

Face Recognition in Ideal and Noisy Conditions Using Support Vector Machines, PCA and LDA

Miloš Oravec¹, Ján Mazanec¹, Jarmila Pavlovičová²,
Pavel Eiben¹ and Fedor Lehocki¹

¹ Dept. of Applied Informatics and Information Technology, ² Dept. of Telecommunications, Faculty of Electrical Engineering and Information Technology, Slovak University of Technology, Ilkovičova 3, 812 19 Bratislava
Slovak Republic

1. Introduction

In this chapter, we consider biometric recognition based on human face. Biometrics became frequently used in automated systems for identification of people (Jain et al., 2004) and huge interest is devoted to the area of biometrics at present (Jain et al., 2008; Shoniregun & Crosier, 2008; Ross et al, 2006).

Along with well-known methods such as fingerprint or DNA recognition, face image already opened new possibilities. Face recognition has been put into real life by many companies. It is already implemented in image organizing software (e.g. Google's Picasa: <http://www.deondesigns.ca/blog/picasa-3-5-adds-face-recognition/>), web applications (e.g. web photo albums http://picasa.google.com/intl/en_us/features-nametags.html) and even in commercial compact cameras (e.g. Panasonic Lumix). Passports contain face biometric data since 2006 (EU – Passport Specification, 2006).

In the area of face recognition, a class represents all images of the same subject (person). The goal is to implement an automated machine supported system that recognizes well the identity of a person in the images that were not used in a training phase (an initialization and training by representative sample of images precede an evaluation phase). Various applications are possible, e.g. automated person identification, recognition of race, gender, emotion, age etc. The area of face recognition is well described at present, e.g. starting by conventional approaches (PCA, LDA) (Turk & Pentland1991; Marcialis & Roli, 2002; Martinez & Kak, 2001), and continuing at present by kernel methods (Wang, et al., 2008; Hotta, 2008; Wang et al., 2004; Yang, 2002; Yang et al., 2005). Advances in face recognition are summarized also in books (Li & Jain, 2005; Delac et al., 2008) and book chapters (Oravec et al., 2008).

Our aim is to present complex view to biometric face recognition including methodology, settings of parameters of selected methods (both conventional and kernel methods), detailed recognition results, comparison and discussion of obtained results using large face database. The rest of this chapter is organized as follows: In section 2, we present theoretical background of methods used for face recognition purposes - PCA (Principal Component

Analysis), LDA (Linear Discriminant Analysis) and SVM (Support Vector Machines). Section 3 provides information about FERET database (FERET Database, 2001), since large image set from this database including total 665 images is used in our experiments. The face images are first preprocessed (normalization with respect to size, position and rotation and also contrast optimization and face masking). In Section 4, face recognition methods that are used in the rest of the chapter are discussed. We also propose methods utilizing PCA and LDA for extracting the features that are further classified with SVM and compare them to usual approaches with conventional classifiers. Section 5 presents results of recognition systems in ideal conditions. We show that proposed methods result in excellent recognition rate and robustness. Also behavior of presented methods is analyzed in detail and best settings for these methods are proposed. Section 6 is devoted to the influence of input image quality to face recognition accuracy. For this purpose, we use best parameter settings we obtained running 600 tests in ideal conditions. Gaussian noise, salt & pepper noise and speckle noise with various intensities are included. This enables to get insight into face recognition system robustness. Also equivalence of different types of noise from the recognition point of view is discussed.

2. Face Recognition Methods and Algorithms

We use different methods in our single-stage and two-stage face recognition systems: PCA (Principal Component Analysis), LDA (Linear Discriminant Analysis) and SVM (Support Vector Machines). The role of PCA and LDA falls into feature extraction. We use different classifiers that are in the form of both simple metrics and more complex SVMs.

2.1 Principal Component Analysis PCA

This standard statistical method can be used for feature extraction. Principal component analysis PCA (Turk & Pentland, 1991; Marcialis & Roli, 2002; Martinez & Kak, 2001; Haykin, 1994; Bishop, 1995) reduces the dimension of input data by a linear projection that maximizes the scatter of all projected samples. Let $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ be a set of N sample images of dimensionality n belonging to one of c classes $\{X_1, X_2, \dots, X_c\}$. Its covariance (total scatter) matrix is

$$\mathbf{S}_T = \sum_{k=1}^N (\mathbf{x}_k - \boldsymbol{\mu})(\mathbf{x}_k - \boldsymbol{\mu})^T \quad (1)$$

PCA transforms input images to new feature vectors

$$\mathbf{y}_k = \mathbf{W}^T \mathbf{x}_k \quad k = 1, 2, \dots, N, \quad (2)$$

where $\mathbf{W} \in \mathfrak{R}^{n \times m}$ is a transform matrix with orthonormal columns and $\boldsymbol{\mu} \in \mathfrak{R}^n$ is the mean image of all sample images. This yields also in dimensionality reduction ($m < n$). The scatter of the transformed feature vectors $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N\}$ is $\mathbf{W}^T \mathbf{S}_T \mathbf{W}$. In PCA, the projection \mathbf{W}_{opt} maximizes the determinant of the total scatter matrix of the projected samples

$$\mathbf{W}_{opt} = \arg \max_{\mathbf{W}} (\det(\mathbf{W}^T \mathbf{S}_T \mathbf{W})) = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m], \tag{3}$$

where $[\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m]$ is the set of n -dimensional eigenvectors (called eigenfaces when applying PCA to face images) of \mathbf{S}_T corresponding to the m largest eigenvalues $[\lambda_1, \lambda_2, \dots, \lambda_m]$. Thus, PCA maximizes the total scatter - this is the disadvantage of this method.

2.2 Fisher’s Linear Discriminant FLD, Linear Discriminant Analysis LDA

Fisher’s Linear Discriminant (FLD) (Marcialis & Roli, 2002; Martinez & Kak, 2001; Bishop, 1995; Belhumeur et al., 1997; Oravec & Pavlovičová, 2004; Duda & Hart, 1973) shapes the scatter with the aim to make it more suitable for classification. A computation of the transform matrix results in maximization of the ratio of the between-class scatter and within-class scatter.

Between-class scatter matrix \mathbf{S}_B and within-class scatter matrix \mathbf{S}_W are defined by

$$\mathbf{S}_B = \sum_{i=1}^c N_i (\mu_i - \mu)(\mu_i - \mu)^T \tag{4}$$

$$\mathbf{S}_W = \sum_{i=1}^c \sum_{\mathbf{x}_k \in X_i} (\mathbf{x}_k - \mu_i)(\mathbf{x}_k - \mu_i)^T \tag{5}$$

respectively, where N_i is the number of samples in class X_i and μ_i is the mean image of class X_i . The transform matrix \mathbf{W}_{opt} maximizes the ratio of the determinant of the between-class scatter matrix of the projected samples to the determinant of the within class scatter matrix of the projected samples:

$$\mathbf{W}_{opt} = \arg \max_{\mathbf{W}} \frac{\det(\mathbf{W}^T \mathbf{S}_B \mathbf{W})}{\det(\mathbf{W}^T \mathbf{S}_W \mathbf{W})} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m] \tag{6}$$

where $[\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m]$ is the set of generalized eigenvectors of \mathbf{S}_B and \mathbf{S}_W corresponding to the m largest generalized eigenvalues $\{\lambda_1, \lambda_2, \dots, \lambda_m\}$:

$$\mathbf{S}_B \mathbf{w}_i = \lambda_i \mathbf{S}_W \mathbf{w}_i \quad i = 1, 2, \dots, m \tag{7}$$

There are at most $c - 1$ nonzero generalized eigenvalues, i.e. the upper bound of m is $c - 1$ (Belhumeur et al., 1997; Duda & Hart, 1973).

In (Marcialis & Roli, 2002), the eigenvectors of $\mathbf{S}_W^{-1} \mathbf{S}_B$ are the columns of \mathbf{W}_{opt} and the authors show that this choice maximizes the ratio $\det(\mathbf{S}_B)/\det(\mathbf{S}_W)$.

In face recognition, the number of sample images N is typically much smaller than the number of pixels n in each image (so called small sample size problem). This is why $\mathbf{S}_W \in \mathfrak{R}^{n \times n}$ can be singular. The rank of \mathbf{S}_W is at most $N - c$. In (Belhumeur et al., 1997),

authors solve the problem of singular \mathbf{S}_W by proposal of alternative criterion to that of (6). At first, sample images are projected into lower dimensional space using PCA. This results in nonsingular \mathbf{S}_W . PCA reduces the dimension of the feature space to $N-c$, and then standard FLD (6) is applied to reduce the dimension to $c-1$. This method is called Fisherfaces. Then \mathbf{W}_{opt} can be computed as follows:

$$\mathbf{W}_{opt} = \mathbf{W}_{FLD}^T \mathbf{W}_{PCA}^T \quad (8)$$

where

$$\mathbf{W}_{PCA} = \arg \max_{\mathbf{W}} (\det(\mathbf{W}^T \mathbf{S}_T \mathbf{W})) \quad (9)$$

$$\mathbf{W}_{FLD} = \arg \max_{\mathbf{W}} \frac{\det(\mathbf{W}^T \mathbf{W}_{PCA}^T \mathbf{S}_B \mathbf{W}_{PCA} \mathbf{W})}{\det(\mathbf{W}^T \mathbf{W}_{PCA}^T \mathbf{S}_W \mathbf{W}_{PCA} \mathbf{W})} \quad (10)$$

Optimization for \mathbf{W}_{PCA} is performed over $N \times (N-c)$ matrices and optimization for \mathbf{W}_{FLD} is performed over $(N-c) \times m$ matrices. The smallest $c-1$ principal components are discarded in PCA computation.

