Robust continuous subspace learning and recognition*

Danijel Skočaj and Aleš Leonardis
Faculty of Computer and Information Science
University of Ljubljana
Tržaška 25, 1001 Ljubljana, Slovenia
danijel.skocaj@fri.uni-lj.si

Abstract

Učenje je trajen in nepretrgan proces. Znanje o predmetu nadgradimo vsakokrat, ko se z njim ponovno srečamo. Učenje mora biti prav tako robustno, saj se moramo učiti tudi iz nepopolnih podatkov. V tem članku predstavljamo metodo za vizualno učenje na osnovi izgleda, ki upošteva ta dejstva. Učenje je implementirano kot nepretrgan inkrementalen proces, ki neprestano robustno nadgrajuje že zgrajene predstavitve.

1 Introduction

Learning of object and scene representations is an essential part of any cognitive vision system. In the real world, learning is usually continuous, never-ending process. After we have acquired some knowledge about an object, we continuously update this knowledge every time when a new instance, or a new view of the object is encountered. Thus, in the context of cognitive vision, methods for updating already learnt representations are required. In addition, one can also expect that the images of objects and scenes are not always ideal and that they may contain noise or occlusions, thus requiring a robust learning algorithm. In this paper we present a method, which takes these considerations into account and builds the representations of objects and scenes using a robust and continuous approach.

Often, visual learning is approached as appearance-based modeling of objects and scenes, which is commonly realized using Principal component analysis (PCA). However, the standard PCA approach is usually performed in the batch mode, i.e., all training images are processed simultaneously, which means that all of them have to be given in advance. This is inadmissible in a continuous on-line scenario, where the images to be processed are obtained sequentially. We propose a novel method for incremental learning, which processes images sequentially one by one and updates the principal subspace accordingly. We also embed the method for robust recognition [3] in this incremental framework, resulting in a method for robust continuous learning.

The proposed algorithm presents a new approach to robustness in the appearance-based modeling using PCA. Several methods for robust recognition have already been proposed (see e.g., [3]). A few methods for robust learning have also been introduced [1, 6], however, all of them operate in batch mode. The proposed method blurs the border between the training and the recognition stage by building the representation incrementally in a robust manner. Therefore, this algorithm is perfectly well suited for continuous on-line scenarios.

The paper is organized as follows. First, we briefly outline the standard PCA approach and introduce the notation. Then, we propose a method for incremental learning. In section 4 we present a robust version of this method. Then, we present the experimental results. Finally, we summarize the paper and outline some work in progress.

2 Standard PCA

In this section we briefly outline the standard PCA and introduce the notation. Let $\mathbf{x}_i = [x_{1i}, \ldots, x_{Mi}]^\top \in \mathbb{R}^M$ be an individual image represented as a vector, and $\mathbf{X} = [\mathbf{x}_1, \ldots \mathbf{x}_N] \in \mathbb{R}^{M\times N}$ the matrix of all training images. For further processing, we create mean normalized data matrix $\hat{\mathbf{X}}$ by subtracting the mean image $\boldsymbol{\mu} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i$ from the input images. Principal axes are traditionally obtained by performing the eigendecomposition (or, similarly, the singular value decomposition) of the covariance matrix of the input data. They are denoted by $\mathbf{u}_i = [u_{1i}, \ldots, u_{Mi}]^\top \in \mathbb{R}^M$; $\mathbf{U} = [\mathbf{u}_1, \ldots \mathbf{u}_N] \in \mathbb{R}^{M\times N}$. The columns of $\mathbf{U}$, i.e., eigenvectors, are arranged in decreasing order with respect to the corresponding eigenvalues. Usually, only k, $k < N$, eigenvectors (those with the largest eigenvalues) are needed to represent $\hat{\mathbf{x}}$ to a sufficient degree of accuracy as a linear combination of eigenvectors: $\tilde{\mathbf{x}} = \sum_{i=1}^{k} a_i\mathbf{u}_i = \mathbf{U}\mathbf{a}$,

* This work was supported by EU project IST-2000-29375 Cognitive Vision Systems - CogVis.

where $\tilde{\mathbf{x}}$ denotes the approximation of $\hat{\mathbf{x}}$. A coefficient vector characterizing an image is traditionally calculated by a standard projection of the input image into the eigenspace: $\mathbf{a} = \mathbf{U}^\top \hat{\mathbf{x}}$.

3 Incremental PCA

Traditional batch methods for eigenspace learning compute the eigenvectors from all images in one step. Incremental methods, on the other hand, take the training images sequentially, and compute the new eigenspace from the subspace obtained in the previous step and the current input image.

