

Multiclass Pattern Recognition of Facial Images using Correlation Filters

Nisha Chandran S¹, Charu Negi², Poonam Verma³

School of Computing
Graphic Era Hill University
Dehradun, India

Abstract—Pattern Recognition comes naturally to humans and there are many pattern recognition tasks which humans can perform admirably well. However, human pattern recognition cannot compete with machine speed when the number of classes to be recognized becomes tremendously large. In this paper, we analyze the effectiveness of correlation filters for pattern classification problems. We have used Distance Classifier Correlation Filter (DCCF) for pattern classification of facial images. Two essential qualities of a correlation filter are distortion tolerance and discrimination ability. DCCF transposes the feature space in such a way that the images belonging to the same class gets closer and the images from different class moves far apart; thereby increasing the distortion tolerance and the discrimination ability. The results obtained demonstrate the effectiveness of the approach for face recognition applications.

Keywords—Pattern recognition; correlation filter; multiclass recognition

I. INTRODUCTION

There are many daily pattern recognition tasks that come naturally to humans. For example, we can recognise a close friend of ours even after a gap of many years though his features have changed a lot. We can understand a familiar voice even if it is slightly distorted. However, human pattern recognition suffers from three main drawbacks: poor speed, difficulty in scaling, and inability to handle some recognition tasks. Not surprisingly, humans cannot match machine pattern recognition tasks where good pattern recognition algorithms exist. Also, human pattern recognition has limitations when the number of classes to be recognized becomes large. Although humans have evolved to perform well on some recognition tasks such as face or voice recognition, except for a few trained experts, most humans cannot tell whose fingerprint they are looking at. Thus, there are many interesting pattern recognition tasks for which we need machines.

The main goal of pattern recognition is to assign an observation maybe a signal, or an image or a high dimensional object into one of the multiple classes. An important class of pattern recognition applications is the use of biometric signatures like face image, fingerprint image, iris image etc. for person identification [1-.3] The use of two-dimensional (2-D) correlation to detect, locate, and classify targets in observed scenes has been a topic of research for a long time [4-10].

In this paper we analyze the possibility of using correlation filters for solving multiclass classification problems. DCCF design uses a global transformation (correlation filter) to transform the feature space to decrease the intra class distance and to increase the inter class distance. Results of the experiments conducted on benchmark dataset demonstrate the robustness of the proposed method. The rest of this paper is organized as follows. We begin with a discussion on some related work on correlation filters in Section 2. In Section 3, we discuss the salient features of DCCF and outlines the strategies adopted to apply DCCF for a multiclass facial recognition problem. Section 4 provides the experimental results and finally Section 5 concludes the paper.

II. RELATED WORK

Correlation filters have been widely used for several pattern recognition tasks and visual tracking of objects [11]. The advantage of using correlation filters for object tracking tasks is that it can track objects that are rotated or occluded or are with several photometric and geometric challenges. Pattern recognition of complex objects which are partially occluded are also done efficiently using multiple correlation filters [12] [13]. Composite correlation filters [14-17] gives superior performance to Matched Filters as they are designed from multiple reference images. If the reference image set are well represented to incorporate all the possible distortions that are likely to be encountered by an object, then the resulting filter will be distortion tolerant. While designing correlation filters three questions need to be considered: (1) How good is the ability of the filter to suppress clutter and noise. (2) How easy is it to detect a correlation peak. (3) How tolerant is the filter to the distortion of the object.

One of the earliest composite correlation filters proposed was the Synthetic Discriminant Function (SDF) [18]. SDF uses a linear combination of reference images to create a composite image. When the designed filter is correlated with a test image, a peak will be present in the correlation plane if the test image corresponds to the TRUE class to be recognized. For all other inputs belonging to FALSE class, there will be no peak in the correlation plane. For digital pattern recognition applications, an SDF synthesized in the computer is correlated with the test image digitally, whereas for optical pattern recognition applications, the SDF designed digitally is converted to a hologram using multiple-exposure holographic techniques.

Initially, design of SDF did not consider any noise and hence the filters were not noise tolerant. Minimum Variance SDF (MVSDF) [19] was one of the earliest attempts to introduce noise analysis in SDF filter design by maximizing the noise tolerance of the SDF. The original SDF design considered only the cross-correlation values at the origin. This could not ensure that the output had its peak at the origin. Since shift in the reference images resulted in shift in the peak location, there was an ambiguity in peak location when the reference image shifts were unknown. This problem was addressed in the Minimum Average Correlation Energy (MACE) filters [20]. MACE filters could produce sharp correlation peaks at the origin and were more likely to produce a correlation peak at the same location as the shifted input.

