

Generalized Ring Averaging: A New Method for Left and Right Directional Illumination Invariant Face Recognition for Frontal Poses and their Small Variants.

Mrs. Madhuri A. Potey
Head Dept. of computer engg.
D.Y.Patil College of Engg, Akurdi, Pune.

Mrs. Nilam Vaibhav Upasani
Research Student,
D.Y.Patil College of Engg, Akurdi, Pune

ABSTRACT

In this paper a new method is proposed for face recognition when subject is illuminated from left and right direction in a fixed pose named “generalized ring averaging”. We have proposed new features called “generalized ring averaging” features, which is extension to ring features. Ring features are invariant to rotation and used for binary images only. Proposed features are invariant to direction of illumination (left and right) and used for gray scale images. Well known Fuzzy min-max neural network classifier is used for recognition purpose. The proposed method is found better than one of the most popular method used for face recognition called “eigenfaces”, in terms of percentage recognition rate, when compared with same dimensionality of feature vector. The proposed method requires less time to extract features than eigenfaces and recall time per pattern is found comparable to eigenfaces. However, the proposed method is suitable only for frontal poses and its little variants, which are very close to frontal pose and only for left and right direction of illumination by keeping pose and illumination strength around constant.

Index terms

Face recognition, Computer vision, Machine learning, Eigen values, Eigen vectors, Neural nets, Illumination invariants.

1. INTRODUCTION

In these days of theoretical definitions of “time machine”, which is capable of transporting a human in past and future (at least in theory), but some problems of practical importance like robust face recognition are still found to be difficult and are big challenges to the IT industry [1]. Recognizing face under various conditions is not a simple problem, since visual images undergo lot of changes with different illumination conditions, poses, facial expressions and time [2]. We have recognized two challenges in face recognition, which are illumination and pose variations [2].

Many researchers worked on face recognition under various illumination conditions using various approaches and various imaging modes like visual imagery, thermal imagery and range imagery. Y. Adini, Y. Moses and S Ullman [4] experimentally proved that difference in appearance induced by illuminations is larger than differences between individuals.

Thus the problem of illumination invariant face recognition is not simple. H. Moon, P.J. Phillips [5] have proposed a generic modular PCA based system (also called eigenfaces) for face recognition in which they have used three variations of illumination normalization

method based on histogram equalization by adjusting mean and standard deviation of non-masked facial pixels, but the method is not much robust. Thus, some researches started working with thermal imaging to recognize faces under different illuminations. Diago A. Socolinky et.al. [6] worked on thermal long wave infrared images at equinox sensors. They used two algorithms viz. ARENA [7] and eigenfaces to recognize faces under different illuminations and found that thermal images are better than visual to recognize faces under different illuminations, since in the thermal images intrinsic property of face which is invariant to illumination is used. But since thermal imaging sensors are not common up to date of this paper we are trying to cope with the problem using visual images.

This paper is a sincere effort to propose and check a method, which is invariant specific to left and right directions of illuminations. In this paper we have concentrated on illumination invariant face recognition with two restricted directions of illumination viz. left and right. We have proposed new features called “generalized ring averaging”, which is extension to ring features defined in [3]. Ring features, which are invariant to rotation and used for Chinese handwritten character recognition in [3] is applicable only to binary images. But it is redefined for any gray scale image for face recognition under change of direction of illumination.

Remaining parts of this paper are organized as follows. Section II describes preprocessing required for face image like how to manually select face silhouette and scaling. Section III describes generalized ring averaging which can be used as invariant features. Preparation of training set used for generalized ring averaging is discussed in section IV. Experimental results of generalized ring averaging are discussed in section V. Experimental results with PCA are described in section VI.