It is often said that algorithms based of LDA outperform those based on PCA. LDA is insensitive to significant variation in lighting direction (Marcialis & Roli, 2002; Belhumeur et al., 1997), and facial expression (Belhumeur et al., 1997). However in (Martinez & Kak, 2001), authors show that when the training data set is small, PCA achieves better results compared to LDA and that PCA is less sensitive to different training data sets.

2.3 Support Vector Machines SVM

Support vector machines SVM belong to kernel methods (Muller et al., 2001; Hofmann et al., 2008) and play a major role in present machine learning algorithms.

Kernel algorithms map data $[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N] \in \mathcal{R}^p$ from an original space x into a higher dimensional feature space F using a nonlinear mapping Φ (Muller et al., 2001)

$$\phi : \mathcal{R}^p \rightarrow F, \mathbf{x} \rightarrow \phi(\mathbf{x}) \quad (11)$$

An original learning algorithm from original space is used in the feature space. High-dimensional space increases complexity of a problem; fortunately, it can be solved. Computation of a scalar product between two feature space vectors can be done using kernel function k

$$\phi(\mathbf{x}) \cdot \phi(\mathbf{y}) = k(\mathbf{x}, \mathbf{y}) \quad (12)$$

Thus, using kernel functions, the feature space does not need to be computed explicitly, only inner products in the kernel feature space are taken into account. Gaussian radial basis function, polynomial, sigmoidal, and inverse multiquadrics function are used in a role of kernel functions. Every linear algorithm that uses scalar products only can implicitly be executed in high-dimensional feature space by using kernels. Nonlinear versions of linear algorithms can be constructed in this way (Muller et al., 2001).

The basic principle of data separation by SVM is demonstrated on a simplified example in Fig. 1.

SVM accomplishes the task of finding the optimal separating hyperplane by maximizing the margin between the hyperplane and the support vectors. Dashed lines in Fig. 1 containing support vectors are parallel with the separating hyperplane and they run through the samples that are nearest to the separating hyperplane.

The separating hyperplane is defined as

$$\mathbf{w}^T \mathbf{x} + b = 0 \quad (13)$$

where \mathbf{w} is vector of weight coefficients and b is bias. The task of finding optimal separating hyperplane is accomplished by minimizing

$$\mathbf{w}^T \mathbf{w} + C \sum_i \xi_i \quad (14)$$

according to

$$y_i (\mathbf{w}^T \mathbf{x}_i + b) \geq 1 - \xi_i \quad (15)$$

where ξ_i is a slack variable that defines tolerance band around support vector and thus creates so called soft margin. The C variable controls the influence of this tolerance band.

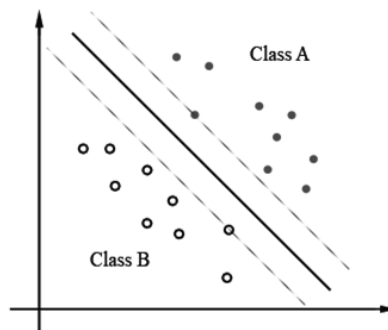


Fig. 1. Separation of data using SVM

Large amount of available papers, e.g. (Wang, et al., 2008; Hotta, 2008; Wang et al., 2004; Yang, 2002; Yang et al., 2005) indicates intensive use of SVMs and other kernel methods (kernel principal component analysis, kernel linear discriminant analysis, kernel radial basis function networks) also in face recognition area.

2.4 Metrics

Mahalanobis (also called Mahalanobis) Cosine (MahCosine) (Beveridge et al., 2003) is defined as the cosine of the angle between the image vectors that were projected into the

PCA feature space and were further normalized by the variance estimates. Let vectors \mathbf{w}_i and \mathbf{w}_j be image vectors in the unscaled PCA space (eigenvectors) and vectors \mathbf{s} and \mathbf{t} their projections in the Mahalinobis space. Using the fact that variance σ_i^2 of the PCA projections of input vectors to vector \mathbf{w}_i equals to eigenvalue λ_i ($\lambda_i = \sigma_i^2$, where σ_i is the standard deviation), the relationships between the vectors are then defined as:

$$s_i = \frac{w_{ii}}{\sigma_i}, \quad t_i = \frac{w_{ji}}{\sigma_i} \quad (16)$$

The Mahalinobis Cosine is

$$S_{MahCosine}(\mathbf{w}_i, \mathbf{w}_j) = \cos(\theta_{st}) = \frac{|\mathbf{s}||\mathbf{t}|\cos(\theta_{st})}{|\mathbf{s}||\mathbf{t}|} = \frac{\mathbf{s} \cdot \mathbf{t}}{|\mathbf{s}||\mathbf{t}|} \quad (17)$$

$$D_{MahCosine}(\mathbf{w}_i, \mathbf{w}_j) = -S_{MahCosine}(\mathbf{w}_i, \mathbf{w}_j)$$

(this is the covariance between the images in Mahalinobis space).

LDAsoft (Beveridge et al., 2003) is LDA specific distance metric. It is similar to the Euclidean measure computed in Mahalinobis space with each axis weighted by generalized eigenvalue λ (also used to compute LDA basis vectors) raised to the power 0.2 (Zhao et al., 1999):

$$D_{LDAsoft}(\mathbf{w}_i, \mathbf{w}_j) = \sum_i \lambda_i^{0.2} (w_{ii} - w_{ji})^2 \quad (18)$$

3. Image database

For our tests, we used images selected from FERET image database (Phillips et al., 1998; Phillips et al., 2000). We worked with grayscale images from Gray FERET (FERET Database, 2001). FERET face images database is de facto standard database in face recognition research. It is a complex and large database which contains more than 14126 images of 1199 subjects of dimensions 256 x 384 pixels. Images differ in head position, lighting conditions, beard, glasses, hairstyle, expression and age of subjects. Fig. 2 shows some example images from FERET database.

We selected image set containing total 665 images from 82 subjects. It consists of all available subjects from whole FERET database that have more than 4 frontal images containing also corresponding eyes coordinates (i.e. we chose largest possible set fulfilling these conditions from FERET database). The used image sets are visualized in Fig. 3.

Recognition rates are significantly influenced by size of a training set. We used 3 different sets of images for training – i.e. two, three and four images per subject in the training set. Two, three or four images for training were withdrawn from FERET database according to their file name, while all remaining images from the set were used for testing purposes.

Prior to feature extraction, all images were preprocessed. Preprocessing eliminates undesirable recognition based on non-biometric data (e.g. “T-shirts recognition” or “haircut

recognition”). Preprocessing includes following basic steps of converting original FERET image to a normalized image:

- Geometric normalization – aligning image according to available coordinates of eyes.
- Masking – cropping the image using an elliptical mask and image borders. In our experiments we tried two different maskings:
 - “face” - such that only the face from forehead to chin and cheek to cheek is visible
 - “BIGface” – leaving more of face surrounding compared to “face” – more potentially useful information is kept.
- Histogram equalization – equalizes the histogram of the unmasked part of the image.



Fig. 2. Example of images from FERET database

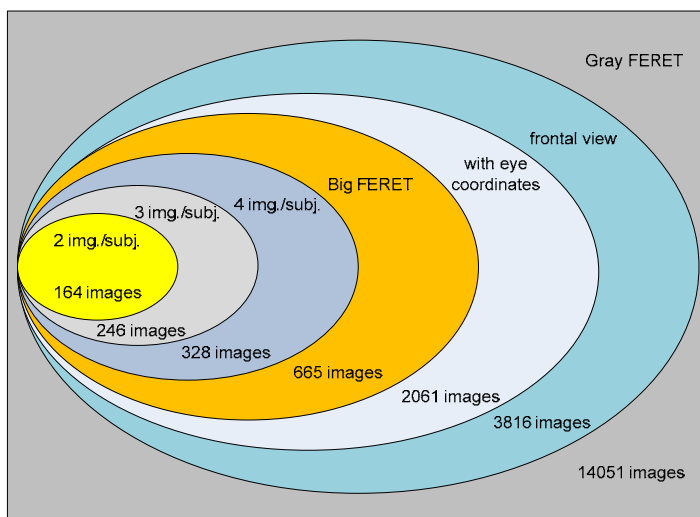


Fig. 3. Visualization of subset of images from FERET database used in our experiments

After preprocessing, the image size was 65x75 pixels. Fig. 4 shows an example of the original image, the image after “face” preprocessing and the image after “BIGface” preprocessing. All images from Fig. 2 preprocessed by “BIGface” preprocessing are shown in Fig. 5.



Fig. 4. Example of original image, image after “face” preprocessing and image after “BIGface” preprocessing



Fig. 5. Images from Fig. 2 preprocessed by “BIGface” preprocessing

4. Examined Methods for Face Recognition

We examined five different setups of face recognition experiments. They contain both single-stage and two-stage recognition systems as shown in Fig. 6:

- In single-stage face recognition (Fig. 6a), SVM is used for classification directly (i.e. there is no feature extraction performed).
- For two-stage face recognition setups including both feature extraction and classification (Fig. 6b - Fig. 6e), we used PCA with MahCosine metrics, LDA with LDASoft metrics and our proposed methods utilizing both PCA and LDA for feature extraction followed by SVM for classification. We propose also optimal parameter setups for the best performance of these methods.