However, it was realized that by exclusively focussing on the correlation-peak values, one neglects the information in the other regions of the correlation plane. In Minimum-squared-error synthetic discriminant function the averaged squared error between the resulting correlation outputs and the desired one is minimized to obtain a desired correlation plane. Maximum-average-correlation-height filter essentially uses this idea to achieve a correlation shape that yield the smallest squared error. Distance Classifier Correlation Filters (DCCF) [21] essentially incorporates the two ideas. In DCCF rather than considering just the peak the entire correlation plane was considered. Applications of initial DCCF had limitations as the approach was limited to just two classes at a time.

III. MULTICLASS PATTERN RECOGNITION OF FACIAL IMAGES

A facial recognition system is a computer application for automatically identifying or verifying a person from a digital image or a video frame obtained from a video source. One of the ways to do this is by comparing selected facial features from the image and a facial database. DCCFs were first proposed by Mahalanobis *et al.* DCCF uses a correlation filter h , designed from a set of training images, to classify a test image into one of a set of predefined classes. As illustrated in Fig. 1, DCCF design uses a global transformation (correlation filter) h which transforms the space in such a way that same class gets closer and images belonging to different classes move apart. This improves distortion tolerances as well as discrimination ability of the correlation filter.

A. Formulation of DCCF Filter

Let X_{ik} represent the 2D Fourier transform of the i^{th} training image of class k ordered as a vector and X_{ik} a diagonal matrix with X_{ik} as its diagonal elements. Let m_k represents the mean vector of class k and M_k the diagonal matrix with m_k as its diagonal elements. If the transformation h has to make the inter-class distance large, then the distance between the mean correlation peak values between the different classes is made as large as possible. This is formulated as the measure $h^+ M h$ where M is given as in equation (1)[16].

$$M = \frac{1}{c} \sum_{k=1}^c (m - m_k) (m - m_k)^+ \quad (1)$$

In the above equation m_k represents the mean vector of class 'k' and m is the mean of all the classes given as in equation (2).

$$m = \frac{1}{c} \sum_{k=1}^c m_k \quad (2)$$

Simultaneously, the transformation h makes each class compact. The compactness of a class after applying h is measured by the metric average similarity measure [9] which is a metric measure of the similarity between the training images and the mean value of the class given by $h^+ S h$ where S is intra-class scatter matrix given as in equation (3).

$$S = \frac{1}{c} \sum_{k=1}^c \frac{1}{N} \sum_{i=1}^N h^+ |X_{ik} - M_k| |X_{ik} - M_k|^* h \quad (3)$$

It follows that the correlation filter h should be so designed that it maximizes the metric.

$$J(h) = \frac{h^+ M h}{h^+ S h} \quad (4)$$

The $J(h)$ is maximized when h is the dominant eigenvector of $S^{-1}M$.

Once the correlation filter h is designed using the training image set, a test image is classified by correlating with h . The distance metric is calculated as the difference between two correlation peaks. The first correlation peak corresponds to the correlation of the filter with the test image. As shown in Fig. 1, this corresponds to the transformation of the test image. The second correlation peak corresponds to the correlation of the filter with the mean vector of class 'k' which corresponds to the transformation of the class 'k'. The test image is assigned to the class which gives the minimum distance.

B. Classification using the DCCF Filter: Method 1

Let z be the test image. The correlation peak that corresponds to the correlation between the test image and the correlation filter is given by $H^* z$. The correlation peak that corresponds to the correlation between the mean vector of class 'k' and the correlation filter is given by $H^* m_k$. The distance metric d_k that gives the distance between the transformed test image and transformed mean vector of class 'k' is given by.

$$d_k = |H^* z - H^* m_k|^2 \quad (5)$$

The given test image is assigned to the class that gives the minimum value of d_k

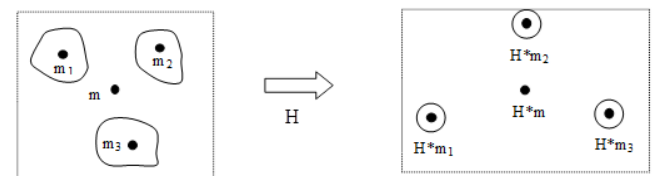


Fig. 1. H Transforms the Feature Space in Such a Way that the Interclass Distance Increases Whereas the Intra class Distance Decreases. Here m_1, m_2, m_3 Represent the mean Value of the Fourier Transform of the Images within each Class and m that of All the Images Combined. $H^+ m, H^+ m_1, H^+ m_2,$ and $H^+ m_3$ Represent the Correlation Peak Values of the Respective Classes.