In this paper we propose a novel method for incremental learning. Let us suppose that we have already built an eigenspace from the first n images. In the step $n+1$ we can calculate a new eigenspace from the *reconstructions* of the first n input images and a new image using the standard batch method. The computational complexity of such an algorithm would be prohibitive, since at each step we would have to perform batch PCA on a set of high-dimensional data. However, identical results can be obtained by using low-dimensional *coefficients* of the first n input images instead of their high-dimensional reconstructions, since coefficients and reconstructed images encompass the same visual variability, i.e., they are the same points represented in different coordinate frames. Since the dimension of the eigenspace is very small, this algorithm is computationally very efficient.

The summarized algorithm for updating an eigenspace looks as follows:

Input:
current mean value $\boldsymbol{\mu}^{(n)}$, current eigenvectors $\mathbf{U}^{(n)}$, current coefficients $\mathbf{A}^{(n)}$, new input image $\mathbf{x}$.[1]

Output:
new mean value $\boldsymbol{\mu}^{(n+1)}$, new eigenvectors $\mathbf{U}^{(n+1)}$, new coefficients $\mathbf{A}^{(n+1)}$, new eigenvalues $\mathbf{l}^{(n+1)}$.

Algorithm:

1: Project a new image $\mathbf{x}$ in the current eigenspace: $\mathbf{a} = \mathbf{U}^{(n)\top}(\mathbf{x} - \boldsymbol{\mu}^{(n)})$.

2: Reconstruct the new image: $\mathbf{y} = \mathbf{U}^{(n)}\mathbf{a} + \boldsymbol{\mu}^{(n)}$.

3: Compute the residuum vector: $\mathbf{r} = \mathbf{x} - \mathbf{y}$. $\mathbf{r}$ is orthogonal to $\mathbf{U}^{(n)}$.

4: Append $\mathbf{r}$ as a new basis vector: $\mathbf{U}_e = \begin{bmatrix} \mathbf{U}^{(n)} & \frac{\mathbf{r}}{||\mathbf{r}||} \end{bmatrix}$.

5: Determine the coordinates of the coefficients in the new basis: $\mathbf{A}_e = \begin{bmatrix} \mathbf{A}^{(n)} & \mathbf{a} \\ \mathbf{0} & ||\mathbf{r}|| \end{bmatrix}$.

6: Perform PCA on $\mathbf{A}_e$. Obtain the mean value $\boldsymbol{\mu}_s$, the eigenvectors $\mathbf{E}_s$ and the eigenvalues $\mathbf{l}_s$.

7: Optionally: Drop the least significant vector of the new basis: $\mathbf{U}_s = \mathbf{E}_s(:, 1:k)$.[2]

8: Project the coefficients to the new basis: $\mathbf{A}^{(n+1)} = \mathbf{U}_s^\top(\mathbf{A}_e - \boldsymbol{\mu}_s \mathbf{1}_{1\times n+1})$.[3]

9: Rotate the enlarged subspace $\mathbf{U}_e$ for $\mathbf{U}_s$: $\mathbf{U}^{(n+1)} = \mathbf{U}_e\mathbf{U}_s$.

10: Update the mean: $\boldsymbol{\mu}^{(n+1)} = \boldsymbol{\mu}^{(n)} + \mathbf{U}_e\boldsymbol{\mu}_s$.

11: New eigenvalues: $\mathbf{l}^{(n+1)} = \mathbf{l}_s(1:k)$.

The initial values of the mean image, the eigenvectors, and the coefficients can be obtained by applying the batch PCA on a small set of images.

It is worth noting, that this algorithm estimates identical principal subspace as the method proposed by Hall et al. [2]. However, the subspace is calculated in a *different way*. A significant advantage of our method is that it is able to treat different images differently, which enables to advance it into a weighted incremental method [5].

We will demonstrate the behavior of the proposed algorithm on a simple 2-D example. The 2-D input space contains 41 points shown as black dots in Fig. 1. The goal is to estimate 1-D principal subspace, i.e. the principal axis. The eigenspace is being built progressively. At each step one point (from the left to the right) is added to the representation and at each step the eigenspace is updated accordingly. Fig. 1 illustrates how the eigenspace evolves during this process. A current principal axis at every sixth step is depicted. The points, which were appended to the model at these steps, are marked with crosses. One can observe, how the origin of the eigenspace (depicted as a square) and the orientation of the principal axis are changing through the time by adapting to the new points which are coming into the process. In the end, the estimated eigenspace, which encompasses all training points, is almost equal to the eigenspace obtained using the batch method.

