

PCA-Based Face Recognition: Similarity Measures and Number of Eigenvectors

Sushma Niket Borade^(✉) and Ratnadeep R. Deshmukh

Department of Computer Science and Information Technology,
Dr. Babasaheb Ambedkar Marathwada University,
Aurangabad 431004, India
sushma.borade@gmail.com,
rrdeshmukh.csit@bamu.ac.in

Abstract. This paper examines the performance of face recognition using Principal Component Analysis by (i) varying number of eigenvectors; and (ii) using different similarity measures for classification. We tested 15 similarity measures. ORL database is used for experimentation work which consists of 400 face images. We observed that changing similarity measure causes significant change in the performance. System showed best performance using following distance measures: Cosine, Correlation and City block. Using Cosine similarity measure, we needed to extract lesser images (30 %) in order to achieve cumulative recognition of 100 %. The performance of the system improved with the increasing number of eigenvectors (till roughly 30 % of eigenvectors). After that performance almost stabilized. Some of the worst performers are Standardized Euclidean, Weighted Modified SSE and Weighted Modified Manhattan.

Keywords: Biometrics · Face recognition · Principal component analysis · Distance measures

1 Introduction

There are many types of personal authentication systems and face recognition is one of the active research areas since last several decades. Several methods have been proposed to recognize faces [1–3]. There are two main categories of face recognition methods: feature-based and appearance-based [1]. Using appearance-based methods, a face image of size $N \times N$ pixels is represented by a vector in N^2 dimensional space. Practically, these spaces are too large to perform robust and fast recognition of faces. To solve this problem, dimensionality reduction is done using Principal Component Analysis (PCA) technique. In 1987, PCA was first used to represent face images by Sirovich and Kirby [4]. Turk and Pentland applied PCA to face recognition and presented eigenfaces method in 1991 [5]. We study the effect of 15 similarity measures on the performance of face recognition using PCA. Following characteristics are used to measure the system performance: area above cumulative match characteristics (CMC), rate of recognition and percent of images needed to extract to achieve cumulative recognition of 100 %.

Organization of the paper is as follows: In Sect. 2, we present face recognition using PCA technique in detail. Various similarity measures are described in Sect. 3. In Sect. 4, experimental work and the results obtained are presented. Section 5 offers the conclusion.

2 Principal Component Analysis

We implemented face recognition using PCA as proposed by Turk and Pentland [5]. Let the gallery set of M face images be $\Gamma_1, \Gamma_2, \dots, \Gamma_M$. The average face image of the whole set is defined by

$$\Psi = \frac{1}{M} \sum_{i=1}^M \Gamma_i \quad (1)$$

Each face image differs from the average face, Ψ , by the vector $\phi_i = \Gamma_i - \Psi$, where $i = 1$ to M . Find covariance matrix C as

$$C = AA^T, \text{ where matrix } A = [\phi_1 \phi_2 \dots \phi_M]. \quad (2)$$

Matrix C is of size N^2 by N^2 . It's computationally expensive to find its N^2 eigenvectors. Therefore form M by M matrix, $L = A^T A$ and get its M eigenvectors, v_i . The most significant M eigenvectors of C are found as:

$$u_l = \sum_{k=1}^M v_{lk} \phi_k, l = 1, \dots, M \quad (3)$$

Now from these M eigenvectors, consider M' ($< M$) eigenvectors (with the highest M' eigenvalues). New probe image is projected into the facespace using operation:

$$w_k = u_k^T (\Gamma - \Psi) \quad (4)$$

for $k = 1, \dots, M'$.

These values of w form projection vector $\Omega = [w_1, w_2, \dots, w_{M'}]$. Probe image is then classified as belonging to closest face class by using some similarity measure.

