Dimension reduction of images using neural networks

Master thesis

J.A. Spierenburg

July 1997

Supervisors:

Prof. Dr. J.N. Kok
Leiden University
Department of Computer Science

Dr. D.P. Huijsmans
Leiden University
Department of Computer Science

Abstract

Dimension reduction can be seen as the transformation from a high order dimension to a low order dimension. An example is the reduction of a cube in 3D (xyz) to a square in 2D (xy), to a point in 1D (x). The main goal of dimension reduction is to concentrate on vital information while redundant information can be discarded. Various ways are developed to reduce dimensions. Reduction methods can be distinguished into linear and non linear dimension reduction. In this thesis we shall present a linear and some non linear dimension reduction strategies which are mainly based on neural networks. A reduced representation of data makes the data easier to handle. With this in mind, we have developed some applications that make use of reduced representation of data. The main application is face recognition which uses a non linear dimension reduction strategy to reduce a face image into no more than say one to five vital values. These values are then used to classify the face image.

Acknowledgments

This Master thesis was developed during February - July 1997 at the department of Computer Science at Leiden University. I would like to thank my supervisors Prof.J.N. Kok and Dr.D.P. Huijsmans for their support and suggestions during this period. The various demonstrations initiated by Prof.J.N Kok were instrumental in improving my presentation skills.
I am grateful to Drs.M. Lamers for his enthusiasm and supply of the numerous papers in the early stages. Further, I would like to mention my colleagues at Leiden University who showed their interest and who helped me to relax when needed. Finally, I want to mention my family, especially my parents, who supported me and backed me up during my studies.

Joost Alexander Spierenburg
July 1997

Contents

Chapter 1

Introduction

In this master thesis research on dimension reduction of images will be described.

Generating a good data representation is of great importance before any further processing can be done. Most applications suffer from high dimensional input data; reduction of the data to a lower dimension makes the data easier to handle. Thus it is important to extract the principal features of the data and remove redundant information in such a way that structural relationships of the data distribution are preserved. This so-called "mapping method" is also known as dimension reduction. There exists a variety of ways to accomplish dimension reduction. We can classify dimension reduction methods in linear and non-linear methods. Principal Components Analysis (PCA) is a well known method for linear dimension reduction. Vector Quantization (VQ) and *five layer* auto-associative Neural Networks (5LN) are examples of non-linear dimension reduction.

We apply the dimension reduction on face images, which are to be considered as non-linear. We use face images because one can then compute and visualize the performance difference between linear and non-linear methods. Another advantage of using face images is that face recognition can be carried out. This recognition of faces is accomplished in the reduced dimension obtained by dimension reduction.

Suppose a face image has dimension 10×10, then one can see this face image as one vector of length 100 *(concatenate the rows of the face image)*. This vector of length 100 is also **one** point in dimension hundred. The PCA method adjusts the origin in dimension 100 and places the new origin at the center of all the face image points. Then the axes are realigned in such a way that the distances between all face image points and the axes are minimized. The VQ method is based on a neural network, which is a system that is capable of learning patterns. The VQ tries to find correlations among the face images. The correlated face images found by the VQ are gathered in clusters. When these clusters are found, one can apply the PCA method within each cluster. The 5LN method is based on another type of neural network. This type of network tries to learn all of the individual face images. The network system must internally store relevant *(correlated)* information on all of the face images

In this research we have the following objectives [1]:

- Compare performance differences computed for the following dimension reduction methods
  - PCA based on the Karhunen-Loève transform
  - Vector quantization with Euclidean partitioning
  - Vector quantization with gradient partitioning *This is an own developed method that uses more information when clustering the face images*
  - Five layer feed-forward neural network
- Real time face recognition using a reduced dimension
- Morphing
- Gender recognition

We reduce the image dimension (our images are of dimension $92 \times 112 = 10304$) to the fifth dimension. The morph as well as gender recognition also use this reduced fifth dimension. The use of only 5 values in representing an image is explained in the chapter 'Five layer networks'.