C. Classification using the DCCF Filter: Method 2

The above equation for d_k can also be written as follows

$$d_k = a + b_k - z^+ h_k - h_k^+ z \quad 1 \leq k \leq C \quad (6)$$

In the above equation, a is the transformed input image and is given as $a = |H^*Z|^2$ and $b_k = |H^*m_k|^2$ is the energy of the transformed k^{th} class mean and $h_k = HH^*m_k$ is considered as the effective filter for class k . This gives us an alternate strategy for classification of a test image using 2D correlation techniques. The third and the fourth term in the expression (6) is implemented as a 2D correlation of the test image 'z' and the effective filter h_k for the class 'k'. The term d_k is minimized when the third and fourth term is maximized which corresponds to the peak correlation value of z with h_k . Shown in Fig. 2 is the schematic for implementing this method using 2D correlation. The test image is correlated with effective filter of each class and the peak value which is the value at the origin of the correlation plane is used to compute the distance d_k using expression (6). As discussed in the previous section, coherent optical processing systems can be used to implement these techniques.

D. Facial Recognition using DCCF

In this section we discuss the classification of 'c' different objects, in this case facial images, using the DCCF. Facial images of one person with different facial expressions form the members of one class. The training images are used to design the filter, which is used to classify a test image. The algorithms for the classification of the facial images using DCCF are given in Table I.

TABLE I. ALGORITHMS FOR MULTICLASS PATTERN CLASSIFICATION USING DCCF

Algorithm to calculate DCCF filter	
Step 1:	Formulate the vectors X_{jk} , m_k and m
Step 2:	Calculate M using Eqn (1)
Step 3:	Calculate Average Similarity Measure S using Eqn (3)
Step 4:	Calculate h as the dominant Eigenvector of $S^{-1}M$
Algorithm to calculate the dominant eigenvector of $S^{-1}M$	
Step 1:	Formulate the matrix E as $E = e_1 e_2 e_3 \dots e_C $
Step 2:	Formulate V as E^+E
Step 3:	Find the eigenvectors and Eigen values of V.
Step 4:	Formulate the matrix P whose columns consist of the normalized eigenvectors of V.
Step 5:	Formulate the square matrix Δ whose diagonal consist of non-negative eigenvectors of V
Step 6:	Compute $\Phi = E P \Delta^{-1/2}$
Step 7:	Compute $\Delta \Phi^+ S^{-1} \Phi$
Step 8:	Calculate the dominant eigenvector a_{\max} of $\Delta \Phi^+ S^{-1} \Phi$
Step 9:	Compute the dominant eigenvector of $S^{-1}M$ as $S^{-1} \Phi a_{\max}$
Algorithm for classification of the test image (Method 1)	
Step 1:	Formulate the square matrix H whose diagonal consist of the vector h
Step 2:	Calculate the distance metric of the test image vector z to the class k mean vector m_k as $d_k = H^+z - H^+m_k ^2$ for $k=1$ to C
Step 3:	Classify the vector z to that class that gives the minimum value of distance metric.
Algorithm for classification of the test image (Method 2)	
Step 1:	Calculate h_k and b_k for each $k=1$ to C as given by expression (6)
Step 2:	Perform the 2D correlation of test image 'z' with the effective filter h_k for each class.
Step 3:	Compute the distance metric d_k for each class 'k' using expression (6).
Step 4:	Classify the vector z to that class that gives the minimum value of distance metric.

