Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 100 (2015) 1167 – 1173

Procedia Engineering

www.elsevier.com/locate/procedia

25th DAAAM International Symposium on Intelligent Manufacturing and Automation, DAAAM 2014

Evaluation of PCA, LDA and Fisherfaces in Appearance-Based Object Detection in Thermal Infra-Red Images with Incomplete Data

David Spulak*, Richard Otrebski, Wilfried Kubinger

UAS Technikum Wien: Department for Advanced Technologies & Mechatronics, Höchstädtplatz 6, 1200 Wien, Austria

Abstract

Object detection in outdoor environments is a challenging task. One is not only confronted with the problem of acquiring a sufficient amount of training images, but also the issue of huge variation in the objects appearance due to changing weather and light conditions. When using appearance-based object detection algorithms, such as in this paper, dimensional reduction of input data is an integral component to reduce computational costs and improve reliability. Based on the probabilistic classification method of Gaussian classifiers this paper examines the effect different dimensional reduction approaches have on the classification performance of thermal infra-red object images with respect to incomplete training data. It is shown that in the detection task at hand, which is to find the rear end of a truck in a thermal infra-red image, a reduction approach that combines principal component analysis (PCA) and linear discriminant analysis (LDA) is less sensitive to missing data.

Peer-review under responsibility of DAAAM International Vienna

Keywords: Principal component analysis; Linear discriminant analysis; Fisherfaces; Machine learning; Object detection

1. Introduction

In the KIRAS project *RelCon* [1] a reliable convoy of autonomous trucks is in development. For that purpose a dependable vision system is a fundamental requirement in order to track the vehicle to follow. The object detection that is the basis for such a system can be implemented in various ways including appearance-based approaches [2].

* Corresponding author. Tel.: +43-1-333-40-77-0; fax: +43-1-333-40-77-269.
E-mail address: spulak@technikum-wien.at

Peer-review under responsibility of DAAAM International Vienna
doi:10.1016/j.proeng.2015.01.480

A tracking system for vehicles in urban and unstructured environments was implemented by [3]. Their approach deploys one monocular visible light camera and makes extensive use of features such as Haar-like, local binary pattern and HOG, utilizing about 30,000 training images to acquire robust classifiers. [4] achieved fast classification in infra-red images by using the machine learning approach AdaBoost on about 3000 training images, deploying an extensive amount of computational power. The appearance-based based object detection algorithm proposed by [5] finds the rear end of a truck in thermal infra-red images, utilizing the Gaussian classifier. With the use of only 600 images, the training of reliable classifiers in this approach is computationally much cheaper compared to weak feature classifiers (e.g. Haar-like, etc.) and AdaBoost. However, the common problem any object detection task has in an outdoor environment is the continuously changing appearance of objects in it, due to changing light and weather conditions. Obtaining a comprehensive training data set is difficult and the encounter of untrained conditions is always possible.

When utilizing appearance-based approaches for classification the dimensional reduction of input images is an essential part of any algorithm. With it, the algorithms robustness and computational complexity can be dramatically improved. In this paper three different methods for dimensional reduction are examined, with special regard to the object detection task in the before mentioned RelCon project and its sensitivity to incomplete training data.

This paper is structured as follows: In Section 2 the object detection task at hand is described in more detail, the current algorithm is explained and the different methods for dimensional reduction are given. After that, in Section 3 the test results will be shown. Finally conclusions are drawn in Section 4.

2. The detection task and current algorithm

In the object detection system that is developed in the RelCon project the rear end of a truck is to be detected within a thermal infra-red image of the size 320×240 pixels. For training purposes a dataset of 6,682 positive training images, which depict the truck in different thermal conditions and environments, and 5,295 negative training images, showing the trucks background scenery, are available. All training images have the size: 50×50 pixels. The positive training images are segmented in three separate datasets representing different hours of the day and thermal conditions, henceforth referred to as *pos_1*, *pos_2* and *pos_3*. The complete negative dataset will be referred to as *neg* in the rest of this paper. Fig. 1 shows a sample of these datasets in their original resolution.

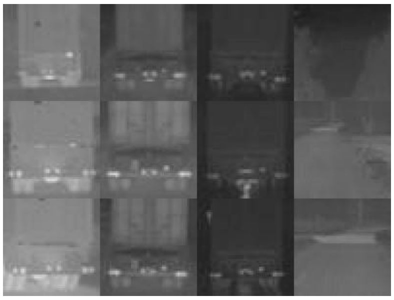

Fig. 1. Training datasets (from left to right column: *pos_1*, *pos_2*, *pos_3, neg*).