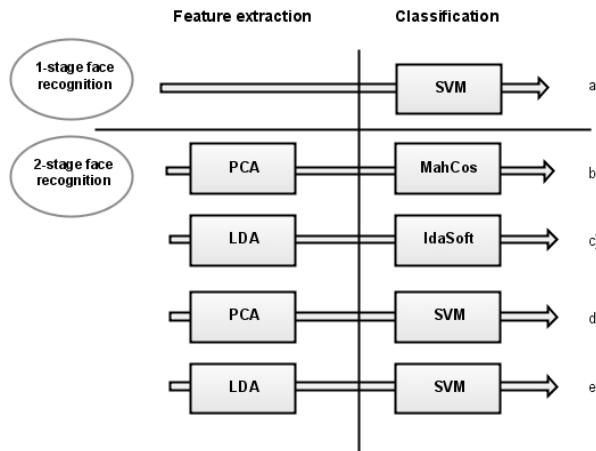


Fig. 6. Methods and classifiers used in our experiments

Last two setups (Fig. 6d) and e)) are our proposed combinations of efficient feature extraction combined with strong classifier. First three setups (Fig. 6a-c)) are the conventional methods, presented for comparison with proposed approaches.

All five setups are significantly influenced by different settings of parameters of the examined methods (i.e. PCA, LDA or SVM). This is the reason we present serious analysis and proposal of parameter settings in following chapters.

We used CSU Face Identification Evaluation System (csuFaceIdEval) (Beveridge et. al., 2003) and libsvm - A Library for Support Vector Machines (LIBSVM, web) that implement mentioned algorithms.

5. Face Recognition Experiments and Results in Ideal Conditions

5.1 Single-Stage Recognition

SVM was directly used for recognizing faces without previous feature extraction from the images (see Fig. 6a)). Input images were of size 65x75 pixels.

In our tests we used SVM with the RBF (radial basis function) kernel

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2\right), \quad \gamma > 0 \quad (19)$$

where $\mathbf{x}_i, \mathbf{x}_j$ are data points (face images) from original space.

It is important to find optimal parameters γ (gamma) and C , because different parameter setups are suitable for solving different problems. $C > 0$ is the penalty parameter of the error term used in a determination of a separating hyperplane with the maximal margin in higher dimensional space by SVM. We used methodology from (Hsu et al., 2008), i.e. parameters search where the best v -fold cross-validation rate performed on training data suggests also the best parameter setup. v -fold cross-validation divides the training set into v subsets of equal size, and sequentially one subset is tested using the classifier that was trained on the remaining $v-1$ subsets. Fig. 7 shows example of the graph we used for parameter search – the dependence of cross validation rate on the parameters C and γ . The best found parameters setups for all training sets and the results are shown in Table 1.

More images per subject in the training set result in better cross-validation rate and also better recognition rate. Difference between face recognition rate using “face” and “BIGface” preprocessing is noticeable only with 2 images per subject, where the result with “BIGface” preprocessing is approx. 5,6% worse than with “face” preprocessing.

It is important to point out that it is not possible to find “universal” values of parameters C and γ that would lead to the best recognition rates independent of used training set and preprocessing type.

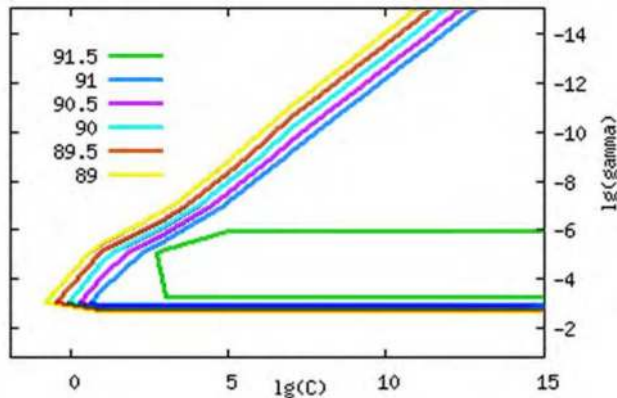


Fig. 7. Example of the output graph – dependence of cross validation rate on the parameters C and γ for training set with 3 images per subject

training set	C	γ	cross-valid.	rec. rate
face, 2img/pers.	0,03125	0,0078125	51,22%	80,04%
face, 3img/pers.	128	3,05176E-05	78,86%	93,79%
face, 4img/pers.	128	3,05176E-05	86,59%	96,74%
BIGface, 2img/pers.	0,03125	0,0078125	64,63%	74,45 %
BIGface, 3img/pers.	8	0,00012207	83,33%	93,56%
BIGface, 4img/pers.	128	3,05176E-05	89,63%	96,74%

Table 1. Recognition rate and optimal SVM parameter setups for used training sets

5.2 Two-Stage Recognition Systems

PCA and LDA algorithms are used to reduce the dimension and extract the features from face images. Using the training set, they produce a transform matrix. For face recognition purposes, we do not need the whole transform matrix and therefore we truncate first or last vectors from the transform matrix. The results of recognition are significantly influenced by parameters “Dropped from front” and “CutOff”.

- **Dropped from front (DPF)** – denotes number of eigenvectors cut from the beginning of transform matrix (first vectors - vectors belonging to the largest eigenvalues). These vectors will not be used by image projection to PCA (or LDA) feature space. Reason to truncate these vectors is based on the assumption that these vectors do not correspond to useful information such as lighting variations (Beveridge et. al., 2003). Our tests were performed for “Dropped from front” values 0, 1, 2, 3, and 4.
- **CutOff (CO)** – represents how many vectors remain in the transform matrix. Reason to truncate last basis vectors (vectors corresponding to the smallest eigenvalues) is to lower the computation requirements and to eliminate unnecessary information that correlates with noise – and as such is meaningless for recognizing faces (Beveridge et. al., 2003). Our tests were performed for CutOff parameter set to 20%, 40%, 60%, 80% and 100%.

preprocessing:	face
number of subjects	82
image resolution:	65x75px

Legend: = Above 96% of images recognized
 = Above 88% of images recognized
 = Above 80% of images recognized

		2 img./person in training set					3 img./person in training set					4 img./person in training set				
		Dropped from front					Dropped from front					Dropped from front				
		0	1	2	3	4	0	1	2	3	4	0	1	2	3	4
Cutoff	20.00%	74,5	73,9	73,7	75,4	75,0	82,8	83,1	81,9	81,9	81,1	91,1	91,1	90,2	90,5	90,8
	40.00%	80,4	80,2	80,0	78,8	79,0	82,1	81,6	81,1	80,2	80,4	90,5	90,8	90,8	90,5	90,8
	60.00%	77,6	77,4	77,6	76,0	75,8	77,1	76,1	76,1	76,1	75,2	89,9	89,6	89,9	89,6	89,9
	80.00%	75,0	74,3	73,9	73,7	74,3	72,3	72,6	71,8	72,6	71,8	84,6	84,0	84,0	84,0	84,0
	100.00%	67,9	67,7	67,5	67,5	67,5	68,3	68,0	67,8	67,1	66,8	78,6	78,6	78,6	78,0	77,2
		max: 80,4 min: 67,5					max: 83,1 min: 66,8					max: 91,1 min: 77,2				
PCA-Mahcosine																
		Dropped from front					Dropped from front					Dropped from front				
		0	1	2	3	4	0	1	2	3	4	0	1	2	3	4
Cutoff	20.00%	82,4	82,0	82,4	82,0	80,4	95,5	94,5	93,8	93,6	92,8	96,7	95,8	95,5	94,4	95,0
	40.00%	82,0	82,4	81,6	81,6	81,6	95,5	95,0	95,2	95,2	95,0	98,2	97,3	97,6	97,3	97,9
	60.00%	87,2	86,6	86,2	84,6	85,0	95,7	95,2	95,5	95,7	95,9	98,5	98,5	97,3	97,6	98,5
	80.00%	91,6	90,8	90,6	91,4	92,2	97,6	97,4	96,9	96,7	96,7	97,6	96,7	97,0	97,3	97,6
	100.00%	93,2	92,6	93,0	92,8	92,4	97,4	97,6	97,6	97,1	97,6	97,3	97,0	97,3	97,9	97,9
		max: 93,2 min: 80,4					max: 97,6 min: 92,8					max: 98,5 min: 94,4				
PCA-SVM																
		Dropped from front					Dropped from front					Dropped from front				
		0	1	2	3	4	0	1	2	3	4	0	1	2	3	4
Cutoff	20.00%	84,8	83,8	84,4	83,4	83,6	89,5	88,1	88,1	86,2	85,9	95,8	94,4	93,8	93,8	93,5
	40.00%	84,8	83,8	84,4	83,4	83,6	90,2	89,3	89,0	88,5	87,6	95,8	95,5	95,3	95,0	94,4
	60.00%	66,1	63,7	56,9	59,3	56,5	93,6	93,6	93,3	92,6	92,1	96,7	96,7	96,4	96,4	96,7
	80.00%	47,9	50,3	51,3	51,5	51,1	82,3	80,9	80,0	80,9	77,3	76,3	64,7	71,5	62,9	57,3
	100.00%	34,7	35,9	32,9	33,5	41,7	57,3	62,1	64,0	62,3	64,7	68,5	72,1	70,9	79,5	76,6
		max: 84,8 min: 32,9					max: 93,6 min: 57,3					max: 96,7 min: 57,3				
LDA-LdaSoft																
		Dropped from front					Dropped from front					Dropped from front				
		0	1	2	3	4	0	1	2	3	4	0	1	2	3	4
Cutoff	20.00%	85,0	86,0	84,2	84,6	82,4	95,5	95,9	95,9	95,0	95,5	97,3	96,1	96,7	96,7	95,8
	40.00%	85,0	86,0	84,2	84,6	82,4	95,5	95,0	94,7	94,7	93,6	96,1	95,5	96,7	96,1	95,8
	60.00%	87,0	87,4	85,0	86,8	86,4	96,9	96,4	96,9	96,4	97,1	97,6	97,6	97,9	97,6	98,2
	80.00%	90,2	89,8	89,6	91,2	89,2	97,1	96,7	96,4	96,7	96,4	97,6	97,6	97,6	97,0	97,3
	100.00%	86,8	88,4	89,2	85,6	84,0	96,4	96,7	95,9	95,0	95,5	97,9	95,8	95,0	95,0	96,1
		max: 91,2 min: 82,4					max: 97,1 min: 93,6					max: 98,2 min: 95,0				
LDA-SVM																

