A Gradient-Based Eigenspace Approach to Dealing with Occlusions and Non-Gaussian Noise

Horst Wildenauer, Thomas Melzer
PRIP
Vienna University of Technology, Austria
{wilde, melzer}@prip.tuwien.ac.at

Horst Bischof
ICG
Technical University Graz, Austria
bischof@icg.tu-graz.ac.at

Abstract

In the recent literature, gradient-based (filtered) eigenspaces have been used as a means to achieve illumination insensitivity. In this paper, we show that filtered eigenspaces are also inherently robust w.r.t. (non-Gaussian) noise and occlusions. We argue that this robustness stems essentially from the sparseness of representation and insensitivity w.r.t. shifts in the mean value. This is also demonstrated experimentally using examples from the field of object recognition and pose estimation.

1. Introduction

Since its inception in the early 1990s, the eigenspace approach to object recognition has received much interest in the vision and object recognition community and is still a very active research topic. It has successfully been employed in various applications such as face recognition [7], illumination planning [4], visual inspection and even visual servoing [5].

The key idea is to represent each object as collection of labelled \((m \times n)\)-images, which show the object under envisaged viewing conditions. In the eigenspace approach, PCA is performed on a set of \(N\) mean-normalized training images

\[
\mathbf{X} = (\bar{x}_1, \ldots, \bar{x}_N) = (\mathbf{x}_1 - \bar{m}, \ldots, \mathbf{x}_N - \bar{m})
\]

(whereby \(\bar{m}\) is the estimated mean and \(\mathbf{x}_i, \bar{m} \in \mathbb{R}^{(m \times n)}\)) to find a low-dimensional representation of the object. If the eigenvectors of the estimated covariance matrix \(\frac{1}{N} \mathbf{X} \mathbf{X}^T\) are given by \((\mathbf{e}_1, \ldots, \mathbf{e}_N)\) (whereby we assume that the \(\mathbf{e}_i\) are sorted in decreasing order according to their eigenvalues), each mean-normalized input image \(\mathbf{x}\) is approximated as linear combination of the first \(k << N\) eigenvectors (eigenimages):

\[
\mathbf{x} \approx \bar{x} = \sum_{i=1}^{k} c_i \mathbf{e}_i. \tag{2}
\]

The coefficients \(c_i = (c_1, \ldots, c_k)\) of the linear expansion of \(\mathbf{x}\) in terms of \((\mathbf{e}_1, \ldots, \mathbf{e}_k)\) are the eigenspace representation of \(\mathbf{x}\).

The purpose of this feature extraction step is twofold: First, it reduces the amount of data needed to represent a single observation (i.e., image) by a factor \(1/k\). Second, PCA identifies the directions that explain best the variability between different object views (in the correlation sense); indeed, it can be shown that among all orthonormal transformations, PCA is optimal in the sense that it minimizes, in the mean square sense, the expected reconstruction error between the original signal \(\mathbf{x}\) and the signal \(\bar{x}\) reconstructed from its low-dimensional representation \(\mathbf{e}\).

Traditionally, the coefficients \(c_i\) in Eq. 2 are obtained as the projections of \(\mathbf{x}\) onto the first \(k\) eigenimages. Although this approach is conceptually simple and elegant, it can not deal with noise and occlusions. Similarly, it can not handle changes in the object’s appearance due to varying illumination, unless these illuminations effects are explicitly incorporated into the eigenspace model (however, this would require that each object pose is acquired under all possible illumination conditions, which is not feasible in most cases).

In this paper, we will discuss an eigenspace approach based on gradient filters, which is not only insensitive to varying illumination, but also inherently robust w.r.t. noise and occlusions. In section 2, we will briefly review the relevant theory, while in section 3, we will demonstrate the
2. Gradient-Based Eigenspaces

The fact that edges are relatively insensitive to the effects of varying illumination has motivated several different implementations of gradient-based eigenspaces (e.g., [8, 2]). Our approach, which was introduced in [1], is based on the observation that a filtered eigenspace representation can be obtained by filtering the original eigenimages rather than performing PCA on filtered versions of the training images: Let \( f \) be a linear filter and let \( * \) denote the convolution operator. Since convolution is a distributive (linear) operation, we have (cf. Eq. 2)

\[
(f * \hat{x}) \approx \sum_{i=1}^{k} c_i (f * e_i).
\]

(3)

Note that Eq. 3 implies that the coefficient vector \( e \) in Eq. 2 can be retrieved from the filtered input image and the filtered eigenimages. However, since the filtering process will, in general, not preserve the orthogonality of the eigenimages, the coefficients can no longer be obtained as projections of the input image onto the eigenimages. Instead, we compute them by solving (in the least squares sense) the overdetermined system of \((m \times n)\) linear equations Eq. 3. In our experiments, we used a set of six steerable gradient filters [6]. Since, in our approach, the coefficients are not affected by the choice of the filter, we can combine the resulting six sets of equations into one large system consisting of \((m \times n) \times 6\) equations and solve them simultaneously.