Within the detection algorithm that was developed in [5] new camera images are segmented into subimages of varying sizes between 200×200 pixels and 50×50 pixels which are scaled down to the same size as the training images. Subsequently all subimages are then classified by an appearance-based approach which allows the localization of the object within the whole camera image.

In appearance-based object classification methods images of the size $n \times m$ are usually represented by a vector of the corresponding size $n \cdot m$. This means, however, that even with very small images all calculations necessary for classification have to be done in high dimensions (In case of the image sizes in this paper: 2500 dimensions). Dimensional reduction of this image data is crucial to reduce the computational cost and improve the robustness of the algorithm it is used in (avoidance of overfitting) [6]. In the current detection algorithm PCA is applied for dimensional reduction [7], [8], [9]. In general PCA projects input data into an orthogonal space, according to the datas variance. This means that the first principal component axis points in the direction of the maximum variance within the data PCA is performed on, the second principal component axis points in the direction of the second highest variance and so on. Therefore, axes q_1 to $q_{n \cdot m}$ have descendent discriminatory power. By selecting the first l principal component axes q_1 to q_l a reduced PCA-space can be chosen. Data can then be projected onto this reduced space of l dimensions.

For both the positive and negative dataset maximum likelihood estimators for the mean and covariance matrix ($\vec{\mu}_{pos}, \Sigma_{pos}, \vec{\mu}_{neg}, \Sigma_{neg}$) can be calculated after projecting the training data into the reduced space. Each subimage can then be classified using the Gaussian classifier [9] (see Eq. 1) [5]:

$$log\big(p(c_i|\vec{x})\big) = -\tfrac{1}{2}(\vec{x}-\vec{\mu}_i)^T \Sigma_i^{-1}(\vec{x}-\vec{\mu}_i) - \tfrac{1}{2}log(|\Sigma_i|) + k, \tag{1}$$

where c_i is class i, $\vec{\mu}_i$ is the mean image of class i, Σ_i is the covariance matrix of class i and $\vec{x}$ is the new image segment that is classified. The constant k is class-independent, thus identical in all classes c_i of this problem. New images are assigned to the class for which their logarithmic likelihood (log $(p\big(c_{pos}|\vec{x}\big)$, log $(p(c_{neg}|\vec{x})))$ is largest – either the positive or the negative class.

2.1. LDA and Fisherfaces

PCA is an unsupervised algorithm, which means that it tries to find discriminatory information, without prior knowledge of any classes that are part of the data. On the other hand LDA is a supervised algorithm that tries to find a projection in which labelled data is optimally separated. To achieve this, LDA optimizes a *within-class* measure, that describes the covariance of the data of each class, and a *between-class* measure, which shows the relation of the class means (see Eq. 2, Eq. 3) [6], [8], [10]:

$$S_W = \textstyle\sum_{n \in C_1}(\vec{x}_n - \vec{\mu}_1)\cdot(\vec{x}_n - \vec{\mu}_1)^T + \sum_{n \in C_2}(\vec{x}_n - \vec{\mu}_2)\cdot(\vec{x}_n - \vec{\mu}_2)^T, \tag{2}$$

$$S_B = (\mu_1 - \mu_2)\cdot(\mu_1 - \mu_2)^T, \tag{3}$$

where S_W is the *within-class scatter matrix* and S_B is the *between-class scatter matrix*. The optimal projection matrix W^* can be found by solving the generalized eigenvalue problem of Eq. 4 [10]:

$$S_W^{-1} S_B W^* = \lambda W^* \tag{4}$$

Many object recognition tasks suffer from the small sample size problem in which the training data set is significantly smaller than the dimensionality of the sample space. In case of LDA this results in a within-class scatter matrix S_W that is singular, thus its invers cannot be calculated. To tackle this problem several alternatives are applicable, one of which is adding a regularizing term to S_W making the within-class scatter matrix non-singular (see Eq. 5) [11]:

$$S_W' = S_W + \epsilon I, \tag{5}$$

where ϵ is a small constant and I is the identity matrix. Another approach is called Fisherfaces, a term that originated from the research field of face recognition. At most the matrix S_W has a rank of $N - c$, where N is the number of images in the training set and c is the number of classes. The idea of Fisherfaces is to first reduce the training data to at least $N - c$ dimensions utilizing PCA. After that LDA is used to reduce the projected data further (see Eq. 6) [10], [12]:

$$W_{opt}^T = W_{LDA}^T W_{PCA}^T, \tag{6}$$

where W_{PCA}^T is the projection onto the reduced PCA-space and W_{LDA}^T the subsequent projection onto the further reduced LDA-space. Together this gives the projection W_{opt}^T which combines both PCA and LDA.