Table 2. Results of experiments for methods PCA+MahCosine, PCA+SVM, LDA+LdaSoft, LDA+SVM with “face” preprocessing (total 300 tests)

Methods utilizing PCA or LDA (Fig. 6b - Fig. 6e) were tested using three training sets with 2, 3 and 4 images per subject. For each method, we tested 25 different parameters DPF and CO setups on three different training sets, what gives total 75 tests per each method and per each type of preprocessing (600 tests in total). Results of these tests are shown in Table 2 and Table 3. The maximal recognition rates are summarized in Fig. 8 and Fig. 9.

preprocessing:	BIGface
number of subjects	82
image resolution:	65x75px

Legend:

- = Above 96% of images recognized
- = Above 88% of images recognized
- = Above 80% of images recognized
- = selected for noise tests

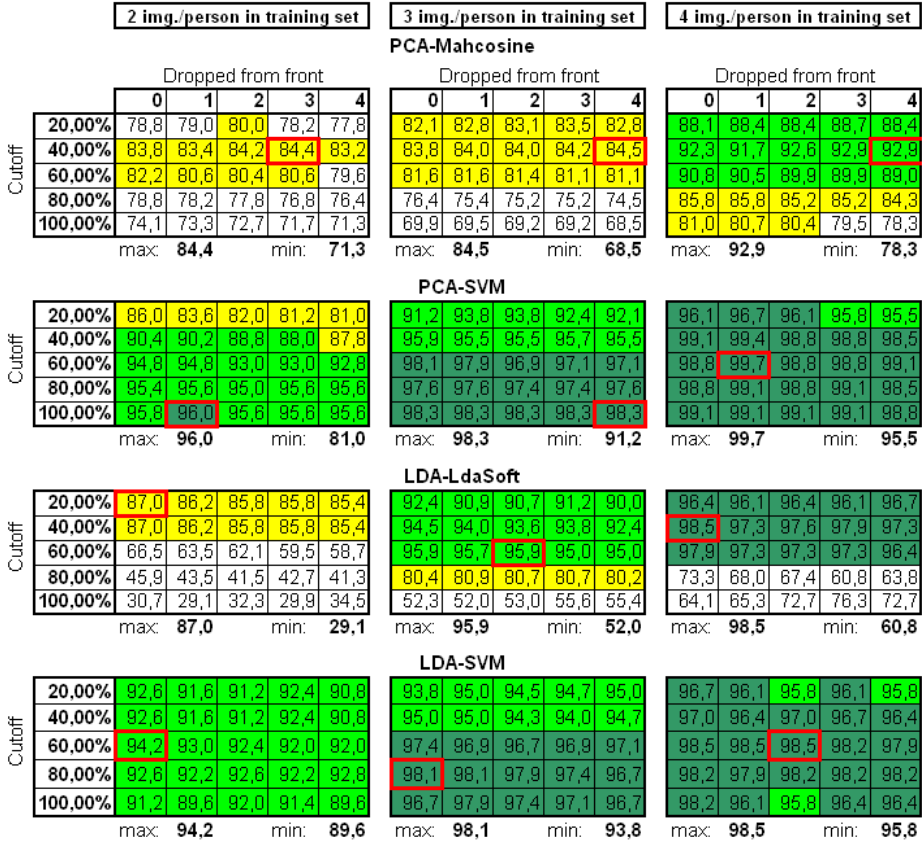


Table 3. Results of experiments for methods PCA+MahCosine, PCA+SVM, LDA+LDAsoft, LDA+SVM with "BIGface" preprocessing (total 300 tests)

5.3 Evaluation of Simulation Results

Based on presented experiments, we can formulate several conclusions:

1. More images in the training stage cause better performance of all methods.
2. LDA+LDAsoft performs better than PCA+MahCosine, but PCA+SVM is slightly better than LDA+SVM.
3. The best performing setups of parameters CO and DPF differ using different preprocessing and number of images per subject in training set. Generally PCA+MahCosine and LDA+LDAsoft perform better for truncating 0-4 first vectors and leaving 20%-60% of the vectors in transform matrix.

4. The recognition rate is most significantly affected by setting of the CO parameter – for PCA+MahCosine and LDA+LDASoft it is better to truncate vectors from the end of the transform matrix leaving only 20% - 60% of the vectors. Methods PCA+SVM and LDA+SVM perform better when leaving more (60% - 100%) vectors of the transform matrix.
5. Results of LDA+LDASoft are more influenced by setting the CO parameter compared to PCA+MahCosine – especially with only 2 images per subject in the training set, where the worst recognition rate is around 30% (see Table 2 and Table 3).
6. Using SVM for classification (methods PCA+SVM and LDA+SVM) makes the recognition rates more stable and less influenced by setting the CO and DPF parameters (see Table 2 and Table 3) and these methods perform better compared to simple PCA+MahCosine and LDA+LDASoft – see Fig. 8 and Fig. 9.

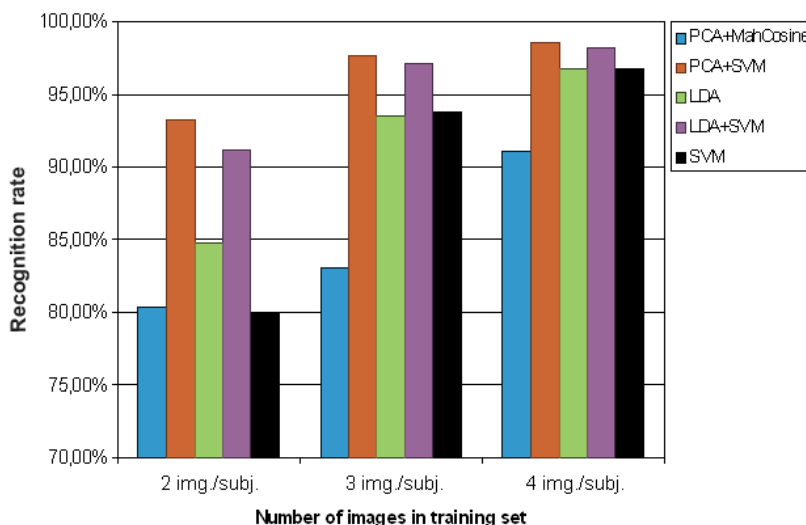


Fig. 8. Graph of maximum recognition rates for methods PCA+MahCosine, PCA+SVM, LDA+LDASoft, LDA+SVM, SVM (left to right) with “face” preprocessing

6. Face Recognition Experiments and Results in Noisy Conditions

In this part of the chapter, we concentrate on the influence of input image quality to face recognition accuracy. Noise and distortions in face images can seriously affect the performance of face recognition systems. Analog or digital capturing the image, image transmission, image copying or scanning can suffer from noise. This is why we study behaviour of discussed methods in the presence of noise.

We include Gaussian noise, salt & pepper noise and speckle noise. Huge effort in removing these types of noise from static or dynamic images in the area of face recognition is documented in the literature, e.g. (Uglov et al., 2008; Reda, & Aoued, 2004; Wheeler et al., 2007). We use these types of noise with various intensities (various parameters).

6.1 Types of Noises

Each image capturing generates digital or analog noise of diverse intensity. The noise is also generated while transmitting and copying analog images. Noise generation is a natural property for image scanning systems. Diverse types of noises exist. Herein we use three different types: Gaussian (Truax, 1999), salt & pepper (Chan et al., 2005), and speckle (Anderson & Trahey, 2006) noises.

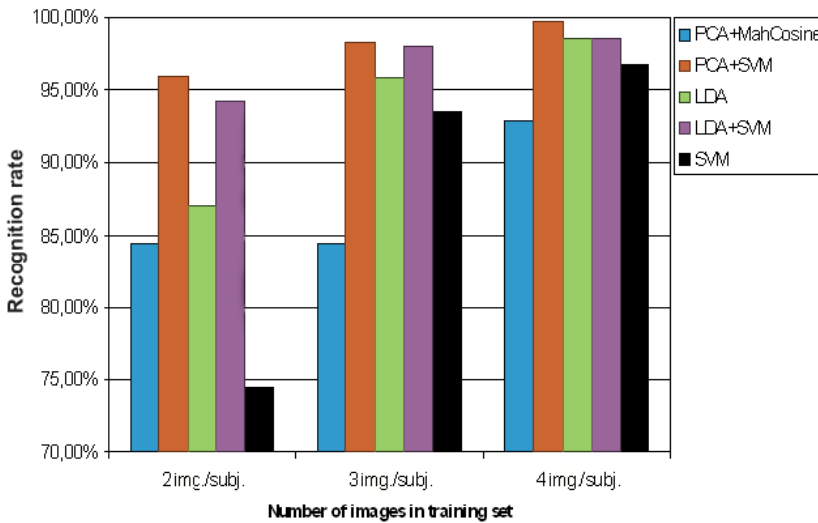


Fig. 9. Graph of maximum recognition rates for methods PCA+MahCosine, PCA+SVM, LDA+LDASoft, LDA+SVM, SVM (left to right) with “BIGface” preprocessing

Gaussian Noise

Gaussian noise is the most common noise occurring in everyday life. The Gaussian noise can be detected in free radio waves or in television receivers. Gaussian noise is produced in analog images that are stored for a long time.

