Image Compression and Face Recognition: Two Image Processing Applications of Principal Component Analysis

Aleš Hladnik

Keywords: principal component analysis, image compression, face recognition, image processing, computer vision

Principal component analysis (PCA) is one of the most widely implemented tools for dimensionality reduction or data exploration used in a variety of scientific and engineering disciplines. It transforms a number of possibly correlated variables into a smaller number of new variables, known as principal components. Since a digital image can be regarded as a two – or more – dimensional function of pixel values and represented as a 2D (grayscale image) or 3D (color image) data array, PCA can be performed on such an m x n matrix. After a brief theoretical introduction the paper will focus on two typical image processing/computer vision applications of PCA. First, we will demonstrate how to apply the method to compress different digital images and, in particular, how the choice of the number of extracted PCs affects the image compression ratio and consequently the image quality. Face recognition is a computer vision topic of active current research with a broad range of applications, including law enforcement and surveillance, personal identification, video content indexing, and access control. Individual steps involve segmentation – detection – of faces from a scene, feature extraction from the face region followed by either recognition or verification. MATLAB implementation of PCA based on a freely available face image database will be shown and discussed.

1. Introduction

The aim of this paper is to demonstrate how two at first glance not quite closely related technical fields – digital image processing and multivariate statistics – can successfully be combined in two applications of high current interest: lossy compression of digital images and human face recognition. One of the many approaches to these applications is via an implementation of Principal component analysis (PCA).

1.1 Principal component analysis

PCA is a multivariate statistical technique frequently used in exploratory data analysis and for making predictive models. It is based on a construction of a new set of variables, referred to as principal components (PCs), which are completely uncorrelated – orthogonal – to each other. PCA can be accomplished either by eigenvalue decomposition of a data covariance (or correlation) matrix or by singular value decomposition (SVD) of a data matrix, usually after mean centering and normalization (Abdi, 2010). Below is a brief outline of the SVD approach (Cao, 2012).

If we have a matrix **A** with *m* rows and *n* columns, with rank *r* and $r \le n \le m$, than SVD transforms **A** into three matrices **U**, **S** and **V** (Figure 1):

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T \qquad (1)$$

where $\mathbf{U} = [u_1, u_2, \ldots, u_r, u_{r+1}, \ldots, u_m]$ is a an *m* x *m* orthogonal matrix, and $\mathbf{V} = [v_1, v_2, \ldots, v_r, v_{r+1}, \ldots, v_n]$ is an *n* x *n* orthogonal matrix.

Column vectors u_i for i = 1, 2, …, *m* and v_i for i = 1, 2, …, *n* form orthonormal sets. **S** is an *m* × *n* diagonal matrix. The diagonal entries, σ_i (i = 1, 2, …, *n*) of **S** are called singular values of matrix **A**. It can be achieved that:

$$\sigma_1 \ge \sigma_2 \ge \ldots \sigma_r \ge 0$$

and (2)

$$\sigma_{r+1} = \sigma_{r+2} = \ldots = \sigma_n = 0$$

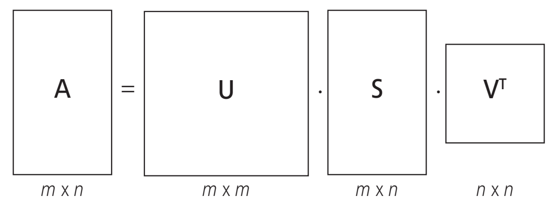
One of the important properties of SVD is that the rank *r* of matrix **A** is equal to the number of its nonzero singular values. Since it is generally the case in image compression applications that the singular values of a matrix decrease quickly with increasing rank, it is possible to compress the matrix data by eliminating the low singular values or the higher ranks.

Before applying SVD to image compression, original image matrix $\mathbf{A} = \mathbf{X}^T$ has to be appropriately transformed. First, we create a new *n* x *m* matrix **Z**:

$$\mathbf{Z} = \frac{1}{\sqrt{n-1}}\mathbf{X}^T \qquad (3)$$

where we subtracted the row average from each entry to ensure zero mean across the rows. Thus, the matrix Z has columns with zero mean. Next, we construct the following *m* x *m* matrix:

$$\mathbf{Z}^T\mathbf{Z} = \left(\frac{1}{\sqrt{n-1}}\mathbf{X}^T\right)^T\left(\frac{1}{\sqrt{n-1}}\mathbf{X}^T\right) = \frac{1}{n-1}\mathbf{X}\mathbf{X}^T \qquad (4)$$



$$\mathbf{S} = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \sigma_{r+1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & \sigma_n \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}$$

Figure 1: Singular value decomposition (SVD, left) and matrix S (right).