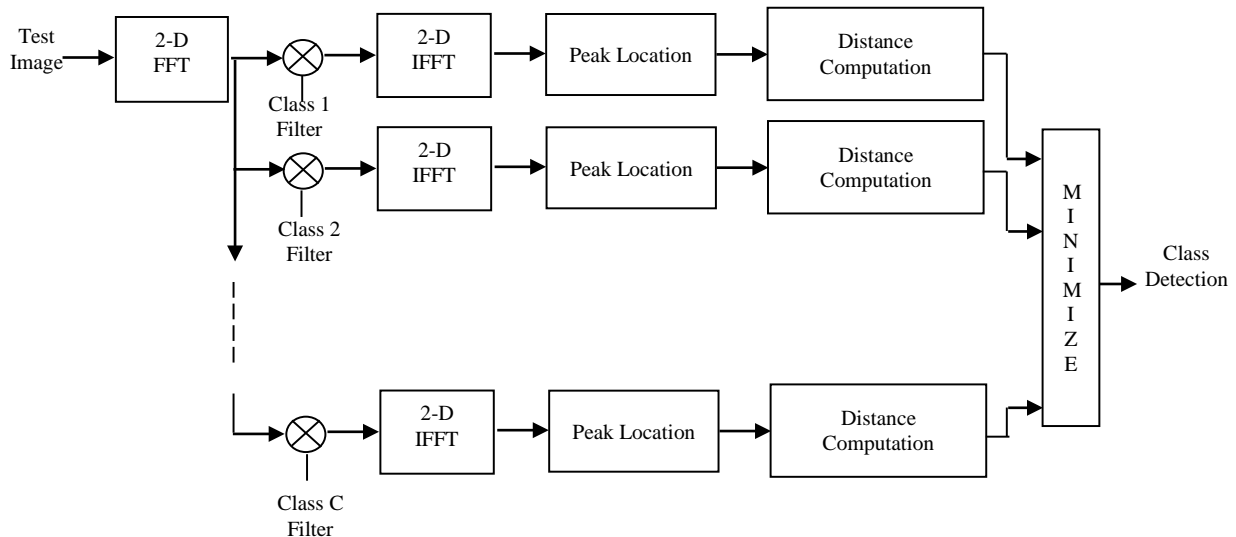


Fig. 2. Schematic Diagram to Illustrate the Design and the use of DCCF's.

IV. EXPERIMENTAL EVALUATION

Experiments are carried out on Essex face database [22] which consists of 3040 facial images of 152 persons. A total of 20 images are available for training from each class. The subjects are set at fixed distance from the camera and were asked to speak, whilst a sequence of 20 images was taken. The speech was used to introduce facial expression variation. The resolution of each image was 180 x 200 pixels. The algorithm discussed in Section 2 was used to design a DCCF that could classify any image in the database to one of the 152 classes. The 20 images in each class were divided into two sets. One set of images were used to train the filter referred to as training images. The other set of images were used to test the filter referred to as test images. Representative images from the database is shown in Fig. 3.

The robustness of the designed filter would depend on the size of the training image set and how well the training image set represents the possible distortions that occur. Obviously, the larger the training image set, the robust the filter will be. But usually there are a limited number of images available. Hence, if most of the images are used to train the filter, we are left with few images to test the filter. If the number of images used to train the filter is less, then the classification error increases. Hence the training and test image set must be chosen judiciously.

Table II gives the results obtained from the study varying the size of the training images set and test image dataset. In each case the total number of images used (i.e.training+testing) is kept a constant. The filter is generated using the training images. The other images are used for testing. The distance of the transformed test image from the transformed mean image of each class is calculated as described in the previous chapter. The test image is classified to the class which gives the minimum value of distance metric.

The total classification error for all classes as a percentage of total facial images are plotted against total number of training images as a percentage of total facial images and is shown in Fig 4. It is observed that when more than 40% images are used for training the classification error is significantly low. Fig. 5 shows the distribution of errors when 95% of the images are used for training. There are only 5 errors out of 152 images used for testing (one image from each of 152 classes) and these 5 errors fall in 5 different classes. In other words, the errors are distributed evenly. Fig. 6 shows the distribution of errors when 90% images of the image dataset is used for training. The 16 errors are distributed over 10 classes with the maximum error per class being 2. Fig. 7 shows the distribution of errors when 85% of the images are used for training. The 25 errors are distributed over 13 classes with the maximum error per class being 3. Fig. 8 shows the distribution of errors 38 errors over 14 classes when 80% of the images are used for training with maximum error per class being 4 and minimum error per class being 1. Fig. 9-15 shows the distribution of errors when the training images decreases from 75% to 45%. The classification errors correspondingly increase from 50 to 175. For the case when training images are 75% of the image dataset, the errors fall in

15 classes with maximum error per class being 5 and minimum error per class being 1. When the training images are 45%, the errors fall into 40 classes with maximum error per class being 11 and minimum error per class being 1. It is observed that for all the cases the error is uniformly distributed across all classes. Fig. 16-23 shows the distribution of errors when the training images decreases from 40% to 5%. The total errors increase from 219 to 2868. These errors fall in 46 classes and 152 classes, respectively. In the worst case when the training images are just 5%, the errors are evenly distributed in all the 152 classes with the average error per class being 19. It may be seen that as the number of training images decreases, the error increases, as expected, and these errors are evenly distributed across multiple classes.



Fig. 3. Representative Images from the Face Database.

TABLE II. CLASSIFICATION ERROR OBTAINED BY VARYING THE NUMBER OF TRAINING IMAGES

Sr. No.	Training Images per class	Total Training images	Test Images Per class	Total Test images	% Training image	Class Error	% Error
1	19	2888	1	152	95	5	3.3
2	18	2736	2	304	90	16	5.3
3	17	2584	3	456	85	25	5.5
4	16	2432	4	608	80	38	6.3
5	15	2280	5	760	75	50	6.6
6	14	2128	6	912	70	67	7.3
7	13	1976	7	1064	65	79	7.4
8	12	1824	8	1216	60	102	8.4
9	11	1672	9	1368	55	129	9.4
10	10	1520	10	1520	50	158	10.4
11	9	1368	11	1672	45	175	10.5
12	8	1216	12	1824	40	219	12
13	7	1064	13	1976	35	305	15.4
14	6	912	14	2128	30	400	18.8
15	5	760	15	2280	25	539	23.6
16	4	608	16	2432	20	751	30.9
17	3	456	17	2584	15	1096	42.4
18	2	304	18	2736	10	1892	69.1
19	1	152	19	2888	5	2868	99.3

The experimental results presented shows that DCCF successfully classifies facial images of 152 persons from the Essex database. The facial images belong to persons from quite different ethnic backgrounds. There are 20 images of each person with different facial expressions. It is seen that the DCCF filter shows good discrimination ability to classify facial expressions of 152 persons while maintaining good distortion tolerance towards a variety of facial expressions present in each class. It is also seen that the classification error is below 10% when at least 40% images from the available dataset are used to train the filter.