We studied face recognition with different Gaussian noise intensity. Gaussian noise was generated with Gaussian normal distribution function which can be written as:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (20)$$

where μ is the mean value of the required distribution and σ^2 is a variance (Truax, 1999; Chiodo, 2006).

Noise parameters settings for our simulations were determined empirically. The mean of Gaussian distribution was set to 0 and we changed the variance. Examples of images corrupted by Gaussian noise can be seen in Fig. 10. The label g0.01 means that the Gaussian

noise of variance 0.01 was applied on the image. The same notation is used also in presented graphs.

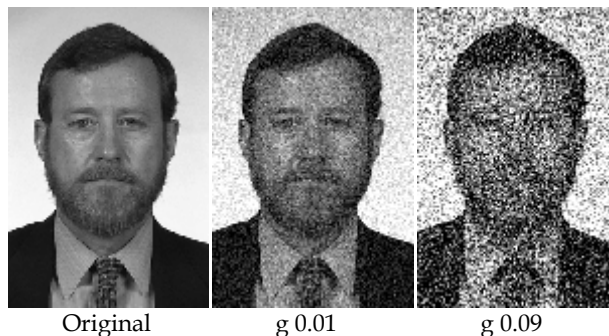


Fig. 10. Examples of images corrupted by Gaussian noise

Salt & Pepper Noise

Salt & pepper noise is perceived as a random occurrence of black and white pixels in a digital image. It can be caused by incorrect data transmission or by a damage of already received data. In CCD and CMOS sensors or LCD displays, the salt & pepper noise can be caused by permanently turned-on or turned-off pixels. Remaining pixels are unchanged. Usually, the intensity (frequency of the occurrence) of this noise is quantified as a percentage of incorrect pixels (Fisher et al., 2003). The median filtering (as a specific case of order-statistic filtering) is used as an effective method for elimination of salt & pepper noise from digital images (Chan et al., 2005).

Noise parameter settings for our simulations vary from 4% of noise intensity (0.04) up to the 30% of damaged pixels. The label sp0.04 means, that the salt & pepper noise of intensity 4% was applied on the image. Examples of images corrupted by salt & pepper noise are shown in Fig. 11.

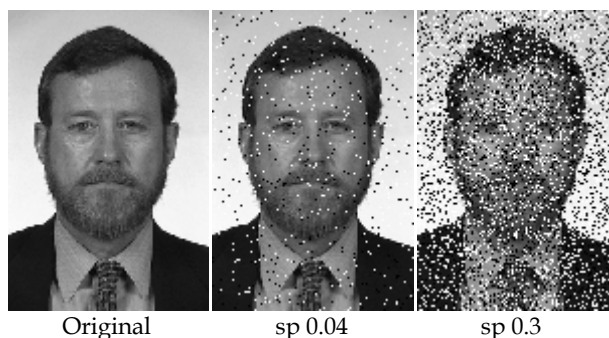


Fig. 11. Examples of images corrupted by 4% and 30% salt & pepper noise

Speckle Noise

This granular noise occurs in ultrasound, radar and X-ray images and images obtained from the magnetic resonance (Chaillan et al., 2007). The multiplicative signal dependent noise is generated by constructive and destructive interference of detected signals. The wave

interference is a reason of multiplicative noise occurrence in the scanned image. The speckle noise is image dependent. Therefore it is very hard (if possible) to find a mathematical model that describes the removal of this noise, especially if we expect the randomness of the input data (Fisher et al., 2003).

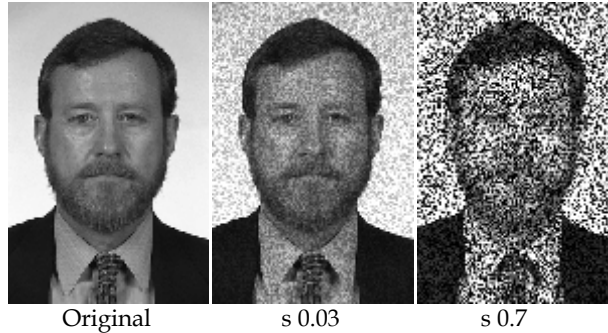


Fig. 12. Examples of images corrupted by speckle noise

The values which determined intensity of noise in our tests were set empirically. The noise was applied according to the following equation

$$S = I + n * I \quad (21)$$

where I is the original human face image and n is the uniform distribution of the noise with zero mean value and variance σ^2 . For our simulations, variance varied from 0.03 to 0.7. The label s0.03 means that the speckle noise of variance 0.03 was applied on the image. Presence of speckle noise in the face image is illustrated in Fig. 12.

For simulation of methods in presence of noise, we use the best parameter settings we obtained running 600 tests in Section 5, i.e. when the methods worked in ideal conditions.

In order to mimic real-world conditions, we use images not distorted by noise for training purposes whilst noisy images are used for testing. Such scenario simulates real-world face recognition conditions.

We concentrate on “BIGface” preprocessed images only, since this preprocessing gives better results compared to “face” preprocessing (this can be seen when comparing Tables 2 and 3). Parameters for settings of the algorithms (CO and DPF) were empirically obtained from Table 3. We selected and used only those parameters for which the recognition experiments were most successful (they are marked by red in Table 3). This was necessary in order to reduce the number of experiments. Using all possible settings from simulations in ideal conditions and combining them with three types of noises with all selected parameters would lead to total 13500 results. Selecting best parameters only lead us to total 540 results. Obtained results are shown in Fig. 13 – 21 along with brief comments.

6.2 Simulation Results for Face Images Corrupted by Gaussian Noise

Simulation results for face images corrupted by Gaussian noise are summarized in Fig. 13 – 15. PCA-MahCosine method is most influenced by increasing the intensity of Gaussian noise. Results for training sets with 2 and 3 img./subj. look alike – recognition rates decrease with higher noise. The effect of the noise for training set containing 4 img./subj. is not so noticeable. Worst results are achieved by PCA-MahCosine method. For training set with 4 img./subj., the results of other 3 methods are almost equal and the recognition rates are surprisingly high even for higher noise intensities and they do not decrease. For 3 img./subj., the best results come from LDA-SVM method, followed by LDA-LDASoft (from intensity of noise >0.01). For training set containing 2 img./subj. only, both SVM methods result in best recognition rates and LDA-SVM is slightly better than PCA-SVM. It is also interesting to notice that there are some cases, when consecutive increase of noise levels resulted in better recognition rates.

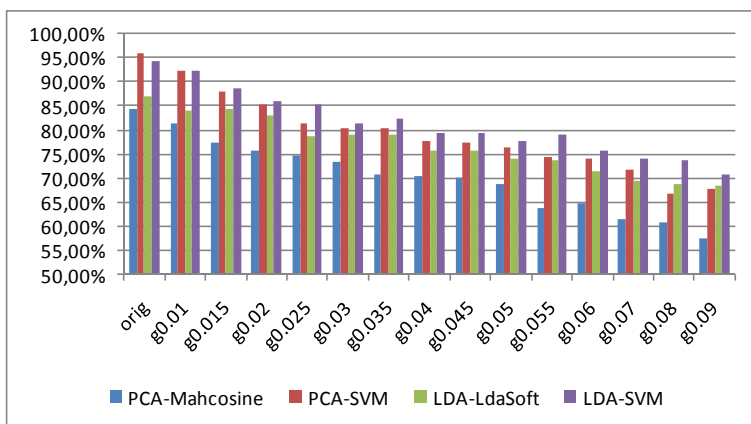


Fig. 13. Recognition rates of examined methods, Gaussian noise, training set 2 img./subj.

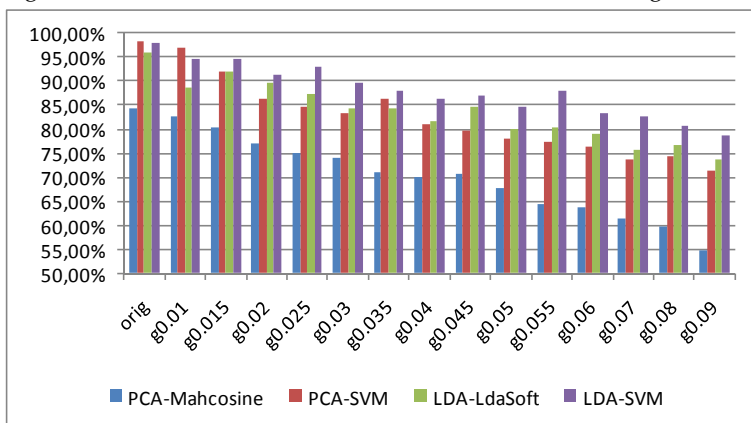


Fig. 14. Recognition rates of examined methods, Gaussian noise, training set 3 img./subj.