It can be shown (Richardson, 2009) that $\mathbf{Z}^T\mathbf{Z}$ is equal to the covariance matrix of **A**, $\mathbf{C}_A$. Here we are able to link the eigenvalue decomposition method of PCA to SVD: the principal components of **A** are the eigenvectors of $\mathbf{C}_A$.

Therefore, if we compute a SVD of the matrix **Z**, the principal components of **A** will be the columns of the orthogonal matrix **V** (Shlens, 2003).

1.2 Image compression

Due to the large amount of data encountered in modern still- and video image applications, compression is an almost inevitable step towards the reduction of the image storage space or transmission time requirements. While lossy compression standards such as JPEG offer a tradeoff between the image quality and the file size, in lossless compression methods – run length encoding, LZW, Chain codes and others – no information reduction takes place, but the image file size may be prohibitively large.

Three types of redundancy typically exist in a digital image and are exploited in image compression techniques: coding-, interpixel- (spatial) and psychovisual redundancy. In practice, both non-redundant and relevant information are discarded to achieve a higher compression ratio (CR) and reduced file size. CR can be computed according to the following formula:

$$\text{CR} = \frac{\text{size of original image}}{\text{size of compressed image}} \qquad (5)$$

While in lossless (reversible) compression typical CR is approximately 2:1 for natural scenes and somewhat higher for document images, this ratio is much higher with lossy compression schemes, such as lossy JPEG.

1.3 Face recognition

Face recognition is a computer vision topic of active current research with a broad range of applications, including crowd surveillance, personal identification, video content indexing, and entrance security. Individual steps involve segmentation – detection – of faces from a scene, feature extraction from the face region followed by either recognition or verification. The main idea of using PCA for face recognition is to transform face images into a small set of characteristic feature images, known as eigenfaces, which are the PCs of the training set' face images. Recognition is performed by projecting a new image into the subspace spanned by the eigenfaces ("face space") and then classifying the face by comparing its position in face space with the positions of known individuals (Turk and Pentland, 1991).

2. Methods and Materials

2.1 Image compression

The workflow of our image compression experiment is presented in Figure 2. Three 24-bit RGB images – referred to as "Hedgehog", "Man" and "Building" – taken with Canon digital camera EOS 400D were converted to 8-bit grayscale images, trimmed to a square region and downsampled to 1024x1024 pixel size. Each image was subject to PCA procedure. The degree of image reconstruction, i.e. image quality, when using an increasing number of PCs was assessed using structural similarity (SSIM) index (Wang et al., 2004) along with the conventional image quality metric peak signal-to-noise-ratio (PSNR). Corresponding compression ratios (CR) were also computed. All image processing routines and computations were performed with MATLAB®. PCA implementation was based on this paper (Richardson, 2009) and consisted of the following main steps:

- Loading an image into a computer memory and displaying it
- Computing the image matrix' row mean and obtaining matrix **X**

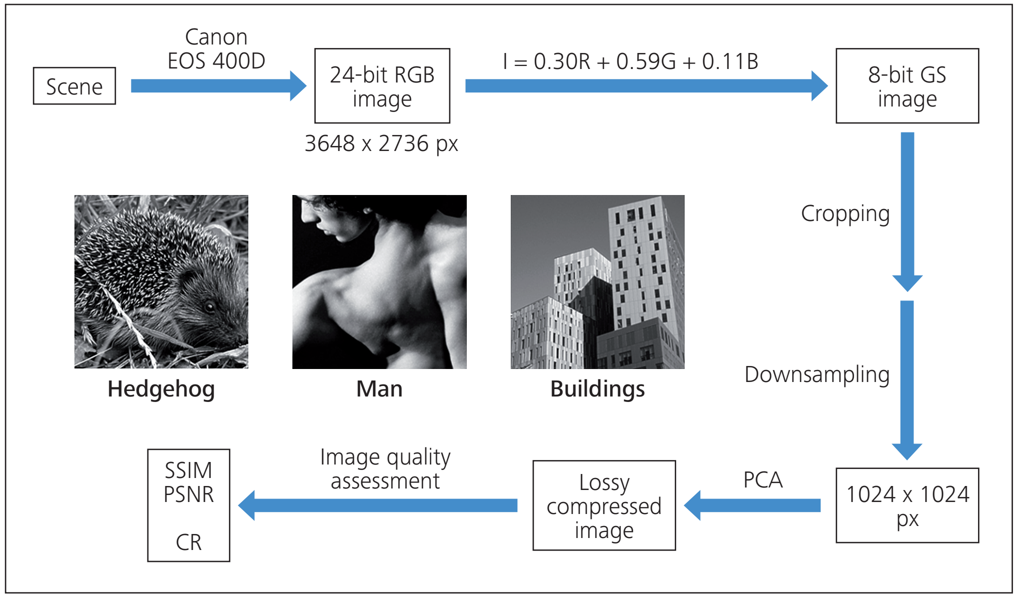


Figure 2: Image compression workflow.