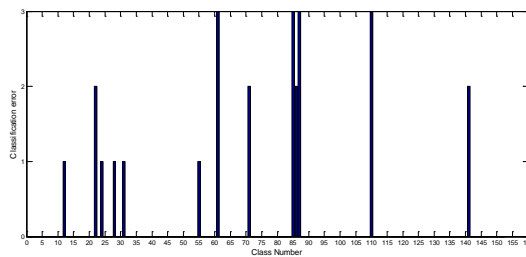


Fig. 7. 85% Training Images.

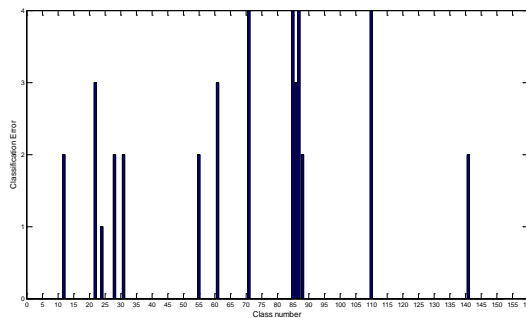


Fig. 8. 80% Training Images.

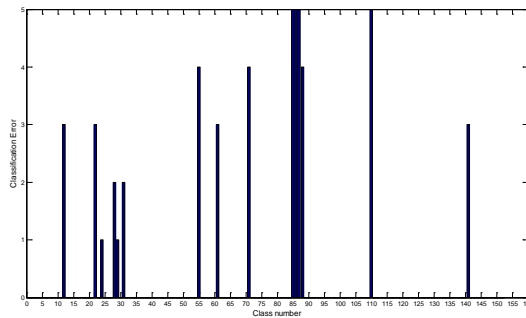


Fig. 9. 75% Training Images.

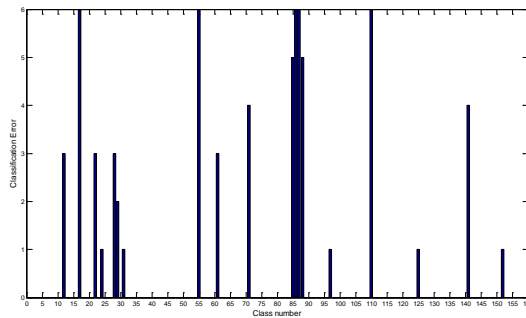


Fig. 10. 70% Training Images.

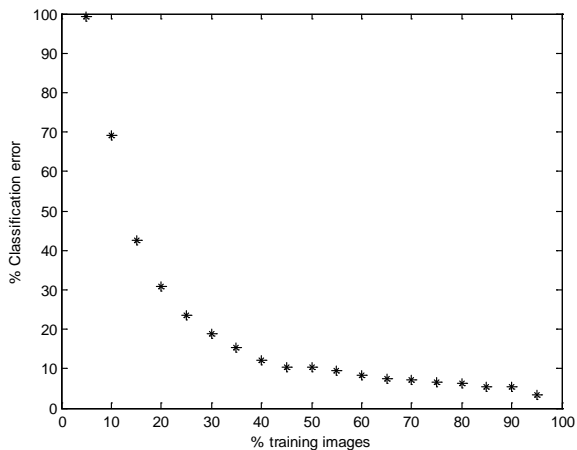


Fig. 4. Graph Showing the Decrease in Classification Error as the Number of Training Images Increases.

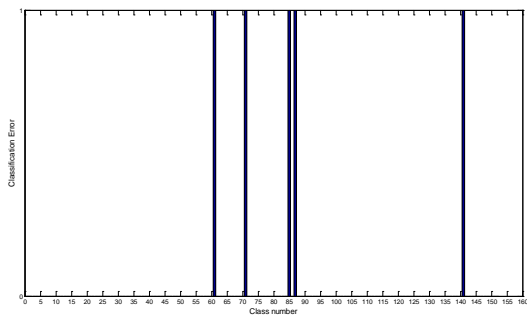


Fig. 5. 95% Training Images.