2. FEATURE EXTRACTION PREPROCESSING

In this paper we have used image database called IRIS Thermal/Visible Face Database of OTCBVS Benchmark Dataset Collection [8] containing simultaneously acquired thermal and visible face images under variable illuminations, expressions, and poses. Although this database contains thermal and visual images but we have considered only visual images. There are 23 subjects, 16 poses of each subject in two directions of illumination viz. left and right. Out of that we have considered only three poses only viz.

frontal, previous to frontal and immediate next to frontal after converting them into gray scale and thumbnails to reduce the computations. Fig.1 shows such thumbnails of subject Bernard in three poses and illuminated from left and right, after converting into gray scale.

After this a sub-image is selected manually from these images using Microsoft paint in such a way that it accommodate as much face as possible, without taking care of hairs, ears and beard which are outside the face silhouette.

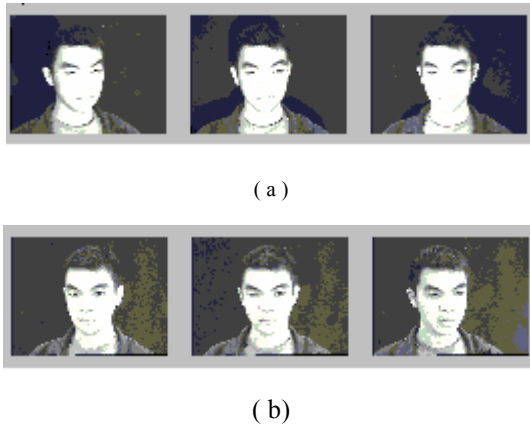


Fig.1. Thumbnail images of subject a) Illuminated from left b) Illuminated from right



Fig. 2. Face silhouette selected from Fig1. (a) and (b)

After this step images will look like as shown in Fig.2. Note that images in Fig 2 may have different sizes but care is taken to have around similar size, since this task is done manually. These images are now scaled to 24 X 24 by using (1).

$$f'(x', y') = f(x/\alpha, y/\beta) \tag{1}$$

where α and β are the scaling factors and calculated as $\alpha=x/24$ and $\beta=y/24$, and $f(x, y)$ is original image before scaling and $f'(x', y')$ is scaled image. This scaled image is used to extract features proposed in this paper called 'generalized ring features'.

3. GENERALIZED RING AVERAGING FEATURES

Consider a face silhouette of $N \times N$ size gray scale pattern, which is defined as $f(x, y)$, where

$$\begin{aligned} x &= -(N-1)/2, \dots, 0, \dots, (N-1)/2 \\ y &= -(N-1)/2, \dots, 0, \dots, (N-1)/2 \end{aligned} \tag{2}$$

then generalized ring features (d) is defined as average value of pixels whose distances (or radii) to the center of image frame are truncated (or rounded) to an integer d . Means face pattern is divided into number of rings, which has some width. Width of a ring can be integer, and face pattern can be divided into at most L rings with width w as given in (3) and (4).

$$\text{generalized-ring-data}(l) = \sum_{(x,y) \in S_l} f(x,y) / P_l \tag{3}$$

where P_l is the total number of pixels in the l th ring $l = 1, 2, \dots, L$, with

$$S_l = \left\{ (x,y) \mid \text{int} \left[\frac{\sqrt{x^2 + y^2} - 1}{w} + 1 \right] = l \right\} \tag{4}$$

where l 's are ring numbers, S_l is the set of pixels in the l th ring, and w is the ring width. The largest ring number L is given by (5).

$$L = \text{int} \left[\left(\frac{(N-1)}{2} - 1 \right) / w + 1 \right] \tag{5}$$

A face of gray pattern of size 100X 100, illuminated from left and right and their rings are shown in Fig. 3, with ring width 10 and largest ring number is 5.