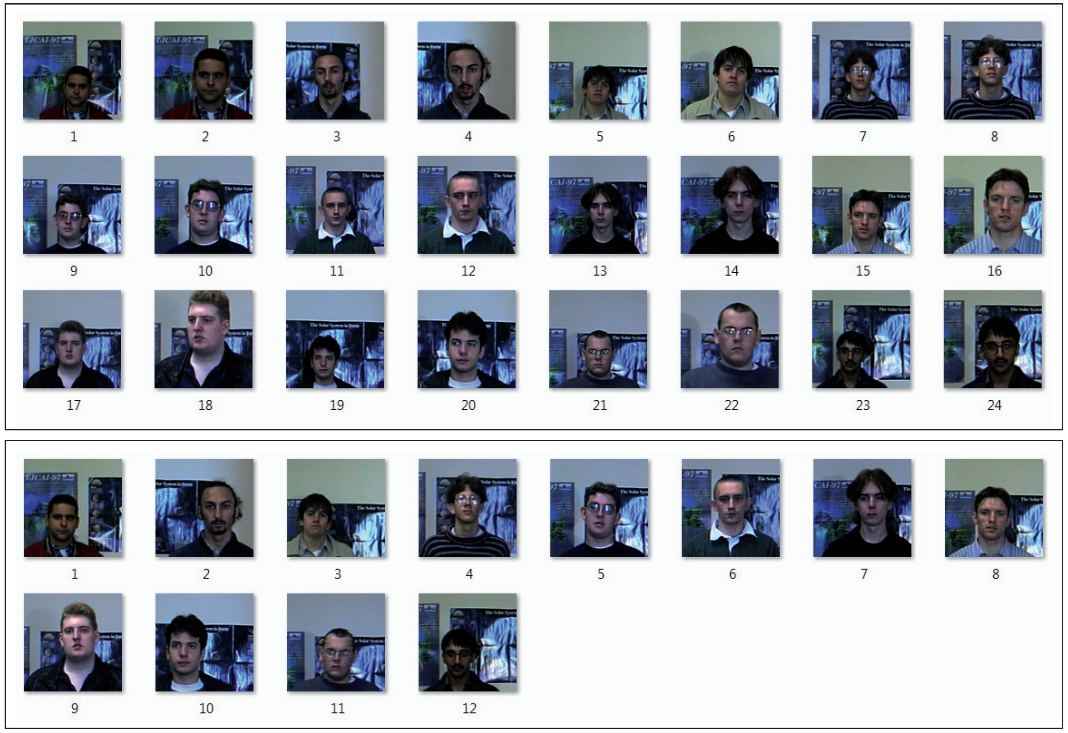


Figure 3: Face96 training (top) and testing (bottom) sets used in the face recognition investigation.

- Creating matrix **Z** (Eq. 3) and performing SVD of **Z**
- Extracting first few (e.g. 10) PCs
- Projecting data onto PCs
- Converting data back to original basis
- Displaying the compressed image

2.2 Face recognition

First, we downloaded three publicly-available human face databases – Faces94, Faces95 and Faces96 (Spacek, 2008). They differ mainly in terms of the head position and scale, background pattern and lighting conditions as discussed in the Results section. Within each database, a dataset consisting of 36 images of 12 people was created of which 24 images were selected for training and 12 for testing purposes (Figure 3). Every person was therefore represented three times, twice in the training set and once in the test set. Recognition rate (RR), i.e. the percentage of human faces correctly recognized/classified by the software, was computed.

We modified the face recognition MATLAB® code written by A. Omidvarnia (MatlabCentral, 2007) that is based on the Eigenface technique described e.g. in (Belhumeur et al., 1997). The code contains three functions: *CreateDatabase*, *Recognition* and *EigenfaceCore*. The first one – *CreateDatabase* – reshapes all of the training set 2D images into 1D column vectors and puts these in a row to construct a 2D matrix "T". Next, the *EigenfaceCore* function is used to determine the most discriminating features among the images of faces. Three quantities – mean of the training dataset, eigenvectors of the training dataset' covariance matrix (= eigenfaces) and matrix of the centred image vectors – are computed from the Matrix "T". Finally, function *EigenfaceCore* is called to compare two faces (one from training and one from testing datasets) by projecting the images into facespace and measuring their Euclidean distance. The output is the number of the recognized image in the training set.