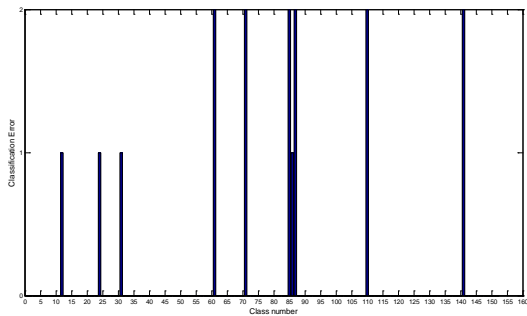


Fig. 6. 90% Training Images.

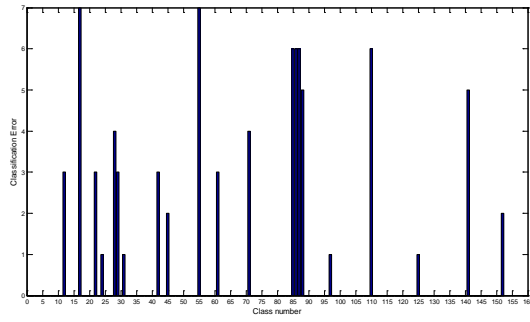


Fig. 11. 65% Training Images.

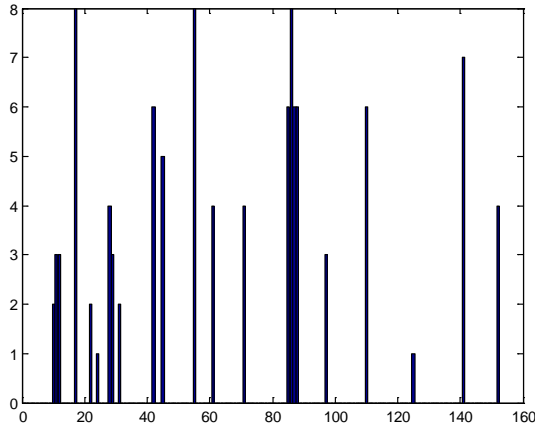


Fig. 12. 60% Training Images.

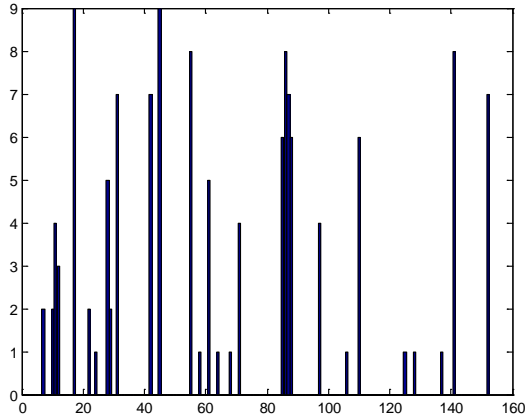


Fig. 13. 55% Training Images.

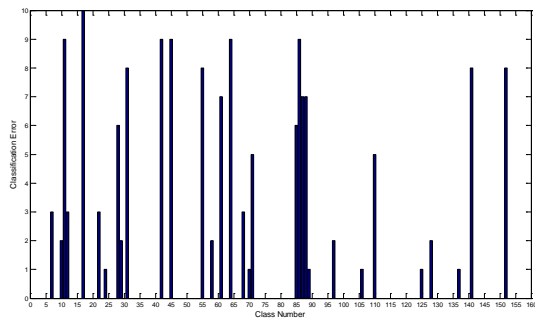


Fig. 14. 50% Training Images.

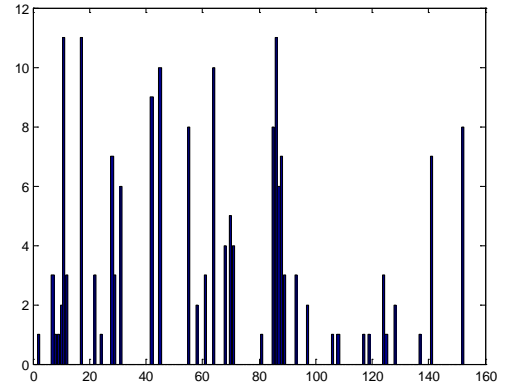


Fig. 15. 45% Training Images.

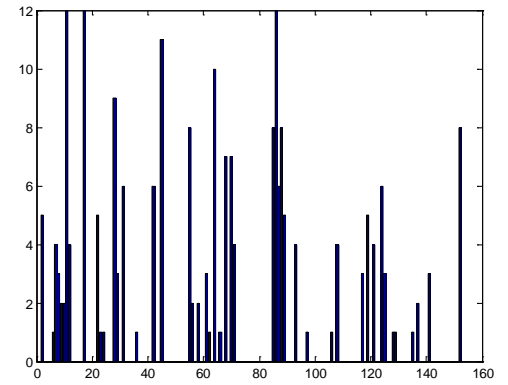


Fig. 16. 40% Training Images.