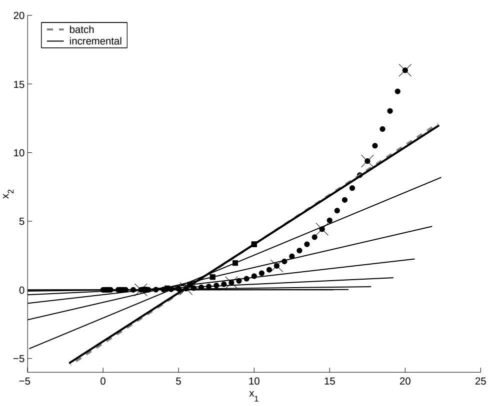


Figure 1: Incremental learning.

[1] Superscript denotes the step (time), which the data is related to. $\mathbf{U}^{(n)}$ denotes the values of the matrix $\mathbf{U}$ at the step n.

[2] Matlab notation. By discarding the least significant basis vector, we preserve the dimension of the eigenspace k.

[3] $\mathbf{1}_{m\times n}$ denotes a matrix of dimension $m \times n$ where every element equals to 1.

4 Robust approach

The incremental algorithm presented in the previous section suffers from similar drawbacks as the standard batch algorithm, including non-robustness to the non-gaussian noise. Here we show, how to embed the incremental method in the robust framework.

In the robust framework we are aware that images may contain outliers. As outliers we treat all pixels, which are not consistent with the information contained in other images. Since at each step we have a current model of the object or scene seen so far, we can detect outliers in the new image and replace them with the values which are yielded by the current model. This is achieved by projecting the new image into the current eigenspace in a robust manner. Instead of a simple projection, a robust procedure based on subsampling and hypothesize-and-select paradigm is used [3]. Coefficients are obtained only (mainly) from inliers, thus their reconstructions tend to the correct values also in outliers. Consequently, the reconstruction error in outliers is high, which makes their detection easier. Thus, to make the incremental learning robust, we first detect outliers in a new image, reconstruct their values and update the eigenspace with restored, outlier-free image using the algorithm described in the previous section.

To demonstrate the behavior of the the robust algorithm, we changed the values of the second coordinate of five points in our 2-D example drastically. Fig. 2 shows that these outlying points pull the origin and the orientation of the estimated principal axis in the wrong direction. On the other hand, the robust method sequentially detects the outlying coordinate values, replaces these values with the reconstructed values (shown as circles) and updates the eigenspace accordingly. The principal axis obtained using this approach is significantly closer to the optimal one shown in Fig 1.

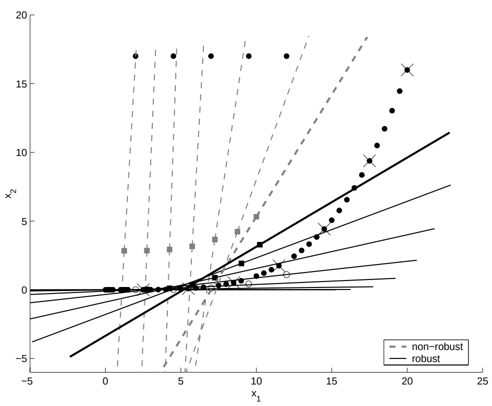


Figure 2: Robust incremental learning.

5 Experimental results

In this section we present the results of experiments in which the proposed method was used for face recognition. Face recognition was indeed one of the first applications of PCA in the computer vision [7]. However, our intention is not to present a new face recognition algorithm. We rather want to demonstrate the principle of continuous learning and to evaluate the proposed learning method.

For the testbed we used the ORL face database [4] consisting of 400 images (10 images of each of 40 subjects), rescaled to the size of 32×32 pixels. First, in training stage, we built an eigenspace from 40 images (one image of each subject; e.g., Fig. 3(a)) using the standard batch method. Some of the obtained eigenimages are depicted in Fig. 3(d). Then, we expressed each training image as a linear combination of eigenimages obtaining the coefficients in the eigenspace. In the recognition stage, we projected all other images (all together 360 training images; e.g. Fig. 3(b)) into this eigenspace and performed the recognition by searching for the closest projected training image.

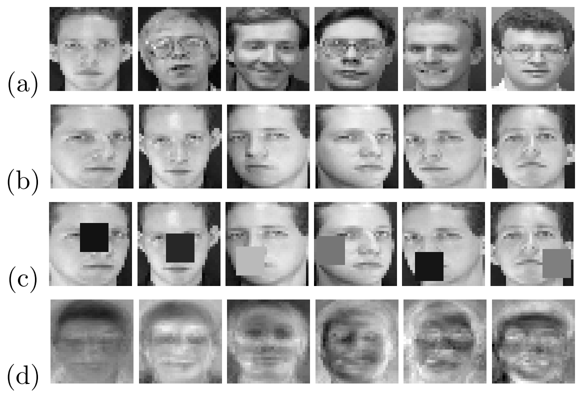


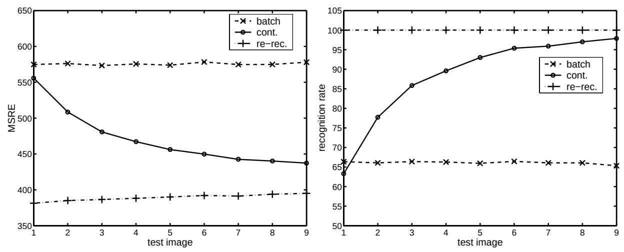
Figure 3: Six images from (a) training set, (b) test set, (c) occluded test set, (d) first six eigenimages.