This thesis is structured as follows. In Chapter 2 it is explained how the Karhunen-Loève transform can be applied to reduce the face images. In Chapter 3 an overview is given of neural networks and all related aspects such as the learning of a network. In the next Chapter 4 it is described how to use five layered networks to reduce the face images. In Chapter 5 an introduction to vector quantization neural networks is given. In Chapter 6 the use of the vector quantisizer which is based on the Euclidean distance is described and Chapter 7 is also dedicated to a vector quantisizer, but this time to a quantisizer using more information on the images when reducing the dimension. The obtained results with the various dimension reduction methods are handled in Chapter 8. The applications morphing, face recognition and gender recognition which are all based on dimension reduction are

[1] The research is mainly inspired by the paper by Kambhatla & Leen [1]

described in respectively Chapter 9, 10 and 11. Finally, the conclusions and recommendations for future research are given in the last Chapter 12.
An overview of the handled dimension reduction methods and applications in this thesis is shown in Figure 1.1.

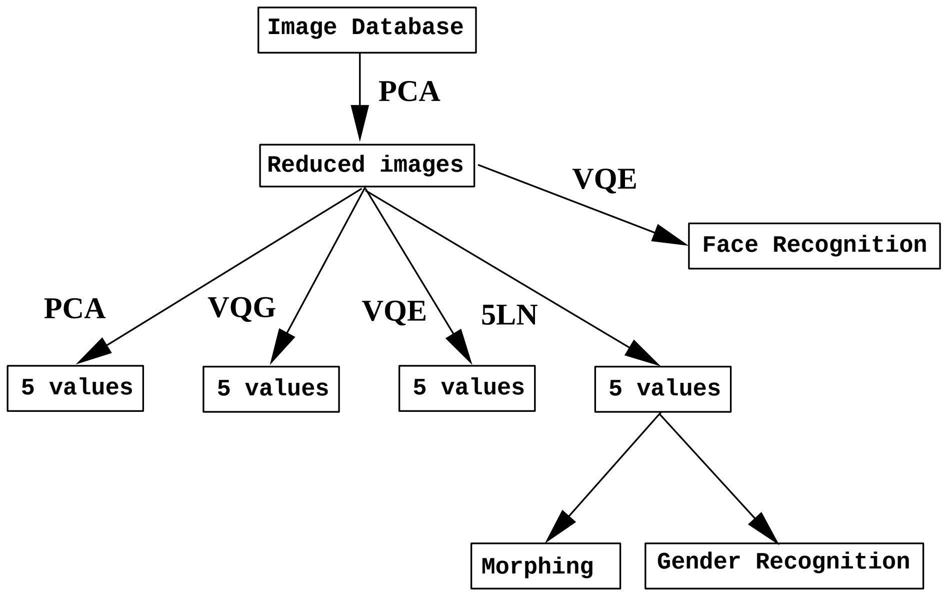

Figure 1.1: Dimension reduction tree and applications

Chapter 2

Karhunen-Loève transform

The KLT *(this transform also is commonly referred to as the eigenvector, principal component or Hotelling transform)* is a well known linear method for dimension reduction. We use the KLT based on the description by Turk & Pentland [2, 3, 4, 5, 6, 7, 8] and Gonzalez & Woods [9]. The goal of the KLT is to create a feature space that spans the significant variations among the face images. The significant features are known as "eigenfaces," because they are the eigenvectors (*principal components*) of the set of face images. The eigenvectors are ordered, such that the first eigenvector accounts for the largest variance and the last eigenvector for the smallest variance among the face images. Each $N \times N$ face image can be regarded as one point in the N^2 dimensional space.