3 Similarity Measures

Consider two feature vectors x and y of dimensions n each. The distances between these feature vectors can be calculated as [6–11]:

1. City block distance (or Manhattan distance):

$$d(x, y) = \sum_{i=1}^n |x_i - y_i| \quad (5)$$

2. Euclidean distance:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (6)$$

3. Squared Euclidean distance (Sum square error, SSE):

$$d(x, y) = \sum_{i=1}^n (x_i - y_i)^2. \quad (7)$$

4. Mean square error (MSE):

$$d(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2 \quad (8)$$

5. Cosine distance:

$$d(x, y) = -\frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}} \quad (9)$$

6. Mahalanobis distance:

$$d(x, y) = \sqrt{(x - y)S^{-1}(x - y)^t} \quad (10)$$

where S is the covariance matrix of the distribution.

7. Standard Euclidean distance:

$$d(x, y) = \sqrt{(x - y)V^{-1}(x - y)^t} \quad (11)$$

where V is the n by n diagonal matrix whose j^{th} diagonal element is $S(j)^2$, where S is the vector of standard deviations.

8. Minkowski distance:

$$d(x, y) = \left(\sum_{i=1}^n |x_i - y_i|^p\right)^{1/p}. \quad (12)$$

where p is a scalar exponent and here $p > 0$.

9. Chebychev distance:

$$d(x, y) = \max_i \{|x_i - y_i|\} \quad (13)$$

10. Correlation distance:

$$d(x, y) = -\frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sqrt{(n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2)(n \sum_{i=1}^n y_i^2 - (\sum_{i=1}^n y_i)^2)}} \quad (14)$$

11. Canberra distance:

$$d(x, y) = \sum_{i=1}^n \frac{|x_i - y_i|}{|x_i| + |y_i|} \quad (15)$$

12. Modified SSE distance:

$$d(x, y) = \frac{\sum_{i=1}^n (x_i - y_i)^2}{\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i^2} \quad (16)$$

13. Modified Manhattan distance:

$$d(x, y) = \frac{\sum_{i=1}^n |x_i - y_i|}{\sum_{i=1}^n |x_i| \sum_{i=1}^n |y_i|} \quad (17)$$

14. Weighted Modified SSE distance:

$$d(x, y) = \frac{\sum_{i=1}^n z_i (x_i - y_i)^2}{\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i^2}, \quad z_i = \sqrt{1/\lambda_i} \quad (18)$$

where λ_i are eigenvalues.

15. Weighted Modified Manhattan distance:

$$d(x, y) = \frac{\sum_{i=1}^n z_i |x_i - y_i|}{\sum_{i=1}^n |x_i| \sum_{i=1}^n |y_i|}, \quad z_i = \sqrt{1/\lambda_i} \quad (19)$$

In this paper we perform identification task. CMC curve graphically represents the performance of the identification system. It is a plot of rank values on the X axis and probability of correct identification at or below that rank on the Y axis [10, 12].

4 Experiments and Results

We tested performance of the system using ORL face image database [13]. Face images are taken at the AT and T Laboratories between April 1992 and April 1994. ORL database contains images of 40 different persons, with 10 images per person. The face images are taken at different times by varying lighting, facial details and facial expressions. All face images are frontal with some pose variation. Each image is of size 112×92 pixels with 256 gray levels.

For experimental work, each face image is resized to 50×40 pixels. Training is done by considering first 5 images and testing is done considering remaining five images per person. This gave us gallery set of 200 images and probe set of 200 images. Experiments are implemented using MATLAB® R2013a. We used nearest mean rule in which we computed a template for each identity in the database. The closest identity is chosen as the match.

The experimental results with 15 similarity measures are listed in Tables 1, 2 and 3. The performance for face recognition is measured by calculating the area above cumulative match characteristic curve (CMCA). If CMCA is smaller, it indicates better recognition performance. We show how many images (in percents) should be extracted to get cumulative recognition rate between 80 to 100 %. If it is smaller, it implies that fewer images need to be extracted to get required cumulative recognition rate. We also find recognition rate that is achieved if the closest match is extracted from the system. If this first one recognition rate is more, it indicates better results. In Tables 1, 2 and 3, we used subscripts to mark the best results.

Table 1. Performance using 20 % of features (40).

