;p.miller;h.wang;h.zhou}@qub.ac.uk

models [14, 13]. Very high accuracies of 93% have been reported, though in a strictly constrained, experimental environment, with idealised data. Intuitively, a static pedestrian image can be considered similar to an image of a frontal view of one of these 3D models and both of these should contain a lot of the same body shape information useful for gender classification.

2 Experiments

2.1 Dataset

In the original full body gender work [3] we created two gender datasets from two publicly available pedestrian recognition datasets, the MIT pedestrian recognition set [10] and the viewpoint invariant pedestrian recognition set (VIPeR) [6]. In this work we have merged the images from both datasets and uniformly scaled them giving us a larger number of images to work with.

Thus the resulting dataset contains 413 images of each gender and all frontal poses. The images are all quite tightly cropped with the subject in the centre of the image, with minimal background. The images are taken from natural pedestrian photographs and as such the subjects are not in standardised poses. Whilst they are all frontal view images, there is some variation in pose and the subjects are not standardly aligned across all images, unlike the FERET [12] face image dataset which has been used for similar analysis in a face context [1]. In this sense and because of the degree of variation in clothing and image clarity, we describe these as 'noisy' images.

2.2 PCA of Full Body Images

The images are converted to grey-scale and then each image is represented as a single column vector. The dataset contains 826 images of 106 x 45 resolution, which results in 826 vectors of length 4770. It is straightforward to perform PCA of these vectors, which results in 826 non-zero eigenvectors of length 4770, however the first 200 components of these eigenvectors accounted for approximately 90% of the variance and thus each image can be efficiently represented using the just 200 components instead of 4770. It is worth pointing out that this is a global computation of PCA, computing components common to both gender classes simultaneously. PCA was performed using a publicly available pattern recognition toolbox. [5]

The well known term *'Eigenfaces'* was coined with respect to visualising the eigenvectors which capture faces variation, and in a similar way the eigenvectors computed on these body images can be visualised as *'Eigenpeople'*.

Because the variations in pose account for a very small amount of the overall variance within the data, the 'Eigenpeople' tend to display a standardised mean pose as the average across all images in the dataset, unless the components accounting for pose are included.

Figure 1. Example Eigenpeople. Top: original body images. Middle: Constructed images using the first 200 components accounting for 90% of the variance. Bottom: Just the first component, with the largest variance showing a standardised pose but no clear gender information.

2.3 Key Components for Gender

Principle component analysis is often used as a dimensionality reduction technique. The components are ordered according to their importance in accounting for the variance in the data. Often a high percentage, usually above 80% of this variance occurs within the first few components and the remaining components can be discarded without significant data loss. However it is possible that the properties of interest could be encoded in components which are perhaps not significant in discriminating the data. Thus, it is well known that the selection of components should be based on their importance for a particular task, rather than their importance in accounting for total variance in the data.

Using the Fisher Ratio, which takes account of between class scatter and within class scatter; it is possible to identify the encoding power of the individual components for the 2 class problem of gender classification.

Figure 2 shows a plot of the estimated encoding power of the first 200 PCA components. Unlike similar analysis of gender in faces where one or two components usually significantly outperform the others [1] the gender from body seems to be represented by a combination of multiple components.

Looking at this graph, and by selecting a threshold cut off point, the top seven components for gender classification in this dataset appear to be (in descending order) components: 14, 3, 5, 37, 20, 33, and 85.

These seven components show much higher encoding power than the rest and visual examination of the corre-

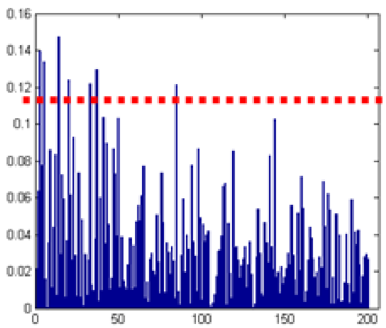


Figure 2. Bar graph of the estimated encoding power of the first 200 components with respect to gender. Dotted line indicates a chosen threshold for selecting the top 7 which are considered important.

sponding 'Eigenpeople' confirms that they appear to effectively capture relevant gender information. Figure 3 shows the effect of reconstructing images in the datasets using only these key components. Even with just the first component, visual gender discrimination is possible, with the differences becoming more pronounced as more key components are included. Figure 4 shows two particularly clear examples where the gender of the reconstructed images can be visually determined using just the top seven identified gender components.