3. Results and Discussion

3.1 Image compression

As Figure 4 demonstrates, the higher the number of PCs used in the reconstruction of the "Hedgehog" image, the better its quality – as indicated by progressively higher SSIM and PSNR values – but, on the other hand, the lower the compression ratio and, consequently, the file reduction gain. The corresponding SSIM maps provide additional information on the location of image regions with poor (black) or good (white) quality with respect to the original image. It can be seen that the most pronounced differences are in the facial area, which corresponds very well to the visual, subjective, perception of these images.

Table 1 shows that SSIM index, unlike CR, depends on the type of image under investigation. With the same number of PCs used for reconstruction, the image with the largest amount of details, i.e. high-frequency components – in our case the "Hedgehog" image – results in the lowest quality and, consequently, the lowest SSIM index when compared to the other two images. This finding is again in a very good agreement with the human-based perception of these images.

Table 1: Comparison of the three 1024 x 1024 test images in terms of the number of PCs, compression ratio, SSIM index and PSNR

		SSIM index			PSNR		
No. of PCs	CR	Hedgehog	Buildings	Man	Hedgehog	Buildings	Man
10	48.8	0.34	0.64	0.77	14.6	21.5	26.0
20	25.0	0.50	0.76	0.88	15.7	24.4	30.9
30	16.8	0.62	0.85	0.92	16.7	26.4	33.6
40	12.6	0.70	0.89	0.95	17.6	27.7	35.5
50	10.1	0.76	0.91	0.96	18.4	28.7	37.0
…	…	…	…	…	…	…	…
100	5.1	0.91	0.97	0.99	21.7	32.3	41.2
200	2.6	0.98	0.99	1.00	27.0	37.2	45.3
512	1.0	1.00	1.00	1.00	41.6	48.8	53.9