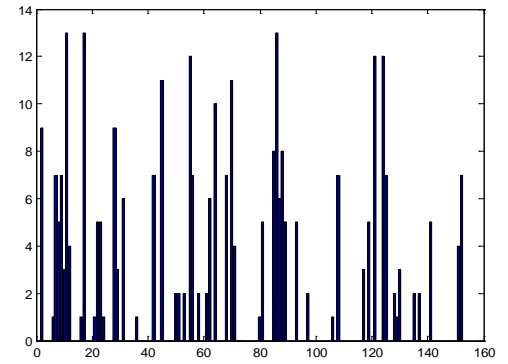


Fig. 17. 35% Training Images.

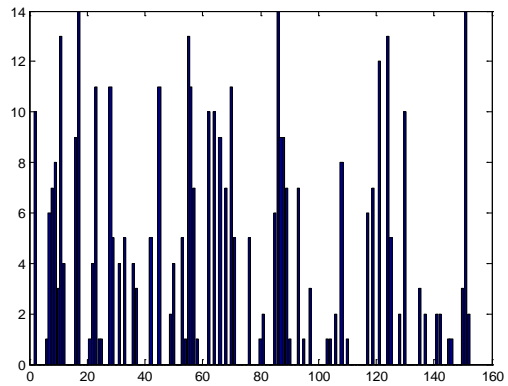


Fig. 18. 30% Training Images.

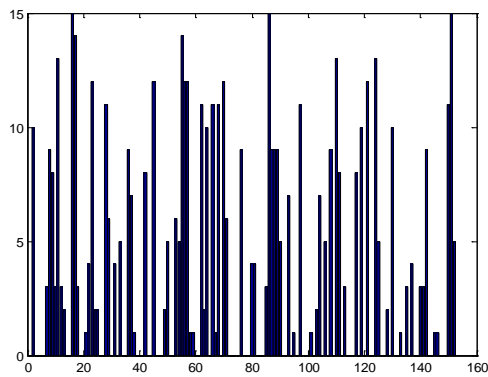


Fig. 19. 25% Training Images.

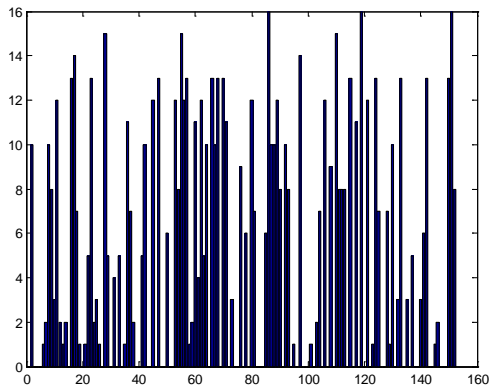


Fig. 20. 20% Training Images.

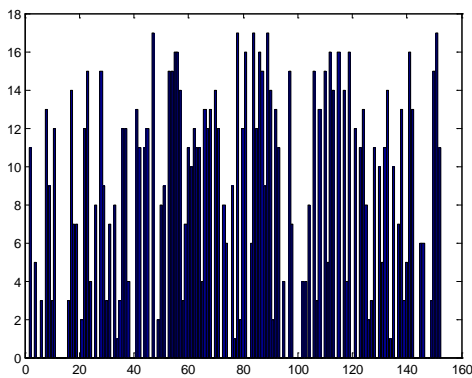


Fig. 21. 15% Training Images.

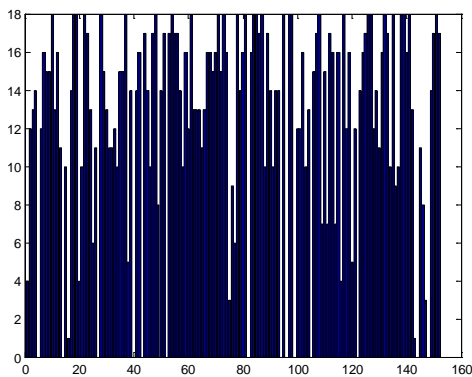


Fig. 22. 10% Training Images.

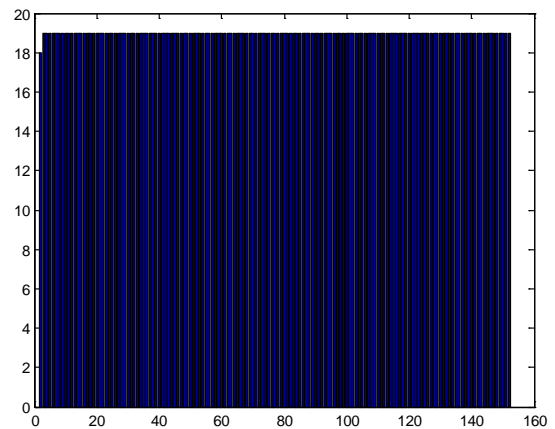


Fig. 23. 5% Training Images.