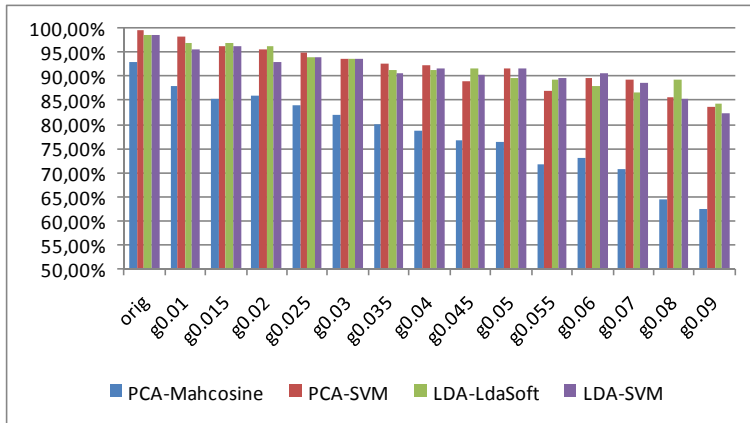


Fig. 15. Recognition rates of examined methods, Gaussian noise, training set 4 img./subj.

6.3 Simulation Results for Face Images Corrupted by Salt & Pepper Noise

Fig. 16 - 19 show results for face images corrupted by salt & pepper noise. Increasing the noise level does not have significant effect till intensity 0.2. Decrease of the recognition rate while increasing the noise intensity is most noticeable for results with 2 img./subj. in the training set. PCA-MahCosine is again the worst method. Best recognition rates are achieved by the methods that use SVM and they both achieved almost equal results. For 3 img./subj., LDA-SVM was slightly better than PCA-SVM. One can again notice, that in some cases consecutive increase of noise levels resulted in better recognition rates.

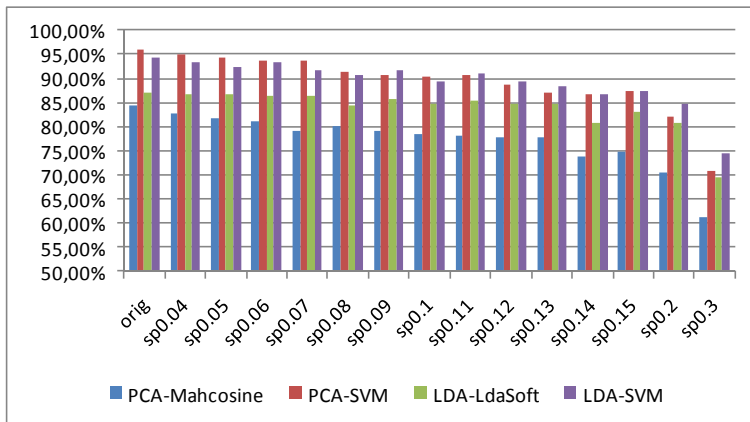


Fig. 16. Recognition rates of examined methods, salt & pepper noise, training set 2 img./subj.

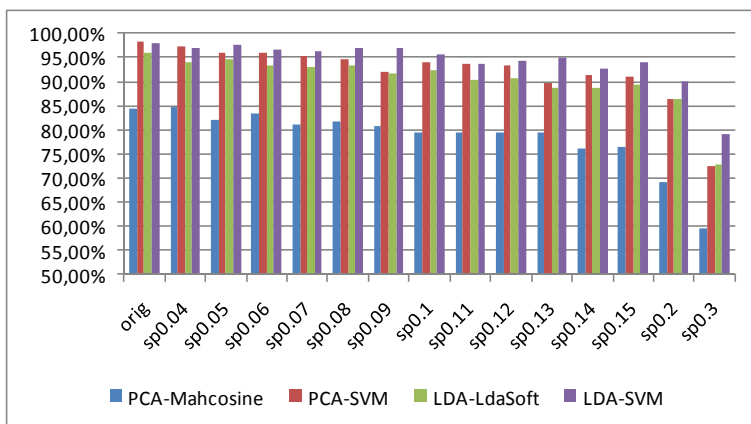


Fig. 17. Recognition rates of examined methods, salt & pepper noise, training set 3 img./subj.

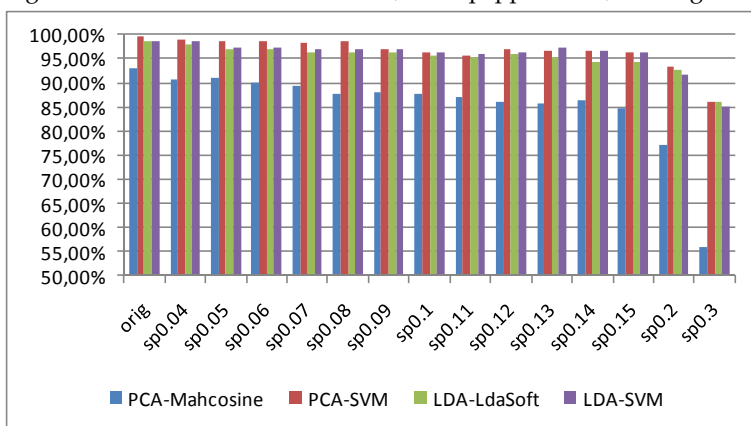


Fig. 18. Recognition rates of examined methods, salt & pepper noise, training set 4 img./subj.

6.4 Simulation Results for Face Images Corrupted by Speckle Noise

Fig. 19 – 21 contains simulation results for face images corrupted by speckle noise. PCA-MahCosine method achieves worst results. Best results can be achieved by LDA-SVM; this is more noticeable for higher noise intensities. For 4 img./subj., the PCA+SVM, LDA+LDAsoft and LDA+SVM methods have almost equal recognition rates. For 3img./subj., the LDA+LDAsoft method is better than PCA+SVM, for 2 img./subj., the PCA+SVM is better than LDA+LDAsoft. For speckle noise, there are not cases when higher noise levels result in better recognition rates. There was an exception for speckle noise of intensity 0.03 for training set 3 img./subj., because recognition by PCA-MahCosine method gives better rate for corrupted images (84.73%) than recognition using the original images (84.5%).

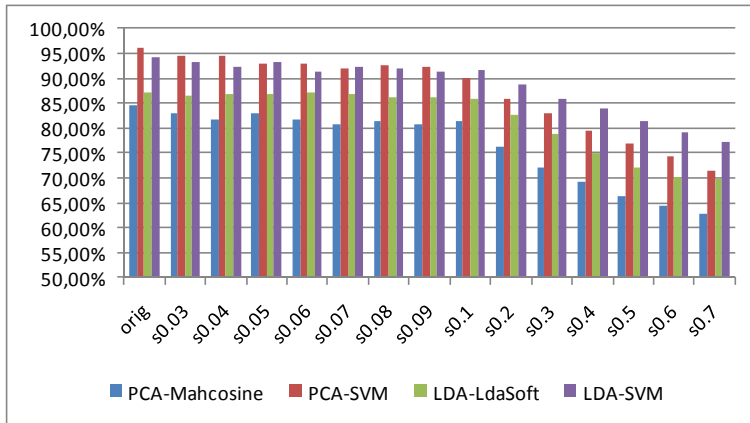


Fig. 19. Recognition rates of examined methods, speckle noise, training set 2 img./subj.

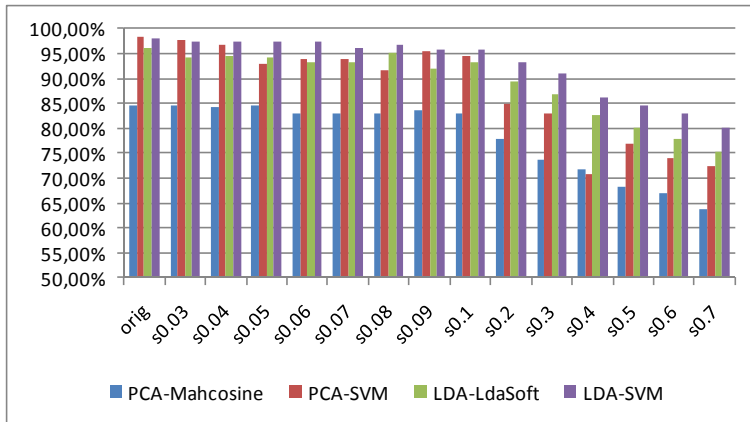


Fig. 20. Recognition rates of examined methods, speckle noise, training set 3 img./subj.

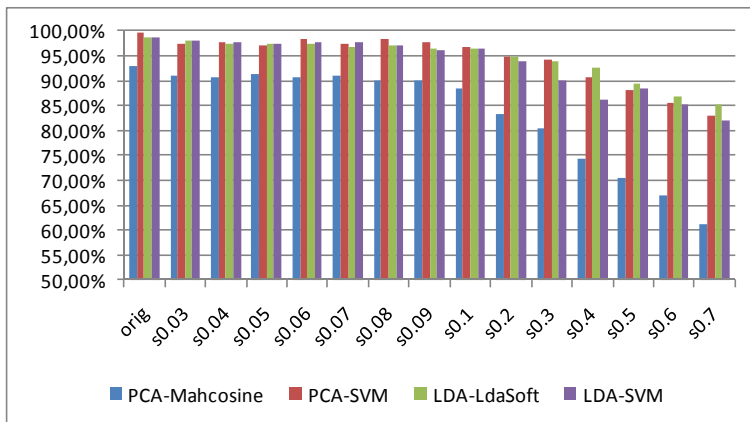


Fig. 21. Recognition rates of examined methods, speckle noise, training set 4 img./subj.

6.5 Equivalence of Different Types of Noise from the Recognition Point of View

After presenting recognition results for different types of noise an interesting question arises: What is the relationship among different noise types? The concrete values of noise parameters do not give the answer – a comparison cannot be based on non-related parameters.