Similarity measure	Rank (%) of images					CMCA 0–10000	First 1 rec
	80	85	90	95	100		
City block	1.62	3.09	4.56	13.75	50 ₄	401.25 ₄	166 ₄
Eucl; SSE; MSE	1.84	3.16	4.47	8.75	37.5 ₂	369.38 ₃	165 ₅
Cosine	0.94	2.5	4.06	8.75	30 ₁	336.25 ₁	170 ₁
Mahalanobis	3.90	5.83	13	30	60 ₅	583.13	151
Std Euclidean	2.31	4.23	10	18.75	67.5	510.63	161
Minkowski	1.84	3.16	4.47	8.75	37.5 ₂	369.38 ₃	165 ₅
Chebychev	4.74	6.25	9.38	20	72.5	516.25	134
Correlation	1.25	2.64	4.03	10	30 ₁	336.88 ₂	169 ₂
Canberra	4.17	6.11	12.5	30	92.5	623.75	148
Modified SSE	−2	3	7.5	23.75	90	531.88	169 ₂
ModManhattan	−0.8	3.33	8.5	16.25	47.5 ₃	431.25 ₅	168 ₃
WeigMod SSE	3.61	6.79	18.33	40	95	728.75	156
WeigMod Man	3.54	8.5	22.5	40	80	740.63	155

Table 2. Performance using 60 % of features (120).

Similarity measure	Rank (%) of images					CMCA 0–10000	First 1 Rec
	80	85	90	95	100		
City block	−0.21	1.88	3.96	9	35 ₂	344.38 ₃	173 ₁
Eucl; SSE;MSE	1.32	2.80	4.26	8.75	35 ₂	358.75 ₄	168 ₄
Cosine	0.88	2.35	3.82	7.5	30 ₁	325.63 ₂	171 ₃
Mahalanobis	16.25	26.5	38.75	67.5	92.5	1215.6	127
Std Euclidean	11	14.25	26.67	43.75	100	915.63	131
Minkowski	1.32	2.80	4.26	8.75	35 ₂	358.75 ₄	168 ₄
Chebychev	4.74	6.25	9.38	20	72.5 ₄	516.25	134
Correlation	0.74	2.21	3.68	6.25	30 ₁	322.5 ₁	172 ₂
Canberra	10	14.06	28.75	50	97.5	950	132
Modified SSE	0	4.17	8.5	21.25	82.5 ₅	536.25	166
ModManhattan	0.91	3.18	5.71	12.5	45 ₃	385.63 ₅	167 ₅
WeigMod SSE	11.5	21.67	35	62.5	95	1095.6	139
WeigMod Man	15	28.75	40	54.38	87.5	1123.8	114

Table 3. Performance using 90 % of features (180).

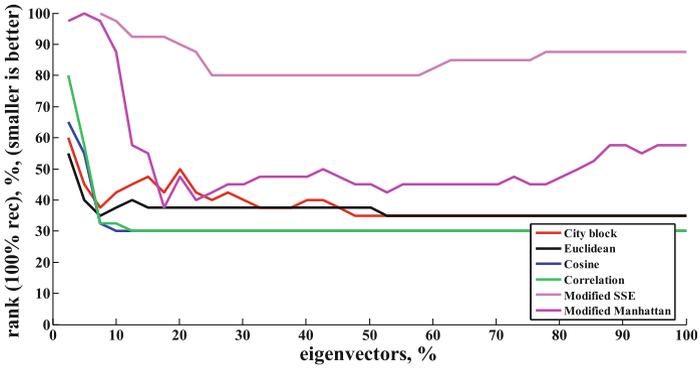
Similarity measure	Rank (%) of images					CMCA 0–10000	First 1 rec
	80	85	90	95	100		
City block	–1.25	1.25	3.75	9	35 ₂	340.63 ₃	175 ₁
Eucl; SSE; MSE	1.18	2.65	4.12	8.75	35 ₂	356.88 ₄	169 ₄
Cosine	0.88	2.35	3.82	6.25	30 ₁	323.13 ₁	171 ₃
Mahalanobis	9.46	12.5	25	34.17	77.5 ₅	838.13	97
Std Euclidean	27.5	34.5	45	60	100	1538.1	87
Minkowski	1.18	2.65	4.12	8.75	35 ₂	356.88 ₄	169 ₄
Chebychev	4.74	6.25	9.38	20	72.5 ₄	516.25	134
Correlation	0.63	2.19	3.75	6.67	30 ₁	323.75 ₂	172 ₂
Canberra	16.25	25.83	37.5	62.5	92.5	1212.5	110
Modified SSE	1.07	4.64	9.06	23.75	87.5	561.25	164
ModManhattan	1	3.5	6.67	15.83	57.5 ₃	422.5 ₅	166 ₅
WeigMod SSE	15	28.75	50	68.75	97.5	1323.8	124
WeigMod Man	28.13	36.67	50	65	90	1516.9	87