Fig. (3a). Face image illuminated from right and scaled to 100X100



Fig. (3b). Ring pixels of Fig. (3a) from ring 1 to ring 5



Fig. (3c). Face image illuminated from left and scaled to 100X100

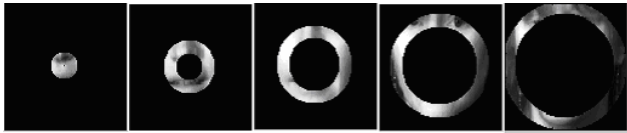


Fig.(3d). Ring pixels of Fig.(3c) from ring 1 to ring 5

If we compare the pixel intensities in ring pixels of Fig. (3b) and Fig. (3d) on a ring-by-ring basis then it is found that in general the count of pixels with same intensities in each ring are approximately same. We have also arrived at the same conclusion by observing other several images. This strengthens basic philosophy proposed in this paper. When an image is illuminated from left then pixels toward left are brighter as compared to pixels towards right hand side of the face and exactly opposite is the case when the image is illuminated from left. Thus average value of pixels in a ring can be used to cope with illumination invariant face recognition when direction of illumination is left and right. Thus to prove “pixels in each ring have around similar average value if illuminated from left and right” following experimentation is done.

In this experimentation we have calculated generalized ring features of third subject of database [8] in visual mode and considered only the frontal pose illuminated from left and right as shown in Fig.4, These features are listed in Table 1 and Table 2, respectively. Preprocessing is done as previously discussed. These features are extracted by setting ring width $w = 2$ with largest ring $L = 6$. For keeping track on change in generalized ring features we are using sample mean μ , standard deviation δ and percentage of spread

of values from the sample mean $\frac{\delta}{\mu} \%$ as discussed in [9].



Fig 4. 24 X 24 images of subject 3 illuminated from left and right

Ring1	Ring 2	Ring 3	Ring 4	Ring 5	Ring 6
241.875	236.863	234.828	243.0	228.4	170.37

Table 1. Generalized ring features when illumination direction is left

Table 2. Generalized ring features when illumination direction is right

Ring1	Ring 2	Ring 3	Ring 4	Ring 5	Ring 6
240.958	251.954	239.881	239.333	230.766	182.425

Mean of generalized ring features of Table1 and Table2 is shown in Table3. Standard deviation in Table4 and percentage of spread of values from respective means is given in Table5.

Table 3. Arithmetic mean of table 1 and table 2.

Ring1	Ring 2	Ring 3	Ring 4	Ring 5	Ring 6
241.416	244.409	237.355	241.166	229.58	176.4

Table 4. Standard deviation of table 1 and table2

Ring1	Ring 2	Ring 3	Ring 4	Ring 5	Ring 6
0.6482	10.6709	3.5728	2.5927	1.6735	8.5206

Table 5. Percentage of spread of values from respective means.

Ring1	Ring 2	Ring 3	Ring 4	Ring 5	Ring 6
0.2685	4.3660	1.5053	1.0751	0.7289	4.8303

By observation of percentage spread of values from respective means $\delta / \mu \%$ as given in Table 5 one can conclude that it is very small in generalized ring features of same subject if the same pose is illuminated from left and right. To minimize percentage of spread of values from respective mean again we are taking average of each component of generalized ring feature of same subject in same pose but illuminated from right and left direction. Calling this method as “generalized ring averaging”. This method is used to recognize faces illuminated from left and right. This method works well only for frontal poses and poses which are very much close to frontal pose.

4. PREPARATION OF TRAINING SET

Consider that we have 23 subjects SK for $K = 1, \dots, 23$. We have considered three poses of each subject and called them pose1, pose2 and pose3. Let the variable P denote a pose then it can assume a value from the set {1, 2, 3}. These poses are small variations to frontal poses. Again the same poses are illuminated from left and right directions. Let variable I denote illumination direction then it can assumes value from the set {1, 2}. Using the notations of K, P

and I any image can be identified uniquely as SKPI, where K = 1, 2, ..., 23, P = 1, 2, 3 and I = 1, 2.

Thus the image database consists of max(K) * max(P) * max(I) = 23*3* 2 =138 images, where max(J) is the maximum value that can be taken by variable J. If we separate this database in two sets based on direction of illumination then each set will contain 69 images with three poses and with the same direction of illumination either right or left. With the above notations we are now ready to define proposed ring averaging method in a formal manner.