In the framework of continuous learning, the border between training and recognition stage is blurred. During the recognition, the system also acquires new information, which should be built into the already acquired knowledge. To achieve this, we sequentially performed the recognition task on individual images. After an image was recognized, we updated the eigenspace with this image using the proposed incremental algorithm. After all test images were used for updating the representation, we performed re-recognition of all test images by projecting them into the obtained eigenspace to determine how well the representation was adapted.

We evaluated the results using two measures; the mean squared reconstruction error and the recognition rate. Fig. 4 depicts the accumulated results for nine sets of 40 test images of all subjects. First we plotted the mean MSRE and recognition rate of 40 test images (one of each subject), which were first presented to the system, and so on. In the end, the results are plotted for a set of images which were presented to the system lastly. In the case of the batch non-continuous learning method, the results for all nine sets of images are almost constant, since the representation remains the same through the entire recognition stage. In contrast, when the proposed continuous method is used, the results are improving through the time, since the newly encountered images update the representation. When the last image of each subject comes into the system, the eigenspace already contains nine projections of the every subject, which makes the recognition much more reliable.

The results are summarized in Table 1(a). The average results for all test images and the results for only the last set of test images are presented. One can observe that the proposed continuous method significantly outperforms the standard batch method. The results are particularly better in the end of the recognition process. This is also reflected in the excellent re-recognition results, where the proposed method correctly re-recognized all training images. In the case of the batch method, re-recognition results would be the same as recognition results, since no extra knowledge is incorporated in the representation during the recognition phase.

We also performed the same tests on the occluded test images (Fig. 3(c)). This task is rather more difficult, since the images, which are used by the continuous method for updating the representation, contain undesirable outliers. However, by using the proposed robust method we mainly detected and avoided the outliers and still achieved good results. The results are presented in Table 1(b) and have similar characteristics as the results for non-occluded images.



(a) (b)

Figure 4: (a) Mean squared reconstruction errors and (b) recognition rates of sets of test images.

6 Conclusion

In this paper we proposed a novel method for incremental learning of eigenspaces and embedded it in the robust framework resulting in an efficient incremental and robust method. Learning is extended in the recognition phase, enabling updating of the representation with the newly acquired information. The proposed method is suitable for continuous on-line learning, where the model is adapting to the input images as they arrive (SLAM principle).

The algorithm is flexible, since it is able to treat each image differently. Therefore, more recent (or more reliable, or more informative, or more noticeable) images can have stronger influence on the model then others. The principles of short-term and long-term memory, forgetting, and re-learning can be implemented and investigated. This principles are the subject of our current and future research.

Table 1: MSRE and recognition rate for (a) non-occluded, and (b) occluded test images.

		all images		last image	
		MSRE	RR	MSRE	RR
(a)	batch	575	66.1	578	65.3
	continuous	471	88.4	437	97.9
	re-recog.	389	100.0	395	100.0
(b)	batch	720	65.1	716	63.6
	continuous	706	79.7	707	89.2
	re-recog.	700	99.8	705	98.7

References

[1] F. De la Torre and M.J. Black. Robust principal component analysis for computer vision. In *ICCV'01*, pages I: 362–369, 2001.

[2] P. Hall, D. Marshall, and R. Martin. Merging and spliting eigenspace models. *PAMI*, 22(9):1042–1048, 2000.

[3] A. Leonardis and H. Bischof. Robust recognition using eigenimages. *Computer Vision and Image Understanding*, 78:99–118, 2000.

[4] F. Samaria and A. Harter. Parameterisation of a stochastic model for human face identification. In *2nd IEEE Workshop on Applications of Computer Vision*, 1994.

[5] D. Skočaj and A. Leonardis. Weighted incremental subspace learning. Submitted, 2002.

[6] D. Skočaj, H. Bischof, and A. Leonardis. A robust PCA algorithm for building representations from panoramic images. In *Computer Vision – ECCV 2002*, volume IV, pages 761–775, May 2002.

[7] M. Turk and A. P. Pentland. Eigenfaces for Recognition. *CogNeuro*, 3(1):71–96, 1991.