Table 4 reports the sorted similarity measures w.r.t the performance of the system using following characteristics: (i) recognition rate, (ii) overall recognition accuracy (i.e. CMCA), (iii) images (in percent) extracted to get 100 % cumulative identification.

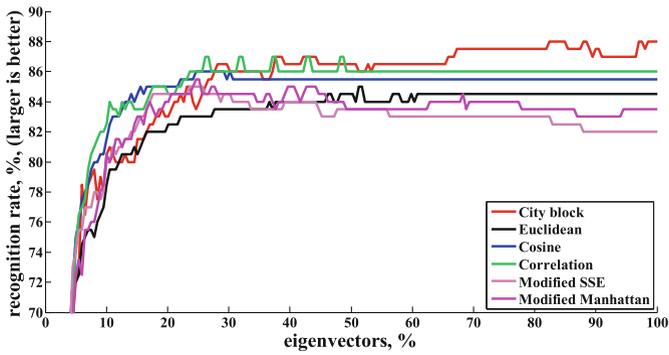
Table 4. Sorted similarity measures with respect to the performance of the system.

No. of eigenvectors	First1	CMCA	Cum100
20 % (40)	Cosine	Cosine	Cosine, Correlation
	Correlation, Mod. SSE	Correlation	Euclidean
	Modified Manhattan	Euclidean	Modified Manhattan
60 % (120)	City block	Correlation	Cosine, Correlation
	Correlation	Cosine	City block, Eucl.
	Cosine	City block	Modified Manhattan
90 % (180)	City block	Cosine	Cosine, Correlation
	Correlation	Correlation	City block, Eucl.
	Cosine	City block	Modified Manhattan

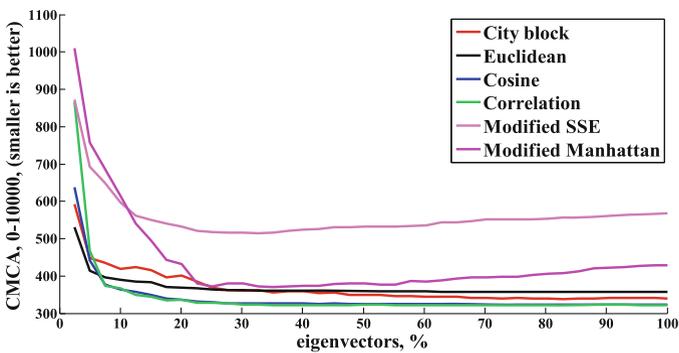
The training set has 200 images (5 images per person). This produces 199 eigenvectors as there will be only $(M - 1)$ meaningful eigenvectors with remaining eigenvectors having associated eigenvalues of zero [5]. Figure 1 shows the variation in the system performance with number of eigenvectors for top 6 performers. We need to extract fewer images (30 %) to achieve 100 % cumulative recognition using Cosine distance measure if 10–90 % of eigenvectors are used and correlation if 30–90 % of eigenvectors are used. Performance with respect to rank increases till 10 % and 12 % of eigenvectors for Cosine and Correlation similarity measures respectively, and later it stabilizes.