2.4 What is encoded by each component?

Intuitively, when thinking about body shape and gender, one would assume that key elements would be broad shoulders for males, and narrow shoulders, but wider hips for females, giving male body shapes a downward pointing wedge shaped appearance and females, more diamond shaped. Studies have shown that males tend to show a higher occurrence of the mesomorph somatotype, one of the characteristics of which is a wedge shaped body. Similarly females displayed a higher rate of endomorphy including wider hips [9]. Also one would assume that the chest area would be key in differentiating between the genders and perhaps head shape indicating the presence/absence of long hair.

Closer examination of the individual key components and what information they encode seems to indicate that they do indeed capture these properties. In order to attempt to determine what a specific component encodes, a series of reconstructed images are generated. First, the average of the components of all images is estimated and an image is reconstructed from these average components. This is the average person. We then modify the component of interest by adding or subtracting the standard deviation of that component multiplied by a constant, leaving the other average components unchanged. The resulting reconstructed images become more or less masculine/feminine depending on the component in question and what it encodes.

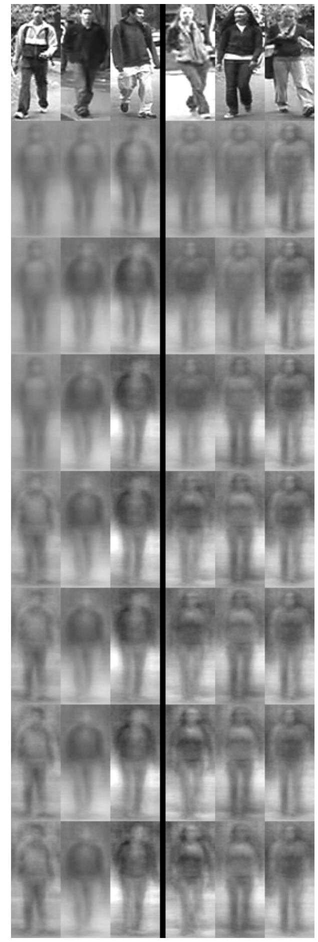


Figure 3. Sample images and their resultant Eigenpeople. As each of the top seven PCA components is included the images become more recognisable as their respective genders.

Figure 4. Example Eigenpeople using just the top seven gender components. Left Male, Right Female.

Figure 5 shows these reconstructed images for component 3, which are particularly clear showing the image becoming progressively more masculine to the extreme left, and more feminine to the extreme right.

It appears to capture information regarding shoulders,

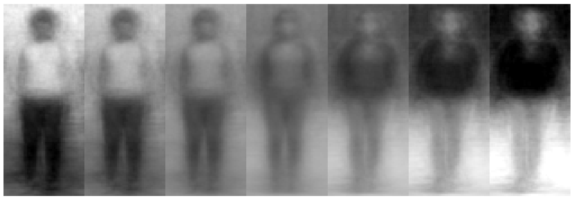
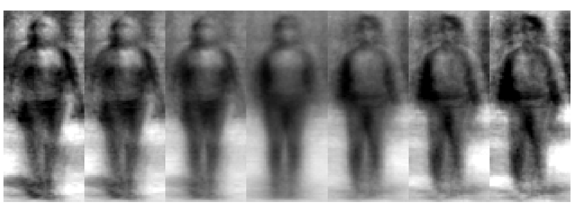


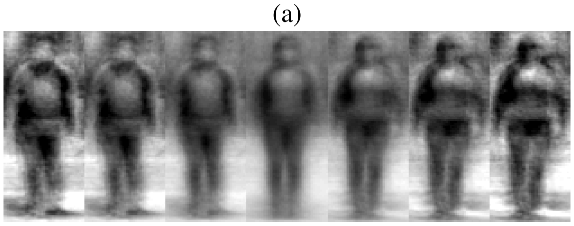
Figure 5. Reconstructed images from alteration of component 3. Components are progressively added/subtracted from the mean in steps ranging from -3 S.D (extreme left) to +3 S.D (extreme right).

showing males to have wider shoulders than females, and it also seems to capture the pose of the legs, where females tend to adopt a stance where their legs are closer together than males.

