Reconstruction of Walking People Images by Principal Component Analysis

Zhipeng Zhao ZHIPENG@PAUL.RUTGERS.EDU
Computer Science Department, Rutgers University, Piscataway, NJ 08854 USA

Abstract

The Principal Component Analysis (PCA) is a useful statistic 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. 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. We can do image data compression by only keeping the first few principal components. In this project, we will show that we can compress the walking people images and later reconstruct the original images without much loss from the reduced data.

1. Introduction

1.1 Representing walking people image

Finding a good representation of walking people images is an active research field in computer vision. A good representation of walking people can be used to build a model, which is very useful in many applications such as walking people tracking, virtual reality and animation.

Many approaches have been taken to find representations. Briefly, they can be put into two categories. One is shape - model based and the other is non-shape model [1]. The shape-model based approach usually builds a 2D or 3D model of the walking people based on the kinematic and shape properties of the human body. Badler's "Bubbleman"[2] and Gavrila's work [3] belong to this category.

The other approach usually describes walking people in statistical terms derived from low level features. Oren[4] applied wavelet transformation to images and uses wavelet coefficients as low-level features.

The approach we take in this project belongs to the second category. We apply principal component transformation to extract the features from original data. Principal component analysis has been applied in many fields. [6]. Pentland etd. al. [5] took the same approach to extract features from faces. Here, we try to apply it to walking people's images.

To demonstrate that we create a good model of walking people, we reconstruct the original images from our model.

1.2 Principal Component Analysis

The shape-model based approaches of representing walking people images tend to ignore the issue of what features of the images are important to represent walking people. This suggests to us that an information theory approach of encoding walking-people images might give insight into the information content of walking people images, emphasizing the significant local and global features.

From the view of information theory, we want to extract the relevant information from walking-people images, encode it as efficiently as possible and use it to represent the original images. A simple way to extract the information contained in an image of walking people is to capture the variation in a set of walking people images and use this information to encode an individual image.

In mathematical terms, this approach is to find the principal components of the distribution of walking people images, or to find the eighenvectors of the covariance matrix of the set of walking-people images, treating each image as a vector in a very high dimensional eigen space. The eigenvectors are ordered, each accounting for a different amount of the variation among the images.

These eigenvectors are a set of features that together characterize the variation between walking people images. In our model, each image can be represented by a linear combination of these eigenvectors. Each image can also be approximated by the first few eigenvectors, which have the largest eigenvalues and therefore account for the most variance within the set of walking people images.

In Pentland's [5] approach, he called these eigenvectors "eigenfaces". And each of face images is a linear combination of eigenfaces.

2. Algorithm

This approach includes the following steps:

1) 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 eigenspace.

2) Project a new walking 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 feature space. The set of weights will be the encoding of the new image in the eigenspace.

3) Reconstruct from the feature space: To test our model, we will use the eigenvectors and the image encoding to transform the image back to the original space and measure the pixel-level distortion.

2.1 Build feature space

Let training images of walking people be a set of centered vectors: X_k, k=1,…, M, $\sum_{k=1}^{M} x_k = 0$, $X_k \in R^N$. PCA diagonalizes the covariance matrix:

$$C = \frac{1}{M}\sum_{j=1}^{M} X_j X_j^T \qquad (1)$$

To do this, we have to solve the eigenvalue equation

$$\lambda Vi = CVi \qquad (2)$$

for eigenvalues $\lambda \geq 0$ and $V_i \in R^N \backslash \{0\}$. The eigenvectors V will be the axes for the feature space. We reorder the vectors V by the corresponding eigenvalues. We choose the first k vectors as principal components and build the feature space with them as the axes.

2.2 Project a new image into feature space

To project a new image $U \in R^N$ into the feature space, we need to first center the image by subtracting the mean of the training data X. Then, we do a simple projection operation.

$$\omega = V^T (U - mean(X)) \qquad (3)$$

The vector ω, which is also an N dimensional vector, can be seen as the new encoding of the image in the feature space. In the feature space, since we build the space with the principal component vectors, and the most significant features will be in the first few axes, only the first few elements in ω is significant.

This suggests that we can also do data compression by applying PCA. If we only keep the first few elements in ω, we can still keep the most important information about the original image.

2.3 Reconstruct image from feature space

The eigenvectors V is orthogonal, $V^T V = E$. From (3), we have

$$V\omega = U - mean(V) \qquad (4)$$

So we can reconstruct the image from the feature space by the following operation:

$$U = V\omega + mean(V) \qquad (5)$$

3. Experiments and results

3.1 Acquire a set of images as training data

We acquired a set of walking people images data and chose 9 images as training data to build the feature space and other 9 images for testing. Each image has 40×53 pixels.

3.2 Build the feature space

In this project, the programming environment is MatLab. For each of the 40×53 images, we reshaped them into a 1×2120 vector. Then, we calculated the covariance matrix C according to (1). Next, we computed the eigenvectors and eigenvalues of the covariance matrix. At the end, we kept the k, k=1..9, most significant eigenvectors and got an 2120×k matrix V. Each column of V is an eigenvector of C.

We tried to use different k in our experiments. We can see that with the increase of k, the reconstructed images will have less error from the original images. According to different problem, we can make a tradeoff between k and the error so that we only need to keep a small k dimensional data while still maintaining a good approximation.

3.3 Project a new image into feature space

We used both testing and training images for projecting into the feature space. For each of images, we got the corresponding encoding.