2.2. Dimensional reduction

In this paper 1,000 training images are used for dimensional reduction, which brings up the small sample size problem mentioned earlier. PCA is not affected by this, however, any reduction method utilizing LDA is and has to be treated accordingly. Therefore, the regularized LDA and PCA-LDA combination of Fisherfaces will be examined in the following. Fig. 2 shows all available training data reduced to three dimensions using different methods of dimensional reduction. When reducing the dimensions with PCA it can be observed, that the different datasets are separated clearly (see Fig. 2a), even those of the same class. This leads to the assumption, that if any of these datasets were missing during the training of classifiers, big errors will be made when assigning objects of the missing datasets to a specific class. In contrast, when utilizing LDA for dimensional reduction even datasets with big thermal differences that are part of the same class are forced together. It can be anticipated that in this case the lack of one of the positive datasets would have little effect on the classification performance. This effect can be seen in Fig. 2b where LDA with regularized within-class scatter matrix was used. However, a large class variance and quite a class overlap of negative and positive images are noticeable. The reason for this is most likely that relative arbitrary features are used to distinguish the positive and negative training images (see Fig. 3). Especially if a small training set is used for calculating the LDA-space the data variance is high after reduction. Utilizing both PCA and LDA for dimensional reduction, similar to Fisherfaces this variance can be significantly reduced (see Fig. 2c). Also the class overlap of the positive and negative class is much lower. The effect of missing training data on PCA, LDA and the combined PCA-LDA approach will be shown in the results section (Section 3) of this paper.

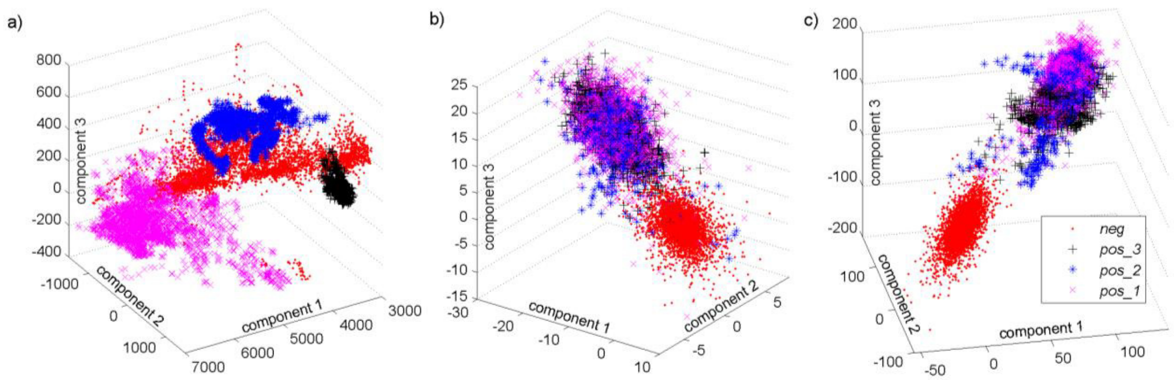

Fig. 2. Training data reduced to 3 dimensions with (a) PCA, (b) LDA regularized, (c) PCA & LDA combined (Fisherfaces).

The eigenvectors calculated with PCA, LDA and Fisherfaces can be seen in Fig. 3. The first three components, derived from the respective dimensional reduction methods are shown here and can also be interpreted as the extracted features. As it can be observed, the features of the truck are easily distinguished in the first three principal components (see Fig. 3a). On the other hand, the features from LDA are very noisy (see Fig. 3b). Fisherfaces also show noisy images, however, the contours of the truck are still recognizable (see Fig. 3c).