Similarly Figures 6 (a) and (b), though the differences are less pronounced than component 3, show components 33 and 37. In both cases here the components appear to be encoding information to do with the chest and hip regions of the body. Component 33 (Figure 6 (a)) seems to encode male information with the image becoming more recognisably male to the extreme right and more female to the extreme left. The reverse can be seen in component 37, where the reconstructed image seems to display the typically recognisable female body curves with repeated addition of the component.



(a)

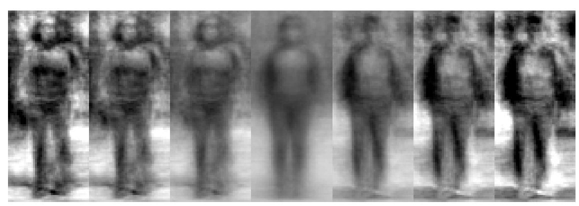


(b)

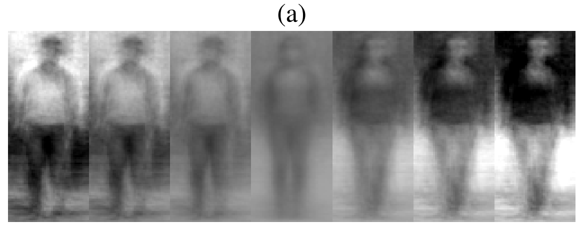
Figure 6. Reconstructed images showing the effect of repeated alteration of component 33 (a) and 37 (b)

This effect of the reconstructed image becoming more or less a particular gender as the specific component is repeatedly added or subtracted from the average can be observed for all the identified key components. For some components it is easier to visually identify their effect than others but it is clear when looking at the most extreme left and right images, most far removed from the average, that gender information is what is being captured. Figure 2 shows that no single component dominates in terms of gender encoding power though, and it is important to look at the effect of combinations of these key components.

Finally we show the reconstructed images for the composites of these top gender components. The bar graph as shown in Figure 2 used to identify the components with the highest encoding power from the Fisher ratios is generated from the absolute values of these ratios. In fact, if the sign of the component is also taken into account it indicates which of the two classes the component is particularly suited towards. This is the reason that repeated addition of some components causes the reconstructed images to become more masculine, while others become more feminine. Of the top seven identified gender components, four of them appear to encode female information and the remaining male. Figure 7 (a) and (b) shows composite images for the male and female components respectively. Here, because there are multiple contributing factors one cannot account for what is specifically encoded. However it is clear that the composite image from the male components becomes visibly more feminine to the extreme left as components are repeatedly subtracted, and more masculine as components are repeated added to the mean image. Similarly, repeated addition of the female components results in a reconstructed image which is distinctly feminine in body shape and appearance.



(a)



(b)

Figure 7. Reconstructed images showing the effect of repeated alteration of multiple key gender components.
(a): key male components 14, 33 and 85
(b): key female components 3, 5, 37 and 20

The female images have wider hips, visible bust shape in the chest area and even the head has taken on a more indistinct shape indicative of long hair. The leg positioning/pose also takes on a more typically feminine shape. In the case of both genders, subtraction of the relevant components seems to result in an image of the opposite gender, indicating that the components are complementary to each other.

2.5 PCA of Edgemaps

The experiments described thus far are based on PCA performed on the raw pixel values of original images and as such are appearance based. The original features used in our previous paper on gender classification [3], dealt with body shape and not appearance. The body shape was initially captured using an edgemap, and then described using histograms of distributed gradients which were computed over a grid of cells that the image was divided into.

In order to shed some light on the shape differences being captured between the genders, we perform similar PCA analysis on edge maps taken from the images. Both Canny and Sobel filters were examined however there was negligible difference observed in their performance.