3.4 Reconstruction from the feature space

We have tried to reconstruct both training images and testing images.

Figure 1 shows the reconstruction of a training image. We can see that with the increase of eigenvectors, the reconstructed images are more and more close to the original image.

To measure the image reconstruction, we used the error between the reconstructed image and original image.

Figure 2 shows the reconstruction error of the testing images. We used the sum of the Euclidian distance between the corresponding points from the reconstructed image and original images as the measure for reconstruction error. For each of the k dimensions, we sum up the errors from the reconstruction of all 9 images. With the increase of k, which is the number of eigenvectors used to build the feature space, we can see the error is decreasing.

Figure 3 shows the reconstruction errors of training images. We can see that when k approaches the number of training images that we use to build the feature space, the error is very close to 0. An intuitive explanation is that we can build another feature space S' using each of the training images as axes. The dimension of the S' is the number of training images. In this feature space, all training images can be represented without loss. The feature space we build from PCA should be better than S' to represent the training images, so when the number of dimensions approaching of that of S, its error is also approaching 0.

4. Conclusion

From our project experiment result, we find that even if we just keep the first 3 or 4 principal components, we can still get a good reconstruction of the images. This means principal component analysis is a good approach of extracting features from the data and therefore, a good way to build model for walking people image. However, principal component analysis is a linear approach, which cannot represent the non-linear features in the data. The improved version of PCA, such as kernel PCA, might be a good complement for it. Kernel PCA will first use non-linear mapping to project the data into a high dimensional space, then it will apply PCA to the data in the high dimensional space. We hope after the nonlinear mapping, the non-linear feature of the original data can be linear in the high dimensional space and can be represented by PCA.

Acknowledgements

I want to thank Prof. Elgammal for his help on my understanding of PCA. He also provided walking-people images as the data for my project

References

[1] Gavrila, D. M. (1999). The visual Analysis of Human Movement: A Survey. *Computer Vision and Image Understanding (*vol 73, no 1, pp. 82-98*)*.

[2] O'Rourke, J. & Badler, N. (1980). Model-based image analysis of human motion using constraints propagation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. (2(6): 522-536).1980.

[3] Gavrila, D. (1996).Vision-based 3-D tracking of humans in action. *PhD thesis* , Department of Computer Science, University of Maryland, 1996

[4] Oren, M., Papageorgion, C., Sinha, P., Osuna. E., & Poggio, T., (1997). Pedestrian detection using wavelet templates. In *Proc. Of IEEE Conference on Computer Vision and Pattern Recognition.* (pp p93-199). San Juan: 1997.

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

[6] Smith, L. (2002). A tutorial on Principal Components Analysis. (http://kybele.psych.cornell.edu/%7Eedelman/Psych-465-Spring-2003/PCA-tutorial.pdf) Maintained by Cornell university, U.S. A

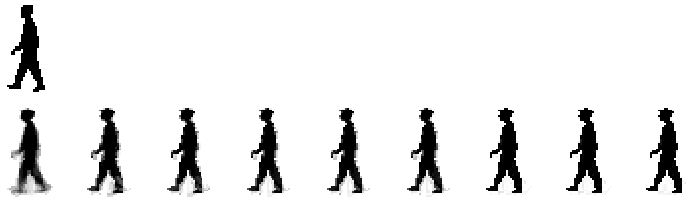

Figure 1:
The above image is the original image. The images below are the reconstructed images. From left to right, we are using k=1 to 9 eigen vectors for reconstruction.

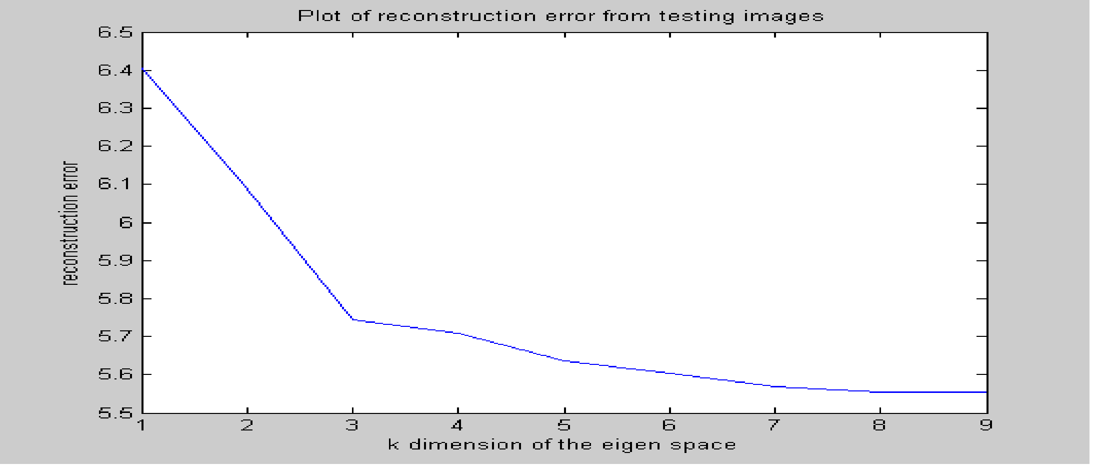

Figure 2:
The reconstruction errors decrease with the increase of the number of principle components used to reconstruct the testing images.

Figure 3:
The reconstruction errors decrease with the increase of the number of principle components used to reconstruct the training images.