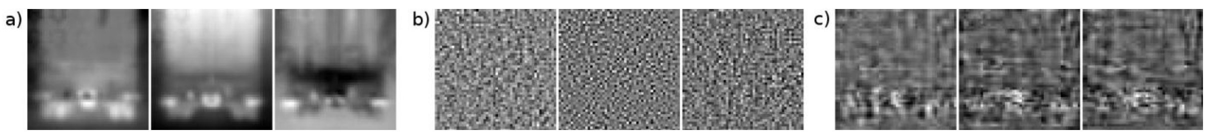

Fig. 3. First three eigenvectors derived from (a) PCA, (b) LDA, (c) Fisherfaces.

3. Results

In order to get to a valid performance study four evaluations were done in an experimental setting, in which the datasets mentioned in Section 2 (*pos_1*, *pos_2*, *pos_3* and *neg*) were used. For the first evaluation 500 positive and 500 negative randomly selected training images from all available datasets were used to derive the projection into lower space (PCA, LDA and Fisherfaces) and to train the Gaussian classifiers. Following this, a cross evaluation was conducted ignoring one positive dataset at a time. In the first cross evaluation dataset *pos_1* (3,148 images) was neglected, again using 500 positive and 500 negative training images for dimensional reduction and classifier training. Similarly the second and third cross evaluations were performed, disregarding datasets *pos_2* (1,571 images) and *pos_3* (1,963 images) respectively. Additionally to the standard Fisherface method the training data was not only reduced to $N - c$ dimensions (998 in this study) before applying LDA, but also to 500, 250 and 100 dimensions. They are labelled *Fish (998)*, *Fish (500)*, *Fish (250)* and *Fish (100)* in the following Tables 1 to 4.

Tab. 1 shows the number of classification errors (number of images assigned to the wrong class: false positives, false negatives) made in each dataset when reduced to three and 20 dimensions. As it can be seen in higher dimensions the classification performance of all dimensional reduction methods are similar. However, if the data is reduced to low dimensions the classification performance is significantly better with an approach that uses LDA.

Table 1. Evaluation 1 (errors in 3 dim. / 20 dim.).

dataset	PCA	LDA	Fish (998)	Fish (500)	Fish (250)	Fish (100)
pos_1	213 / 10	29 / 2	17 / 9	16 / 8	15 / 8	14 / 7
pos_2	0 / 11	29 / 18	20 / 10	18 / 13	19 / 11	15 / 0
pos_3	227 / 0	27 / 7	19 / 5	6 / 4	7 / 4	10 / 3
neg	238 / 14	31 / 16	17 / 5	0 / 1	0 / 3	5 / 7

In Tables 2, 3 and 4 the results of the cross evaluations are shown. In this study PCA seems to be most sensitive to the lack of dataset *pos_1* (see Tab. 2). If this set is disregarded almost all images within this set are misclassified when reduced to three dimensions. Even with 20 dimensions the about 30 % of all images in this set are assigned to the wrong class. With regularized LDA this can be improved significantly to a misclassification rate of around 10 % in three and about 13 % in 20 dimensions. The Fisherface method outperforms PCA, especially with highly reduced data. Standard Fisherfaces bring the misclassification rate down to about 12 % in three dimensions and to about 8 % in 20 dimensions. If the data reduction with PCA is intensified and images are reduced to less than 998 dimensions before using LDA, the classification performance can be improved even further. The best results in this test are achieved by projecting the training images into 500 dimensional PCA-space before applying LDA. Here the error rate is diminished to around 5 % in three dimensions and about 3 % in 20 dimensions.

Table 2. Cross evaluation 1 (errors in 3 dim. / 20 dim.).

dataset	PCA	LDA	Fish (998)	Fish (500)	Fish (250)	Fish (100)
pos_1	3,125 / 693	326 / 402	379 / 278	162 / 95	106 / 156	64 / 552
pos_2	45 / 0	19 / 11	15 / 10	9 / 7	16 / 6	11 / 0
pos_3	18 / 0	6 / 4	7 / 1	1 / 1	1 / 1	1 / 1
neg	513 / 1	11 / 1	7 / 1	1 / 2	2 / 0	3 / 2

With cross evaluation two (see Tab. 3), where training set *pos_2* is neglected, PCA has a relatively high misclassification error of around 21 % in 20 dimensions within the disregarded dataset. Surprisingly the error rate in 3 dimensions is only 1.5 %. When utilizing LDA alone, the error rate improves significantly to only about 4 % in 20 dimensions. The combined PCA-LDA approach lowers the classification error rate to below 6 % in all tested configurations. With a PCA reduction to 250 dimensions before using LDA, the misclassification rate is reduced to only about 2 %.