The training feature matrix M is calculated as shown in (6).

$$M_{(23*3), L} = \bigcup_{K=1}^{23} \bigcup_{P=1}^3 \frac{(V_L(S_{KP1}) + V_R(S_{KP2}))}{2} \tag{6}$$

where VL (SKP1) is the generalized ring feature vector of K th subject in P th pose and subject is illuminated from left, similarly VR (SKP 2) is the generalized ring feature vector of K th subject in P th pose and subject is illuminated from right. Operator U is union operator to unite features (collection of feature vectors) of all the subjects in different poses. Suffix L in (6) is the largest ring number. In general, if we have α subjects and β poses then the training feature matrix M will be calculated as shown in (7).

$$M_{(K*P), L} = \bigcup_{K=1}^{\alpha} \bigcup_{P=1}^{\beta} \frac{(V_L(S_{KP1}) + V_R(S_{KP2}))}{2} \tag{7}$$

This matrix M will have α *β rows and number of columns will be decided by setting the ring width w and calculating the largest ring number L by using (5).

Testing sets 1 T and 2 T are the generalized ring features and calculated as shown in (8) and (9). Both T1 and T2 are of α *β rows and number of columns will be decided by setting the ring width w and calculating the largest ring number L by using (5). Training and testing sets are normalized in such a way that each component must lie in interval [0,1], that can be done by finding maximum component from feature matrix and then dividing each component by the maximum. It brings feature space from Rn to In, where In is the unit dimensional hyper cube. The normalization loses some absolute information about a pattern but keeps relative information about patterns unchanged [10].

$$T_1 = \bigcup_{K=1}^{\alpha} \bigcup_{P=1}^{\beta} V_L(S_{KP1}) \tag{8}$$

$$T_2 = \bigcup_{K=1}^{\alpha} \bigcup_{P=1}^{\beta} V_L(S_{KP2}) \tag{9}$$

5. EXPERIMENTAL RESULTS OF PROPOSED METHOD

For the purpose of face recognition using above features we have used well known and widely accepted neural classifier “fuzzy min-

max neural network” (FMN) proposed by Patrick Simpson [10]. Training set is used to train the FMN and testing sets T1 and T2 are used for recognition purpose. Experimentation is done by changing the ring width and then training FMN and finding percentage recognition rate for each case along with training time and recall time per pattern. We have also cared about timing analysis of feature extraction method.

FMN training is nothing but adjusting the maximum size of hyperbox θ, in such a way that for minimum number of hyperboxes created all the training patterns must be correctly classified. Then testing sets are applied to find recognition rate.

When FMN is trained with w = 2 (6 features), it has created 63 hyperboxes with θ = 0.022 and required 0.235 seconds to train and given 59.4203 % recognition for testing set T1 and 63.7681% recognition for testing set T2. When FMN is trained with w = 3 (4 features), it has created 57 hyperboxes with θ = 0.018 and taken 0.25 seconds to train and given 43.4783 % recognition for testing set T1 and 55.0725% recognition for testing set T2. When FMN is trained with w = 4 (3 features), it has created 61 hyperboxes with θ = 0.0135 and used 0.265 seconds to train and given 40.5797 % recognition for testing set T1 and 42.02% recognition for testing set T2. These results are shown in Table 6 along with timing analysis of FMN in Table 7.

Table 6. Performance of generalized ring averaging

Ring width w	Dimensio -nality Of feature vector	Maximum Size of Hyperbox	Hyperboxes created	Percentage recognition for illumination direction	
				Left	Right
2	6 66	0.022	63	59.420	63.7681
3	4	0.018	57	43.478	55.0725
4	3	0.0135	61	40.579	42.02

Table 7. Timing analysis of fmn for generalized ring averaging.

Ring Width w	Number of features	Training Time (seconds)	Recall time per pattern (seconds) in various illumination direction		
			Left	Right	Avg
2	6	0.235	0.030	0.027	0.029
3	4	0.25	0.014	0.014	0.014
4	3	0.265	0.015	0.014	0.015

The upper bound on average time required for calculating generalized ring features of a given image is found to be 0.1109 seconds and is shown in Table8.