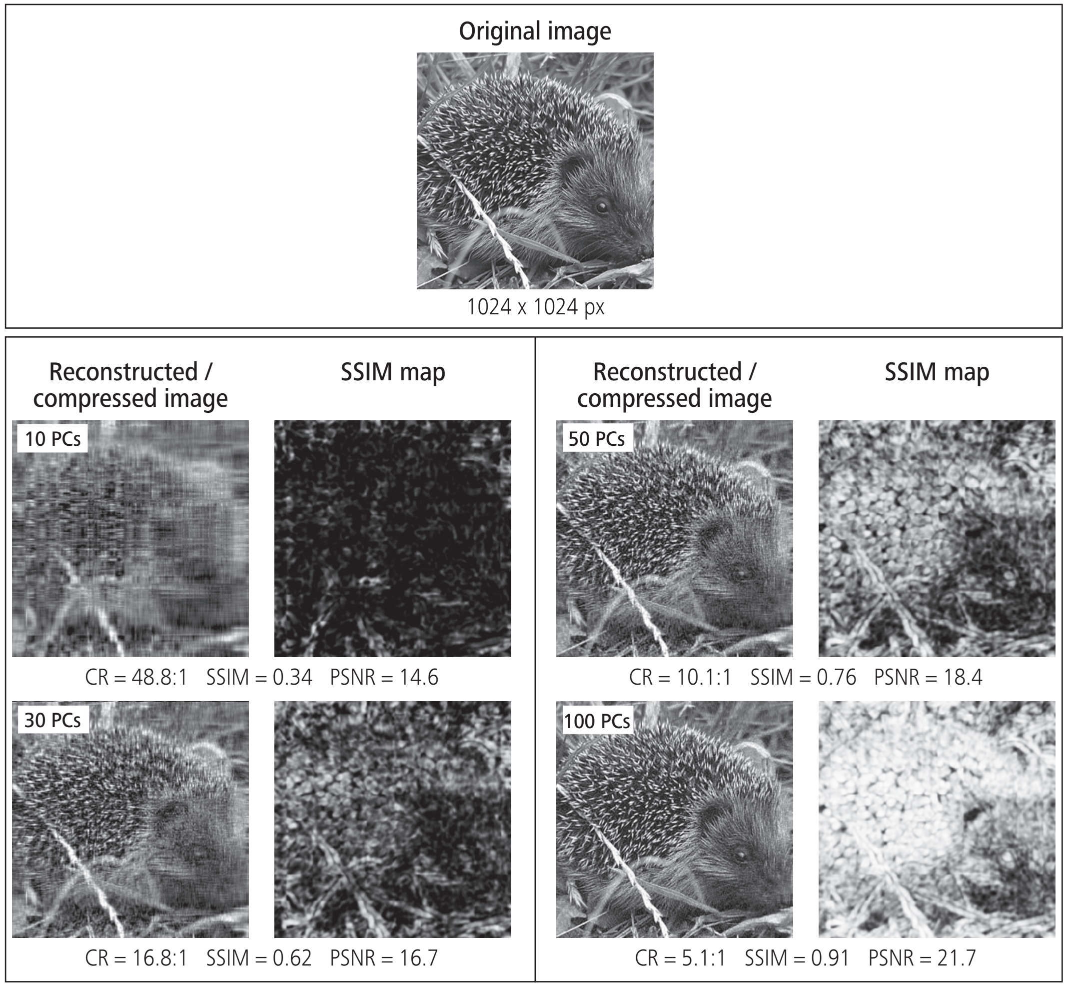


Figure 4: Results of the "Hedgehog" image compression using 10, 30, 50 and 100 PCs.

3.2 Face recognition

RR values for the three investigated datasets are given in Table 2. In case of the Faces94 dataset, 12 female subjects sat at a fixed distance from the camera and were asked to speak, during which photographs were taken. Consequently, their facial expression varied somewhat, while there were only minor variations in the background, position of head or face, as well as lighting conditions. As a result, RR was very high, with only one person out of 12 being falsely recognized.

Both Faces95 and Faces96 dataset images were generated in a similar manner: for each of the 12 male persons, a sequence of three digital images was produced. During the sequence, the subject took one step forward towards the camera, so that the movement introduced significant head scale variations between the images of the same individual. The difference between the Faces95 and Faces96 datasets was in the image background: with the former dataset it was monotonous, unstructured (red curtain), while it was complex and highly structured (glossy posters) with the latter one (see Figure 3). Apparently, it was this lack of image background structure that was responsible for an extremely low RR in case of Faces95 images; with a

Table 2: Recognition rate for the three datasets used in the face recognition study.

	Dataset	Faces94	Faces95	Faces96
Recognition Rate (RR)	Ratio	11/12	2/12	12/12
	Percentage	92	17	100

highly patterned background of Faces96 images, RR was 100%. The results will be presented in more detail in a forthcoming article.

4. Conclusions

It is a well-known fact that the choice of the best compression method always depends on the actual application. Among the most important factors to consider are input image characteristics (data type, previous processing), desired output (compressed) image quality, compression scheme computational complexity and presence of artifacts (blocks, blur, edge noise). PCA is one of the many possible approaches used to reduce image file size at the expense of its perceptual quality.

PCA has also been successfully implemented in advanced machine vision/pattern recognition applications such as face recognition. As our study demonstrated, image/scene characteristics, e.g. background pattern or human face scale, can decisively influence the effectiveness of a face recognition system.

5. References

- Abdi, H./Williams, L.J. (2010): "Principal component analysis", Wiley Interdisciplinary Reviews: Computational Statistics, vol. 2 (4), 433–459.
- Belhumeur, P.N./ Hespanha, J.P./ Kriegman, D.J. (1997): "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19 (7), 711-720.
- Cao, L. (2012): "Singular Value Decomposition Applied To Digital Image Processing", URL: http://www.lokminglui.com/CaoSVDintro.pdf (last access: 3.5.2013).
- MatlabCentral (2007): "PCA-based Face Recognition System", URL: http://www.mathworks.com/matlabcentral/fileexchange/17032-pca-based-face-recognition-system (last access: 3.5.2013).
- Richardson, M. (2009): "Principal Component Analysis", URL: http://people.maths.ox.ac.uk/richardsonm/SignalProcPCA.pdf (last access: 3.5.2013).
- Shlens, J. (2003): "A Tutorial on Principal Component Analysis; Derivation, Discussion and Singular Value Decomposition", URL: http://www.cs.princeton.edu/picasso/mats/PCA-Tutorial-Intuition_jp.pdf (last access: 3.5.2013).
- Spacek, L. (2008): "Collection of Facial Images", URL: http://cswww.essex.ac.uk/mv/allfaces (last access: 3.5.2013).
- Turk, M.A./Pentland, A.P. (1991): "Face Recognition Using Eigenfaces", Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 586-591.
- Wang, Z./Bovik, A.C./Sheikh, H.R./Simoncelli, E.P. (2004): "Image quality assessment: From error visibility to structural similarity", IEEE Transactions on Image Processing, vol. 13 (4), 600-612.

(received for the first time: 01-10-12)

Aleš Hladnik

Dr., Ass. Prof.,
Chair of Information and Graphic Arts Technology,
Faculty of Natural Sciences and Engineering,
University of Ljubljana,
Slovenia

ales.hladnik@ntf.uni-lj.si