Table 3. Cross evaluation 2 (errors in 3 dim. / 20 dim.).

dataset	PCA	LDA	Fish (998)	Fish (500)	Fish (250)	Fish (100)
pos_1	80 / 4	20 / 4	9 / 5	6 / 3	5 / 2	5 / 3
pos_2	26 / 329	72 / 65	75 / 87	87 / 69	84 / 39	90 / 72
pos_3	11 / 0	10 / 7	11 / 8	11 / 11	10 / 8	7 / 8
neg	66 / 0	26 / 7	16 / 4	3 / 7	3 / 6	8 / 14

In the third cross evaluation (see Tab. 4), surprising results were encountered in higher dimensions. Using PCA all images from dataset *pos_3* are assigned to the wrong class in three dimensions. However, the error is reduced to only 1.5 % in 20 dimensions. LDA has an error rate of about 20 % in three and 10 % in 20 dimensions. Also the Fisherface approach performs worse than PCA in high dimensions, with error rates between 10 % and 3 %.

Table 4. Cross evaluation 3 (errors in 3 dim. / 20 dim.).

dataset	PCA	LDA	Fish (998)	Fish (500)	Fish (250)	Fish (100)
pos_1	74 / 2	49 / 10	34 / 7	14 / 8	14 / 5	13 / 4
pos_2	11 / 7	84 / 35	76 / 12	24 / 13	23 / 9	3 / 0
pos_3	1,963 / 30	389 / 185	210 / 91	164 / 142	191 / 104	265 / 50
neg	77 / 2	22 / 7	12 / 4	1 / 3	3 / 6	1 / 13

The evaluations show that LDA and Fisherfaces are more consistent than PCA, when confronted with incomplete data. This corresponds with the results in [12], investigating dimensional reduction methods in visible-light face recognition under variable illumination. Taking the cross-evaluation results into account, tests were carried out in the object detection algorithm at hand using complete and incomplete training data. The training data was reduced to 20 dimensions by PCA, regularized LDA and the Fisherfaces method. For the Fisherfaces approach the data was first reduced to 250 dimensions with PCA, before being reduced to its final 20 dimensions with LDA.

Fig. 4 shows the output of the object detection algorithm, with the object detected by the algorithm marked by a white square and the number 0. All other matches, that where declared as positive matches, but have a lower logarithmic likelihood (compare Eq. 1) are marked with black squares and are numbered according to their score.

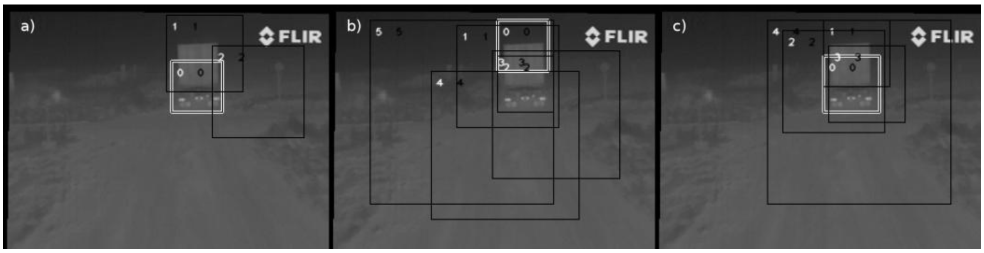

Fig. 4. Found matches in object detection algorithm with projection space (a) PCA, (b) LDA, (c) Fisherfaces (250).

As it can be seen the algorithm working with PCA shows very few false positive matches. Except the final match only two other possible matches are found (see Fig. 4a). Using the pure LDA approach the detection results show a significant increase of false positive detections (see Fig. 4b). This is most likely due to the before mentioned relative arbitrariness of the features used in LDA to distinguish positive and negative samples. Using only LDA to reduce images is feasible in this application since the best match is still the truck. However, the matches the algorithm finds are distributed much wider than with the algorithm relying on PCA only, which makes the accuracy of this approach very poor. Finally the combined dimensional reduction approach shows more false positive matches than the approach using only PCA but less than with pure LDA. All declared matches contain the truck and a convergence of the different sized matches towards the final match can be observed (see Fig. 4c).