Table 8. Timing analysis of proposed method

Ring width	Feature extraction time per pattern under illumination from		Average Time (in seconds)
	Left	Right	
2	0.1039	0.1039	0.1039
3	0.12	0.1019	0.1109
4	0.1005	0.1006	0.1006

$$C = \frac{1}{M} \sum_{n=1}^M \Phi_n^T \Phi_n \tag{12}$$

where Φ^T is the transpose of Φ_n and C is the covariance matrix is of size $N \times N$. Then we have collected first 6 principle components u_k , by considering Eigen values and Eigen vectors of C . Then the weights are calculated by transforming an image I as in (13).

$$\omega_k = u_k^T (I - \Psi) \tag{13}$$

where ω_k is a vector of k components and $k=1,2,\dots,6$ in this case. These eigenfaces are shown in Fig.6.



Fig.6. Eigen faces of training set.

6. FEATURE EXTRACTION AND EXPERIMENTATION WITH EIGENFACES.

We want to compare our method with simple and well-known method used for face recognition called “eigenfaces” also called PCA (Principle Component Analysis) discussed in [11] and [12]. Performance comparison of proposed method is done with the PCA by applying same face database, in same size of images, in same sequence of application of patterns, using the same dimensionality of feature vector (i.e. number of features per pattern) and using the same training set.

Since we are using face image database of 23 subjects, with 3 poses closes to frontal and 2 illumination directions viz. left and right.

Each subject is now represented by average of weights of its all images, present in the training set. Thus each subject S_z is now represented as (14).

The training set in this case, is set of images after taking averaging of images of a subject in a fixed pose but illuminated from left and right. Thus there will be 69 images in the training set. Then the average image Ψ of training set of face database considered is given by (10).

$$\Psi = \frac{1}{M} \sum_{n=1}^M I_n \tag{10}$$

where $M = 69$ in this case, I_n is the n th image in the training set of face database. The average image is as shown in Fig.5.

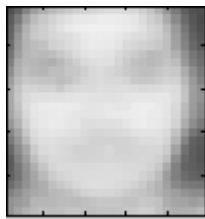


Fig. 5. Average image of the training set.

Each face image differs from average face Ψ by a vector Φ_i as shown in (11).

$$\Phi_i = I_i - \Psi \tag{11}$$

where $i = 1, 2, \dots, M$ and Φ is having size $M \times N^2$, where $N \times N$ is image size. The covariance matrix C is calculated as in (12).

$$S_z = \frac{(\omega_{z1} + \omega_{z2} + \omega_{z3})}{3} \tag{14}$$

where $z = 1, 2, \dots, 23$ and ω_{zi} is the weights of z th subject in i th pose. So we have 23 subject prototypes as knowledge base of the system. Now each image from the database is applied as input to calculate its weights using first six principle components and by using nearest neighbor classifier then 40.5797% face images which are illuminated from left are recognized correctly and 50.7246% of face images which are illuminated from right are recognized correctly. Thus average recognition using six principle components is 45.6522%. We have repeated the same experimentation by selecting first four principle components. It is found that 34.7826% images, which are illuminated from left, are recognized correctly and 40.5797% recognition is achieved for images illuminated from right. Thus by using four features PCA gives average recognition rate of 37.6812%. We have got 24.6377% and 28.9855% recognition for images illuminated from left and right, respectively, when first three principle components are used along with average recognition rate of 26.8116%. All these results are mentioned in Table 9.

Table 9. Performance analysis of pca method.

Dimensionality of feature vector (number of principal components used)	Recall time per pattern Under illuminating from (seconds)		Average Time (seconds)
	Left	Right	
6	0.1513	0.1488	0.15
4	0.1433	0.1481	0.1457
3	0.1492	0.1499	0.1496

By observing average recognition rate of proposed method and PCA by comparing Table 6 and Table 12, it is obvious that proposed method is better than the PCA when compared.