Let the set of face images be $\Gamma_1, \Gamma_2, \Gamma_3, \ldots, \Gamma_M$. The average image of the set is defined by $\Psi = \frac{1}{M}\sum_{n=1}^{M}\Gamma_n$. Each face image differs from the average by the vector $\Phi_i = \Gamma_i - \Psi$. The covariance matrix C of the face image set is defined as

$$C = \frac{1}{M}\sum_{n=1}^{M}\Phi_n\Phi_n^T \tag{1}$$

Because C is real and symmetric, finding the M orthonormal vectors u_n which describe the face images best is possible. Element c_{ii} is the variance of Φ_i and element c_{ij} is the covariance between elements Φ_i and Φ_j. If elements Φ_i and Φ_j are non-correlated, their covariance will be 0 and,therefore $c_{ij} = c_{ji} = 0$.

The kth vector u_k is chosen such that

$$\lambda_k = \frac{1}{M}\sum_{n=1}^{M}\left(u_k^T\Phi_n\right)^2 \tag{2}$$

is maximum, subject to

$$u_l^T u_k = \begin{cases} 1, & \text{if} \quad l = k \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

in which the eigenvectors are u_k, and the eigenvalues are λ_k of the covariance matrix C in (1).

The eigenvectors v_i of C are used to create the eigenfaces u_l of our face images set:

$$u_l = \sum_{k=1}^{M} v_{lk}\Phi_k \quad l = 1 \dots M \tag{4}$$

A face image Γ is transformed into its eigenface components (*projected into "face space"*) by a simple operation,

$$\omega_k = u_k^T(\Gamma - \Psi) \tag{5}$$

for $k = 1, \dots, M'$. The eigenfaces span an M'-dimensional subspace of the original N^2 image space. See Figures 2.1 and 2.2 for examples of a transformation from points to eigen/face-space. The weights form a vector $\Omega^T = [\omega_1, \omega_2 \dots \omega_{M'}]$ which describes the contribution of the eigenfaces in forming the face-space projection of the input face image. So each image has it's own weight vector of length M' which describes it's contribution to each eigenface. The "inverse projection" or reconstruction, from M' to M can be reconstructed from (5) which results in

$$\Gamma = u_k^T w_k + \Psi \quad \text{for} \quad k = 1 \dots M' \tag{6}$$

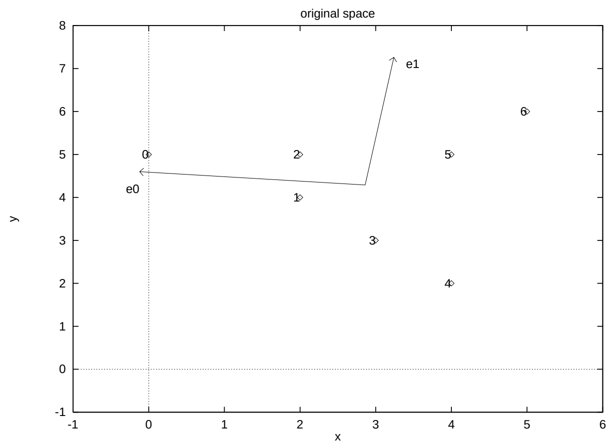

Figure 2.1: Original points

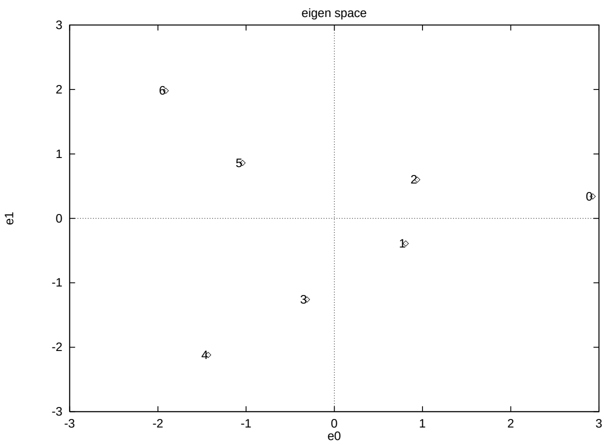

Figure 2.2: The points of Figure 2.1 in eigen/face-space