V. CONCLUSION

In this paper correlation based pattern recognition is adopted for classification of facial images. A good correlation filter must have distortion tolerance and discrimination ability in equal measure. Distance Classifier Correlation Filter is designed, with a global transformation H which maximises the separation between different classes and which minimizes the spread of each class is found. Experimental results show DCCFs work perfectly well for face recognition applications. However, the training set used to develop the filter should be large and well represented. Percentage of training images must be at least 40% of the available dataset for errors to be reasonably low. We find that as the images used to train the filter decreases the increase in classification error is distributed evenly across multiple classes. Our future work would involve the study of DCCF based classification technique for various other biometrics like fingerprint, iris, etc.

REFERENCES

- [1] R.O. Duda, P.E. Hart and D.G. Stork, Pattern Classification, 2nd edn. New York, John Wiley, 2001.
- [2] K. Fukunaga, Introduction to Statistical Pattern Recognition, 2nd edn., New York, Academic Press, 1990.
- [3] G. F. Schils and D. W. Sweeney, "Optical processor for recognition of three-dimensional targets viewed from any direction," J. Opt. Soc. Am. A 5, 1309-1321 (1988).
- [4] M. Turk and A. Pentland, "Eigen faces for recognition," Journal of Cognitive Neuroscience, 3, 1991, 71-86.
- [5] B.V.K. Vijaya Kumar, "Tutorial survey of composite filter designs for Optical correlators," Applied Optics, 31, 1992, 4773-4801.
- [6] B. V. K. Vijayakumar, A. Mahanalobis, and R. D. Juday, Correlation Pattern Recognition, Cambridge University Press, Cambridge, 2005.
- [7] B. Braunecker, R. W. Hauck, and A. W. Lohmann, "Optical character recognition based on no redundant correlation measurements," Appl. Opt. 18, 1979, 2746-2753.
- [8] C. Hester and D. Casasent, "Multivariate technique for multiclass Pattern recognition," Applied Optics, 19, 1980, 1758-1761.
- [9] R. R. Kallman, "Direct construction of phase-only filters," Appl. Opt. 26, 1987, 5200-5201.
- [10] S.M. Kay, Fundamentals of Signal Processing: Estimation Theory Upper Saddle River, NJ, Prentice Hall, 1993.
- [11] D.S. Bolme, J.R. Beveridge, B.A. Draper, and Y.M. Lui, "Visual object tracking using adaptive correlation filters", in proc. IEEE computer society conference on computer vision and pattern recognition, 13-18 June 2010 San Francisco, CA, USA, pp. 2544-2550.

- [12] T. Liu, G. Wang and Q. Yang, "Real-time part-based visual tracking via adaptive correlation filters", in proc. IEEE Conference on Computer Vision and Pattern Recognition, 7-12 June 2015, Boston. pp. 4902-4912.
- [13] H. Kiani Galoogahi, A. Fagg, and S. Lucey, "Learning background-aware correlation filters for visual tracking", in Proc. IEEE international conference on computer vision, 22 -29 Oct 2017 ,Venice, Italy, pp. 1135-1143.
- [14] H.L. Van Trees, Detection, Estimation and Modulation Theory: Part I New York, John Wiley, 1968.
- [15] C.R. Rao, "Information and accuracy attainable in the estimation of Statistical Parameters," Bulletin of the Calcutta mathematical Society, 37, 1945, 81-91.
- [16] C.R. Rao, Linear Statistical Inference and its Applications, New York, John Wiley, 1973.
- [17] H. J. Caulfield and W. T. Maloney, "Improved discrimination in optical character recognition", Appl. Opt., 8, 1969, 2354-2356.
- [18] Z. Bahri and B. V. K. ViJaya Kumar, "Generalized synthetic discriminant functions," J. Opt. Soc. Am. A 5, 1988, 562-571.
- [19] B.V.K. Vijaya Kumar, "Minimum variance synthetic discriminant Functions," Journal of the Optical Society of America A., 3, 1986, 1579-84.
- [20] A. Mahalanobis, B.V.K. Vijaya Kumar, D. Casasent," Minimum Average Correlation Energy filters," Applied Optics, 26, 1987, 3633-3640.
- [21] A. Mahalanobis, B.V.K.Vijaya Kumar, and S.R.F. Sims, "Distance Classifier correlation filters for multi-class automatic target recognition," Applied Optics, 35, 1996, 3127- 3133.
- [22] Face Database Available [Online] <http://cswww.essex.ac.uk/mv/allfaces/faces94.html>.