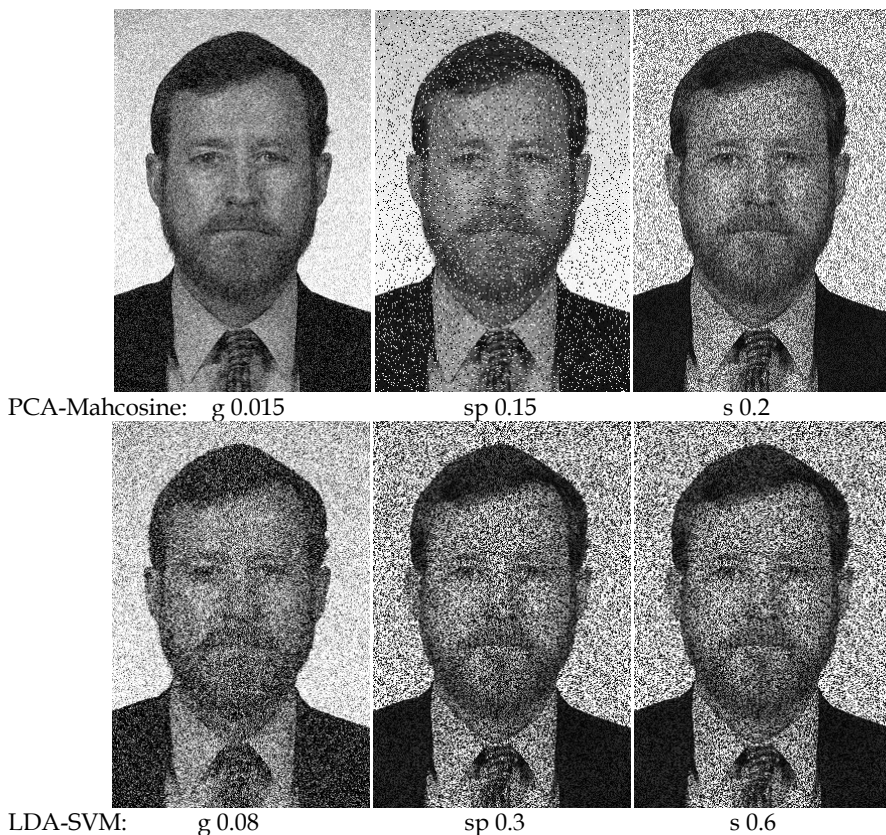


Fig. 22. Example of the subject, for who all the studied methods (here shown PCA-MahCosine and LDA-SVM) result in recognition accuracy about 85 % (see Table 4 for exact noise type and intensity)

One possible solution can be based exactly on results of machine face recognition. This approach is illustrated in Fig. 22 and in corresponding Table 4. Fig. 22 shows images of the subject corrupted by different types of noises. The noise parameters are chosen in such manner that all studied methods (PCA-MahCosine, PCA-SVM, LDA-LDASoft, LDA-SVM) result in recognition accuracy near 85 %. Table 4 specifies each noise type and its corresponding parameter. PCA-MahCosine and LDA-SVM methods are included in Fig. 22, since PCA-SVM and LDA-LDASoft methods are visually similar to LDA-SVM. Fig. 22 thus shows equivalence of different types of noise from the face recognition point of view of

PCA-MahCosine and LDA-SVM methods. But this is equivalency of noise types from machine point of view. It should be even more interesting to compare recognition ability of machine learning methods and humans.

method	Gaussian noise	Recognition rate in %	Salt&pepper noise	Recognition rate in %	Speckle noise	Recognition rate in %
PCA-Mahcosine*	g0.015	85,16%	sp0.15	84,87%	s0.2	83,38%
PCA-SVM	g0.08	85,76%	sp0.3	86,05%	s0.6	85,46%
LDA-LdaSoft	g0.09	84,27%	sp0.3	86,05%	s0.7	85,16%
LDA-SVM*	g0.08	85,16%	sp0.3	85,16%	s0.6	85,16%

Table 4. Types and intensity of noise resulting in recognition rate about 85 % (for training set 4img./subj.).

* included in Fig. 22

7. Conclusion

We examined different scenarios of face recognition experiments. They contain both single-stage and two-stage recognition systems. Single-stage face recognition uses SVM for classification directly. Two-stage recognition systems include PCA with MahCosine metric, LDA with LDASoft metric and also methods utilizing both PCA and LDA for feature extraction followed by SVM for classification. All methods are significantly influenced by different settings of parameters that are related to the algorithm used (i.e. PCA, LDA or SVM). This is the reason we presented serious analysis and proposal of parameter settings for the best performance of discussed methods.

For methods working in ideal conditions, the conclusions are as follows: When comparing non-SVM based methods, higher maximum recognition rate is generally achieved by method LDA+LDASoft compared to PCA+MahCosine; on the other hand LDA+LDASoft is more sensitive to method settings. Using SVM in classification stage (PCA+SVM and LDA+SVM) produced better maximum recognition rate than standard PCA and LDA methods.

Experiments with single-stage SVM show that this method is very efficient for face recognition even without previous feature extraction. With 4 images per subject in training set, we reached 96.7% recognition rate.

The experiments were made with complex image set selected from FERET database containing 665 images. Such number of face images entitles us to speak about general behavior of presented methods. Altogether more than 600 tests were made and maximum recognition rates near 100% were achieved.

It is important to mention that the experiments were made with "closed" image set, so we did not have to deal with issues like detecting people who are not in the training set. On the other hand, we worked with real-world face images; our database contains images of the same subjects that often differ in face expressions (smiling, bored, ...), with different hairstyles, with or without beard, or wearing glasses and that were taken in different session after longer time period (i.e. we did not work with identity card-like images).

We also presented recognition results for noisy images and graphically compared them to results for non-distorted images. In this way, the insight on face recognition system robustness is obtained.

Independently on noise type or its parameter, the PCA-MahCosine method gives the lowest success in face recognition compared to all tested methods. Using other methods, the results were significantly better. Methods that use SVM classifier achieve globally better results for each training set. On the other hand, SVM-based methods need a lot of time to search for optimal parameters, while PCA-MahCosine method is the fastest.

By our work, we continue in our effort to offer complex view to biometric face recognition. In (Oravec et al., 2008) besides detection of faces and facial features, we presented feature extraction methods from face images (linear and nonlinear methods, second-order and higher-order methods, neural networks and kernel methods) and relevant types of classifiers. Face recognition in ideal conditions using FERET database is contained partly in (Oravec et al., 2009) and in this chapter.

Our work on presented methods now further continues in evaluating their sensitivity and behavior in non-ideal conditions. First our contribution to this area which includes presence of noise is covered in this chapter. Our future work will comprise partially occluded faces and also faces extracted from static images and/or video streams transmitted with errors or loss of data, where some parts of face image are missing (block or blocks of pixels) or an error-concealment mechanism is applied prior to recognition (Pavlovičová et al., 2006; Polec et al., 2009; Marchevský & Mochňáč, 2008).

Our future work will also be focused on a psychological experiment trying to find relationship for mentioned types of distortions from the point of view of recognition ability of humans and machines (as an extension of the aspect of noise for machine recognition that is outlined in section 6.5).

8. Acknowledgements

Research described in this paper was done within the grants 1/0718/09 and 1/0883/08 of the Slovak Grant Agency VEGA.

Portions of the research in this paper use the FERET database of facial images collected under the FERET program, sponsored by the DOD Counterdrug Technology Development Program Office.

9. References

- Anderson, M. E. & Trahey, G. E. (2006). A beginner's guide to speckle. *A seminar on k-space applied to medical ultrasound* [online], Available on: <http://dukemil.bme.duke.edu/Ultrasound/k-space/node5.html>, Dept. of Biomedical Engineering, Duke University
- Belhumeur, P. N.; Hespanha, J. P. & Kriegman, D. J. (1997). Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection, *IEEE Trans. Pattern Analysis and Machine Intelligence*, pp. 711-720, Vol. 19, No. 7, July 1997