Various calculation tasks involved in PCA are mentioned in Table 10. From the Table 10 it is clear that calculation of average image Ψ , calculation of difference matrix Φ and calculating weights ω_k for all subjects by averaging weights of all of their images in training set are found to be time consuming tasks. Recall time per pattern exhibited by the PCA method is tabulated in Table 11. By comparing average recall time per pattern for a given number of features of proposed method and of PCA as given in Table 7 and Table 11, it is found that proposed method is equiporable to PCA as compared with recall time per pattern.

Table 10. Timing analysis of pca.

Sr. No.	Computational Task	Time required in seconds
1	Calculation of average image	14.329
2	Calculation of difference matrix	14.1410
3	Calculation of covariance matrix when difference matrix is already calculated	0.0470
4	Finding weights (using six principle components)	17.766

Table 11. Recall timing analysis of pca.

Dimensionality Of feature vector (number of principle components used)	Percentage recognition for Illumination direction		
	Left	Right	Average
6	40.5797	50.7246	45.6522
4	34.7826	40.5797	37.6812
3	24.6377	28.9855	26.8116

7. CONCLUSIONS

It is concluded that proposed method of generalized ring averaging is better than PCA when compared on the basis of percentage recognition using same dimensionality of feature vector. It is also noted that as we go on reducing ring width then we will achieve better recognition but relatively more time is needed to extract the features and more recall time per pattern is needed as compared to large value of ring width. The proposed method is suitable only for frontal poses and its little variants, which are very close to frontal pose and only for left and right direction of illumination by keeping pose and illumination strength around constant. By observing feature extraction time of both the method it is concluded that PCA requires more time since there are various time consuming tasks like finding average image, finding weights and etc. But when we compared both the methods using recall time per pattern we can say that they are comparable.

8. REFERENCES

- [1] W. Zhao, Robust image based 3D face recognition, Phd thesis, University of Maryland, 1999.
- [2] W. Zhao, R. Chellappa, A. Rosenfeld, P.J. Phillips, Face Recognition: A Literature Survey, ACM Computing Surveys, 2003, pp. 399-458.
- [3] Hung-Pin Chiu and Din-Chang Tseng, "Invariant handwritten chinese character recognition using fuzzy Min-max neural networks", Pattern recognition letters, Vol.18, 1997,pp.481-491.
- [4] Y. Adini, Y. Moses and S Ullman, "Face recognition: The problem of compensating for changes in illumination direction", IEEE transactions on Pattern Analysis and Machine Intelligence, 19(7):721-732, 1997.
- [5] H. Moon, P.J. Phillips, "Computational and Performance aspects of PCA-based Face Recognition Algorithms", Perception, Vol. 30, 2001, pp. 303-321.
- [6] Diego A. Socolinky, Lawrence B. Wolff, Joshua D. Neuheisel and Christopher K. Eveland, "Illumination invariant face recognition using thermal infrared imagery", available at <http://www.equinoxsensors.com/publications/face.pdf>.
- [7] Terence Sim, Rahul Sukthankar, Matthew D. Mullin and Shumeet Baluja, "High-performance memory based face Recognition for visitor identification", in proceedings of IEEE conf. Face and gesture recognition, Grenoble, 2000.
- [8] IRIS thermal/visual face database available at <http://www.cse.ohio-state.edu/otcbvbs-bench/>
- [9] Khotanzad A. and Y.H.Hong "Invariant image recognition by zernike moments", IEEE Trans. on pattern analysis and machine intelligence, Vol.12 No.5, may 1990.
- [10] Patrick K. Simpson, "Fuzzy Min-Max Neural Networks – Part 1: Classification", IEEE Transactions on Neural Network, vol. 3, pp 776-786, September 1992.
- [11] Matthew A. Turk and Alex P. Pentland, " Face Recognition Using Eigenfaces", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3-6 June 1991, Maui, Hawaii, USA, pp. 586-591.
- [12] M. Turk, A. Pentland, " Eigenfaces for Recognition", Journal of Cognit.