Ideally, if perfect edge maps could be taken from the images, fully capturing the body in every image, then key landmark co-ordinates could be identified as in the case of the 3D model work [14]. Every image would have a complete set of corresponding landmarks and the feature vectors used to describe the images could consist of distance metrics between these. Unfortunately, due to the nature of the dataset available this is not feasible. Perfect edge maps for every image cannot be automatically computed and often individual limbs are missing and even in some cases, it is hard to identify when looking at the edgemap that a human figure has been found (see Figure 8). The non-standard poses of the figures in the dataset also means that edgemaps cannot be computed to have fully corresponding co-ordinates across all images.

Figure 8. Sample images and their corresponding Canny edge maps

However, what is clear is that when the male and female mean edgemap images are computed, the resulting images show marked gender differences, and if asked, a human observer would conclude correctly which was which (see Figure 9).

PCA can be performed on these edgemap images in a similar fashion to the raw pixel value experiments. The resulting reconstructed eigenimages, are no longer binary due to the way in which the data is mean adjusted during PCA analysis. Key gender components can be identified in this space also, and repeated adding or subtracting of these components to the average image also results in a noticeable shift in the shape of the reconstructed images, showing expected gender differences. The gender differences are not as distinct as the appearance based pixel analysis, however the intuitive shape differences between the genders one would expect to find, are still present. As can be seen in Figure 10 (a), the female component seems to be encoding information to do with the waist and as the image becomes more feminine to the extreme right, expected female body curves become evident. In the case of Figure 10 (b), which is a male component, the shape becomes more distinctly male as the component is repeatedly added. In this case, repeated subtraction does not result in a more clearly female shape as was observed with some other components, but rather an indistinct image. This indicates that not all the key gender components encode relevant information about both classes and some are specifically suited to describing features of one class or the other.

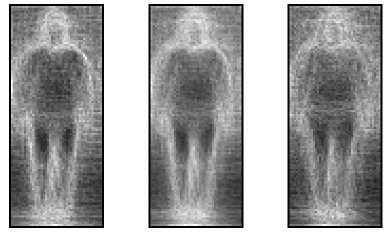
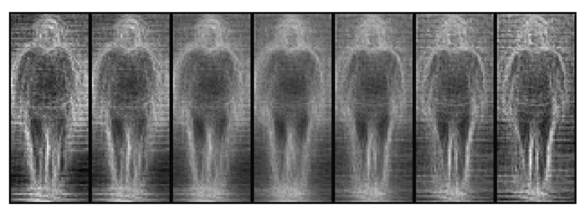
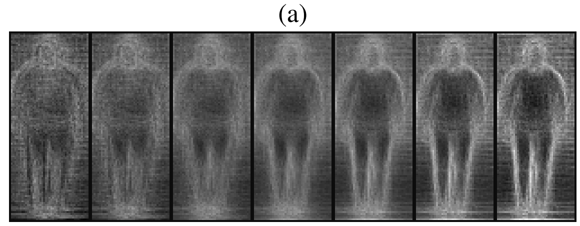


Figure 9. Average images from computing the mean of the edgemap images. Left: mean male edgemap, Middle: global mean, Right: mean female edgemap.



(a)



(b)

Figure 10. Reconstructed images showing the effect of repeated alteration of two key gender components identified from PCA analysis of Canny edge map images. (a) female (b) male

3 Classification Results

Full body gender classification is a much more difficult problem than the facial equivalent due to a wide degree of variation in a number of factors, and also in the case of our data, non-standard body poses. The data we use here is from real world photographs, non-standardised poses, relatively low resolution and with a degree of background clutter. As such, 90th percentile accuracies like those reported by Wuhrer *et al.* [14] on their idealised 3D model data are not feasible with this approach on this data. Our highest previously reported results of 80% for gender classification of this dataset , combined body shape and colour cues and in this work we only consider components relating to body shape. Our highest reported accuracies for body shape alone were 75%.