- Beveridge, R.; Bolme, D.; Teixeira, M. & Draper, B. (2003). The CSU Face Identification Evaluation System User's Guide, Version 5.0, Technical Report., Colorado State University, May 2003, <http://www.cs.colostate.edu/evalfacerec/algorithms/version5/faceIdUsersGuide.pdf>
- Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*, Oxford University Press, Inc., ISBN 0 19 853864 2, New York
- Chaillan, F.; Frascini, C. & Courmontagne, P. (2007). Speckle Noise reduction in SAS imagery, *Signal Processing*, Vol. 87, No. 4., page numbers 762-781, April 2007, ISSN 0165-1684
- Chan, R. H., Ho, C. W., Nikolova, M. (2005). Salt-and-pepper noise removal by median-type noise detector and edge-preserving regularization, *IEEE Trans. Image Processing*, Vol. 14, page numbers 1479-1485
- Chiodo, K. (2006). Normal Distribution, *NIST/SEMATECH e-Handbook of Statistical Methods*, <http://www.itl.nist.gov/div898/handbook/eda/section3/eda3661.htm>
- Delac, K.; Grgic, M. & Bartlett, M. S. (2008). Recent Advances in Face Recognition, *IN-TECH*, 2008, <http://intechweb.org/book.php?id=101>
- Duda, R.O. & Hart, P.E. (1973). *Pattern Classification and Scene Analysis*, John Wiley & Sons, Inc., ISBN-10: 0471223611, New York
- EU - Passport Specification (2006), Biometrics Deployment of EU-Passports, Working document (EN) - 28/06/2006, http://ec.europa.eu/justice_home/doc_centre/freetravel/documents/doc/c_2006_2909_prov_en.pdf
- FERET Database (2001). <http://www.itl.nist.gov/iad/humanid/feret/>, NIST
- Fisher, R., Perkins, S., Walker, A. & Wolfart, E. (2003). Noise Generation, *Hypermedia Image Processing Reference*, <http://homepages.inf.ed.ac.uk/rbf/HIPR2/noise.htm>
- Haykin, S. (1994). *Neural Networks - A Comprehensive Foundation*, Macmillan College Publishing Company, ISBN-10: 0139083855, New York
- Hofmann, T.; Schölkopf, B. & Smola, A. J. (2008). Kernel Methods in Machine Learning, *Annals of Statistics*, Vol. 36, No. 3, 2008, page numbers 1171-1220
- Hotta, K. (2008). Robust Face Recognition under Partial Occlusion Based on Support Vector Machine with Local Gaussian Summation Kernel, *Image and Vision Computing*, Vol. 26, Issue 11, November 2008, page numbers 1490-1498
- Hsu, C. W.; Chang, C.C. & Lin, C.J. (2008). A Practical Guide to Support Vector Classification, *Technical report*, Department of Computer Science, National Taiwan University. July, 2003, last updated October 2, 2008, <http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>
- Jain, A. K.; Flynn, P. & Ross, A. A. (2008), (Eds). *Handbook of Biometrics*, Springer, ISBN: 978-0-387-71040-2
- Jain, A. K.; Ross, A. & Prabhakar, S. (2004). An Introduction to Biometric Recognition. *IEEE Trans. Circuits and Systems for Video Technology*, Vol. 14, No. 1, January 2004, page numbers 4-20
- Li, S. Z. & Jain, A. K. (2005) (Eds.). *Handbook of Face Recognition*, Springer, ISBN# 0-387-40595-x New York
- LIBSVM - A Library for Support Vector Machine, <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>
- Marchevský, S. & Mochnáč, J. (2008). Hybrid Concealment Mechanism, *Acta Electrotechnica et Informatica*, Vol. 8, No. 1, 2008, page numbers 11-15, ISSN 1335-8243

- Marcialis, G. L. & Roli, F. (2002). Fusion of LDA and PCA for Face Recognition, *Proc. of the Workshop on Machine Vision and Perception, 8th Meeting of the Italian Association of Artificial Intelligence (AI*IA)*, Siena, Italy, September 2002
- Martinez, A. M. & Kak, A. C. (2001). PCA versus LDA. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 23, No. 2, February 2001, page numbers 228-233
- Mazanec, J.; Melišek, M.; Oravec, M. & Pavlovičová, J. (2009). Face Recognition Using Support Vector Machines, PCA and LDA with the FERET Database, *Acta Polytechnica Hungarica*, submitted for publication in 2009
- Muller, K.; Mika, S.; Ratsch, G.; Tsuda, K. & Scholkopf, B. (2001). An Introduction to Kernel-Based Learning Algorithms. *IEEE Transactions on Neural Networks*, Vol. 12, No. 2, March 2001, page numbers 181 – 201
- Oravec, M. & Pavlovičová, J. (2004). Linear Discriminant Analysis for Face Images, *Proc. of the 27th Int. Conf. Telecommunications and Signal Processing TSP-2004*, pp. 146-149, Brno, Czech Republic, September 2004
- Oravec, M.; Beszédeš, M. & Rozinaj, G.(2008). Detection and Recognition of Human Faces and Facial Features, book chapter In: *Speech, Audio, Image and Biomedical Signal Processing Using Neural Networks*, Editors: Bhanu Prasad and S. R. Mahadeva Prasanna, page numbers 283-306, Springer-Verlag, Germany, ISBN 978-3-540-75397-1
- Oravec, M.; Mazanec, J.; Melišek, M. & Pavlovičová, J. (2009). Face Recognition Using Support Vector Machines, PCA and LDA with the FERET Database, *Acta Polytechnica Hungarica*. Submitted for publication in 2009.
- Pavlovičová, J.; Oravec, M.; Polec, J.; Keleši, M. & Mokoš, M. (2006). Error Concealment using Neural Networks for Block-based Image Coding. *Radioengineering*, Brno University of Technology, Dept. of Radio Electronics, Vol. 15, No. 2, June 2006, page numbers 30–36, ISSN 1210-2512
- Phillips, P. J.; Moon, H.; Rizvi, S. A. & Rauss, P. J. (2000). The FERET Evaluation Methodology for Face Recognition Algorithms, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 22, page numbers 1090-1104
- Phillips, P. J.; Wechsler, H.; Huang, J. & Rauss, P. (1998). The FERET database and evaluation procedure for face recognition algorithms, *Image and Vision Computing*, Vol. 16, No. 5, page numbers 295-306
- Polec, J.; Pohančenič, M.; Ondrušová, S.; Kotuliaková, K. & Karlubíková, T. (2009). Error Concealment for Classified Texture Images, *IEEE Region 8 EUROCON 2009*, pp.1348-1353, ISBN 978-1-4244-3861-7, St. Petersburg, Russia, May 2009, Piscataway : IEEE
- Reda, A.; Aoued, B. (2004). Artificial Neural Network-based Face Recognition, *Proc. of First International Symposium on Control, Communications and Signal Processing*, pp. 439-442
- Ross, A. A.; Nandakumar, K. & Jain, A. K. (2006). *Handbook of Multibiometrics*, Springer, ISBN: 978-0-387-22296-7
- Shoniregun, C. A. & Crosier, S. (2008). *Securing Biometrics Applications*, Springer, ISBN: 978-0-387-69932-5
- Truax, B. (1999). (Ed.) Gaussian noise In: *Handbook for Acoustic Ecology*, Available on: http://www.sfu.ca/sonic-studio/handbook/Gaussian_Noise.html Cambridge Street Publishing

- Turk, M. & Pentland, A. (1991). Eigenfaces for Recognition, *Journal of Cognitive Neuroscience*, Vol. 3, No. 1, 1991, page numbers 71-86.
- Uglov, J.; Jakaite, L.; Schetinina, V. & Maple, C. (2008). Comparing Robustness of Pairwise and Multiclass Neural-Network Systems for Face Recognition, *EURASIP Journal on Advances in Signal Processing*, Vol. 2008, Article ID 468693, doi:10.1155/2008/468693 <http://www.hindawi.com/journals/asp/2008/468693.html>
- Wang, J.; Lin, Y.; Yang, W. & Yang, J. (2008). Kernel Maximum Scatter Difference Based Feature Extraction and its Application to Face Recognition, *Pattern Recognition Letters*, Vol. 29, Issue 13, October 2008, page numbers 1832-1835
- Wang, Y.; Jiar, Y.; Hu, C. & Turk, M. (2004). Face Recognition Based on Kernel Radial Basis Function Networks, *Asian Conference on Computer Vision*, pp. 174-179, ISBN 10: 8995484209, Korea, January 2004, Asian Federation of Computer Vision Soc.
- Wheeler, F. W.; Liu, X.; Tu, P. H. & Hoctor, R. T. (2007). Multi-Frame Image Restoration for Face Recognition, *Proc. of IEEE Workshop on Signal Processing Applications for Public Security and Forensics, SAFE '07*, pp. 1-6, April 2007
- Yang, M. (2002). Kernel Eigenfaces vs. Kernel Fisherfaces: Face Recognition Using Kernel Methods, *IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 215-220, Mountain View, California
- Yang, M.; Frangi, A. F.; Yang, J. Y.; Zhang, D. & Jin, Z. (2005). KPCA Plus LDA: A Complete Kernel Fisher Discriminant Framework for Feature Extraction and Recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27, No. 2, February 2005, page numbers 230-244, ISSN: 0162-8828
- Zhao, W.; Chellappa, R. & Phillips, P. J. (1999). Subspace Linear Discriminant Analysis for Face Recognition, *Technical Report CAR-TR-914*, Center for Automation Research, University of Maryland, College Park



Face Recognition

Edited by Milos Oravec

ISBN 978-953-307-060-5

Hard cover, 404 pages

Publisher InTech

Published online 01, April, 2010

Published in print edition April, 2010

This book aims to bring together selected recent advances, applications and original results in the area of biometric face recognition. They can be useful for researchers, engineers, graduate and postgraduate students, experts in this area and hopefully also for people interested generally in computer science, security, machine learning and artificial intelligence. Various methods, approaches and algorithms for recognition of human faces are used by authors of the chapters of this book, e.g. PCA, LDA, artificial neural networks, wavelets, curvelets, kernel methods, Gabor filters, active appearance models, 2D and 3D representations, optical correlation, hidden Markov models and others. Also a broad range of problems is covered: feature extraction and dimensionality reduction (chapters 1-4), 2D face recognition from the point of view of full system proposal (chapters 5-10), illumination and pose problems (chapters 11-13), eye movement (chapter 14), 3D face recognition (chapters 15-19) and hardware issues (chapters 19-20).

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Milos Oravec, Jan Mazanec, Jarmila Pavlovicova, Pavel Eiben and Fedor Lehocki (2010). Face Recognition in Ideal and Noisy Conditions Using Support Vector Machines, PCA and LDA, Face Recognition, Milos Oravec (Ed.), ISBN: 978-953-307-060-5, InTech, Available from: <http://www.intechopen.com/books/face-recognition/face-recognition-in-ideal-and-noisy-conditions-using-support-vector-machines-pca-and-lda>

INTECH

open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2010 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](#), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.