When using classifiers that were trained with incomplete data (e.g. without dataset *pos_1*) tests show, that the performance of the algorithm using Fisherfaces for dimensional reduction is hardly influenced. The reliability of the object detection is practically the same, even in images of untrained thermal conditions. Similarly, but less accurate, the regularized LDA approach also performs almost identically with and without dataset *pos_1*. On the other hand, the detection performance is strongly reduced when pure PCA is used for data reduction. The already poor object detection rate is highly dependent on the number of subimages the whole image is fragmented into. Increasing the number of subimages improves the detection rate slightly but significantly raises the computational requirements of the algorithm. The results of the Fisherface approach cannot be matched.

4. Conclusion

In object detection tasks dealing with the unstructured, ever-changing outdoor environment, getting a comprehensive representation of all possible situation in the form of training data is difficult. Therefore, depending on how representative the training data is, compared to the conditions likely to encounter, knowledge of how different dimensional reduction methods deal with incomplete data is essential. The effects these methods have on appearance-based object classification in infra-red images were examined in this paper. Special focus was put on the algorithms performance when confronted with incomplete training data due to changing thermal conditions of the environment. For classification the Gaussian classifier was used.

It shows, that a combined approach using both PCA and LDA for dimensional reduction, the so-called Fisherfaces, is less sensitive to changing thermal and weather conditions in infra-red images and therefore more reliable when the problem of incomplete training data is encountered. However, if the data is complete, the results of the algorithm using PCA for data reduction is more reliable. Determined by the confidence the designer of an object detection algorithm has in the completeness of his training data, either PCA or Fisherfaces should be used to achieve the best possible results. The next steps regarding the detection task at hand will be further test on the real system in order to determine how general the collected training data is and which dimensional reduction method to put to use.

Acknowledgements

This project has been funded by the Austrian Security Research Programme KIRAS – an initiative of the Austrian Federal Ministry for Transport, Innovation and Technology (bmvit).

References

[1] FFG – Österreichische Forschungsgesellschaft mbH. (2014) Reliable control of semi-autonomous platforms (relcon). [Online]. Available: http://www.kiras.at/gefoerderte-projekte/detail/projekt/reliable-control-of-semi-autonomous-platforms-relcon/.

[2] S. Sivaraman, M. M. Trivedi, Looking at vehicles on the road: a survey of vision-based vehicle detection, tracking, and behavior analysis, Intelligent Transportation Systems, IEEE Transactions on, vol. 14, no. 4, Dec. 2013, pp. 1773-1795.

[3] C. Fries, T. Luettel, H.-J. Wuensche, Combining model- and template-based vehicle tracking for autonomous convoy driving, in: Intelligent Vehicles Symposium (IV), 2013 IEEE, June 2013, pp. 1022-1027.

[4] W. Woeber, M. Kefer, W. Kubing, and D. Szuegyi, Evaluation of daylight and thermal infar-red based detection for platooning vehicles, in: B. Katalinic (Eds.), Annals of DAAAM for 2012 & Proceedings of the 23rd International DAAAM Symposium, DAAAM International, Vienna, Austria, 2012, pp. 719-722.

[5] W. Woeber, D. Szuegyi, W. Kubinger, and L. Mehnen, A principal component analysis based object detection for thermal infra-red images, in: ELMAR, 2013 55th Internationial Symposium, Sept. 2013, pp. 357-360.

[6] A. M. Martinez and A. Kak, Pca versus lda, Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 23, no. 2, Feb. 2001, pp. 228-233.

[7] J. Shlens, A tutorial on principal component analysis, CoRR, vol. abs/1404.1000, 2014. [Online]. Available: http://arxiv.org/abs/1404.1000

[8] C. M. Bishop, Pattern recognition and machine learning, Springer, New York, 2006.

[9] A. R. Webb, K. D. Kopsey, Statistical Pattern Recognition, third ed., John Wiley & Sons, Hoboken, 2011.

[10] C. Zhou, L. Wang, Q. Zhang, X. Wei, Face recognition based on pca image reconstruction and lda, Optik – International Journal for Light and Electron Optics, vol. 124, no. 22, Nov. 2013, pp. 5599-5603.

[11] A. Martinez (2011) Fisherfaces. Scholarpedia, 6(2):4282., revision #91266.

[12] P. Belhumeur, J. Hespanha, and D. Kriegman, Eigenfaces vs. fisherfaces: recognition using class specific linear projection, Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 19, no. 7, Jul 1997, pp. 711-720.