(a)



(b)



(c)

Fig. 1. Performance of the system with respect to number of eigenvectors (a) Cumulative 100 % recognition, (b) Recognition rate, (c) CMCA (Color figure online)

System achieved best performance using Cosine similarity measure (336.25–364.38) if 10–20 % of eigenvectors are used, 323.13 if 90 % of eigenvectors are used and Correlation (322.5–323.75) if 30–60 % of eigenvectors are used. For CMCA, top performance is shown by Correlation and Cosine similarity measures. Performance with these measures increases until approximately 30 % of eigenvectors are used, and then performance almost stabilizes.

We achieved largest recognition rates using Correlation (82.5–86 %) if 10–30 % of eigenvectors are used, Cosine distance (85–86 %) if 20–30 % of eigenvectors are used and City block distance (86–87.5 %) if 30–90 % of eigenvectors are used. Best recognition rate is achieved using City block measure. It shows increase in recognition rate with number of eigenvectors. The variation in the performance indicates that selecting similarity measure is a critical decision in designing a PCA-based face recognition.

5 Conclusions

This paper investigates 15 different similarity measures for face recognition using PCA. We examined the performance of the system by varying the number of eigenvectors. The experiments are conducted on ORL face database which has 400 face images. The best identification performance is reported using following similarity measures: Cosine, Correlation and City block. Using Cosine distance we need to extract fewer images to achieve 100 % cumulative recognition than using any other similarity measure. This research shows the effect of similarity measures on the performance of the system. It is observed that, as number of eigenvectors increased, recognition rates also increased. This observation is consistent with prior studies [10]. Performance of the system increased till roughly 30 % of eigenvectors. After that, it almost stabilized. Standardized Euclidean, Weighted Modified SSE and Weighted Modified Manhattan are worst performers to name a few with CMCA of (510.7–1538), (687.5–1323.8) and (734.4–1516.9) for 10–90 % of eigenvectors.

References

1. Zhao, W., Chellappa, R., Phillips, P.J., Rosenfeld, A.: Face recognition: a literature survey. *ACM Comput. Surv.* **35**, 399–458 (2003)
2. Jafri, R., Arabnia, H.R.: A survey of face recognition techniques. *J. Inf. Process. Syst.* **5**, 41–68 (2009)
3. Rao, A., Noushath, S.: Subspace methods for face recognition. *Comput. Sci. Rev.* **4**, 1–17 (2010)
4. Sirovich, L., Kirby, M.: Low-dimensional procedure for the characterization of human faces. *J. Opt. Soc. Am.* **4**, 519–524 (1987)
5. Turk, M., Pentland, A.P.: Eigenfaces for recognition. *J. Cogn. Neurosci.* **3**, 71–86 (1991)
6. Yambor, W., Draper, B.: Analyzing PCA-based face recognition algorithm: eigenvector selection and distance measures. In: Christensen, H., Phillips, J. (eds.) *Empirical Evaluation Methods in Computer Vision*, pp. 39–51. World Scientific Press, Singapore (2002)

7. Mathwork Help for Similarity Measures. <http://www.mathworks.com/help/stats/pdist2.html>
8. Perlibakas, V.: Distance measures for PCA-based face recognition. *Pattern Recogn. Lett.* **25**, 711–724 (2004)
9. Miller, P., Lyle, J.: The effect of distance measures on the recognition rates of PCA and LDA based facial recognition. In: *Digital Image Processing*. Clemson University, Clemson
10. Moon, H., Phillips, P.J.: Computational and performance aspects of PCA-based face recognition algorithms. *Perception* **30**, 303–321 (2001)
11. Borade, S.N., Deshmukh, R.R.: Effect of distance measures on the performance of face recognition using principal component analysis. In: Berretti, S. (ed.) *Intelligent Systems Technologies and Applications, AISC*, vol. 384, pp. 569–577. Springer, Heidelberg (2016)
12. Biometrics Testing and Statistics. www.biometrics.gov/documnets/biotestingandstats.pdf
13. ORL Face Database. <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>