In [1], it was shown that filtered PCA, combined with the robust approach discussed in [3], is insensitive to illumination effects and can tolerate significant levels of noise and occlusion. In this paper, we focus on the properties of filtered PCA itself and show that a gradient-based eigenspace representation, in addition to being illumination insensitive, is also inherently robust w.r.t. noise and occlusions without using the robust algorithm.

3. Experiments

We have extensively evaluated our algorithm on a database of 5 objects (Fig. 1), whereby each object is represented by 72 views; these views were acquired in 5 degree-increments by placing the object on a turntable. The resulting images were of size \(128 \times 128\) with 256 different greyvalues ranging from 0 to 1. Only 36 views were used to build the eigenspace; unless stated otherwise, the eigenspace dimension (number of eigenvectors) was 30. As performance criteria, we used the coefficient error (L2-

\[\text{distance between true and recovered coefficients}\]) and the recognition rate.

**Salt and Pepper Noise** To have a systematic performance evaluation we took the total set of 300 images and added uniformly distributed salt and pepper noise ranging from 0 to 90 percent (see Fig. 2(a)). This experiment was repeated 10 times. Fig. 3(a) shows the average coefficient error and its standard deviation for the filtered and the standard approach. In Fig. 3(b) the obtained recognition rates are depicted.

**Occlusions** In this experiment, we used again all 300 images and distorted them with homogeneous occlusions with greyvalues ranging from 0 to 1 (maximum brightness). The size of the occlusions was increased from 0 to 90 percent of the object’s size (Fig. 2(b)). Fig. 4 shows the coefficient errors obtained with the filtered and standard method, while the corresponding recognition rates can be seen in Fig. 5. Note the high sensitivity of the standard approach w.r.t. the greyvalue of the occlusion.

![Figure 1. The 5 objects used in the experiments.](image1.png)

![Figure 2. Examples of noise and occlusions used in the experiments.](image2.png)
Figure 3. Influence of salt and pepper noise on the standard and the filtered eigenspace.

Figure 4. Comparison of coefficient errors for occluded images.
4. Discussion

In the previous section, we have shown experimentally that our version of filtered PCA compares favorably to the standard eigenspace approach. The experiments indicate that filtered PCA is not only illumination-insensitive (as discussed in [1]), but also inherently robust w.r.t. non-Gaussian noise and occlusions. This robustness stems essentially from the following two properties of filtered PCA:

Sparseness of Representation The gradient-filtered eigenimages and the gradient-filtered input images encode high-frequency information. However, unless the appearance of the training objects is dominated by high-frequency texture or structure elements, the number of such edge pixels will be small compared to the size of the eigenimages. Consequently, distortions in the input image due to occlusions or noise are likely to result in spurious gradients that will not coincide with the true edges encoded in the eigenspace model. Due to the sparse nature of gradient maps, such spurious gradients will be orthogonal (or almost orthogonal) to the eigenspace and thus will have no (or only small) influence on the solution vector \( \mathbf{c} \).

Insensitivity w.r.t. the Mean It has been standard practice in the eigenspace approach to rescale the input image \( \mathbf{x} \) to unit length before subtracting the mean-image. This makes the approach insensitive to changes in the overall brightness level and ensures that the pre-stored mean-image \( \mathbf{m} \) and the actual input image are of comparable magnitude. However, uneven illumination, saturation effects or large occlusions will pull away the computed scale factor \( s = \frac{1}{\| \mathbf{x} \|} \) from its true value, which can make the subtraction of \( \mathbf{m} \) practically meaningless (for too large \( s \)) or, worse, make the relative magnitude of the input image so small that \( \mathbf{x} = \mathbf{x} - \mathbf{m} \) is dominated by \( \mathbf{m} \), and not by the actual observation.

Filtered PCA is far less susceptible to such mean-effects than other eigenspaces approaches. This is due to the fact that gradient filters basically compute the difference of gray-values in a small local neighborhood; the mean, however, will be relatively homogeneous in such small neighborhoods and therefore will be canceled by the subtraction.

References