Whilst direct classification is not the focus of this work, we did use an SVM classifier in a 5 fold cross validation setup to perform gender classification on the top seven identified key components for gender classification from the raw pixel data in order to test their encoding power. A mean accuracy of 63% is reported which while not as high as the HOG based approach is better than random guess and confirms that relevant gender information is captured in these components. Only 59% accuracy is reported on classification of the top seven components from the edgemap data though this is understandable as the appearance based, raw pixel data showed a much clearer distinction when the reconstructed images were visualised. Combining the appearance based components with the edgemap based data using the two stage SVM approach presented by Zhang *et al.* in [15] resulted in a higher overall accuracy than either alone of 66%.

4 Conclusion

We present an analysis of full body static images using PCA of both raw pixel data and of the appearance of binary edgemaps in order to identify the key components which indicate gender. The work identifies seven key components in each case which when visualised conform to intuitive ideas about what the human observer would look for when trying to classify gender manually. Classification is also performed using these PCA features to confirm the effectiveness of out analysis and demonstrate the encoding power of the identified components.

Limitations of the available datasets for gender analysis have also been identified and will be the focus of future work.

5 Acknowledgement

The authors acknowledge funding from the Department for Employment and Learning Northern Ireland (DEL), the Queens University Belfast Research Support Package D8203EEC, and from EPSRC grant EP/E028640/1 ISIS.

References

[1] S. Buchala, N. Davey, T. M. Gale, and R. J. Frank. Principal component analysis of gender, ethnicity, age, and identity of face images. In *IEEE ICMI*, 2005.

[2] L. Cao, M. Dikmen, Y. Fu, and T. S. Huang. Gender recognition from body. In *MM '08: Proceeding of the 16th ACM international conference on Multimedia*, pages 725–728, New York, NY, USA, 2008.

[3] M. Collins, J. Zhang, P. Miller, and H. Wang. Full body image feature representations for gender profiling. In *IEEE Workshop on Visual Surveillance (ICCV)*, 2009.

[4] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 886–893 vol. 1, 2005.

[5] V. Franc and V. Hlaváč. Statistical pattern recognition toolbox for Matlab. Research Report CTU–CMP–2004–08, Center for Machine Perception, K13133 FEE Czech Technical University, Prague, Czech Republic, June 2004.

[6] D. Gray, S. Brennan, and H. Tao. Evaluating appearance models for recognition, reacquisition, and tracking. *Performance Evaluation of Tracking and Surveillance (PETS). IEEE International Workshop*, 2007.

[7] I. Jolliffe. *Principal Component Analysis*. Springer, October 2002.

[8] E. Makinen and R. Raisamo. Evaluation of gender classification methods with automatically detected and aligned faces. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(3):541–547, 2008.

[9] M. J. Muoz-Cachn, I. Salces, M. Arroyo, L. Ansotegui, A. M. Rocandio, and E. Rebato. Body shape in relation to socio-economic status in young adults from the basque country. *Collegium antropologicum*, 31(4):963–8, 2007.

[10] M. Oren, C. Papageorgiou, P. Sinha, E. Osuna, and T. Poggio. Pedestrian detection using wavelet templates. In *Computer Vision and Pattern Recognition*, pages 193–199, 1997.

[11] A. J. O'toole, K. A. Deffenbacher, D. Valentin, K. Mckee, D. Huff, and H. Abdi. The perception of face gender: The role of stimulus structure in recognition and classification. *Memory and Cognition*, 26:146–160, 1997.

[12] P. J. Phillips, H. Wechsler, J. Huang, and P. J. Rauss. The feret database and evaluation procedure for face-recognition algorithms. *Image and Vision Computing*, 16(5):295, 1998.

[13] A. Ruto, M. Lee, and B. Buxton. Comparing principal and independent modes of variation in 3d human torso shape using pca and ica. In *ICA Research Network International Workshop*, 2006.

[14] S. Wuhrer, C. Shu, and M. Rioux. Posture invariant gender classification for 3d human models. *Computer Vision and Pattern Recognition Workshop*, 0:33–38, 2009.

[15] J. Zhang, M. Marszalek, S. Lazebnik, and C. Schmid. Local features and kernels for classification of texture and object categories: A comprehensive study. *IJCV*, 73(2):213–238, June 2007.