Face recognition using fractional coefficients and discrete cosine transform tool

Mourad Moussa, Maha Hamila, Ali Douik
Networked Objects, Control and Communication Systems (NOCCS) Laboratory, Department of Applied informatics,
National Engineering School of Sousse, Sousse University, Tunisia

Article Info

ABSTRACT

Face recognition is a computer vision application based on biometric information for automatic person identification or verification from image sequence or a video frame. In this context DCT is the easy technique to determine significant parameters. Until now the main object is selection of the coefficients to obtain the best recognition. Many techniques rely on premasking windows to discard the high and low coefficients to enhance performance. However, the problem resides in the shape and size of premask. To improve discriminator ability in discrete cosine transform domain, we used fractional coefficients of the transformed images with discrete cosine transform to limit the coefficients area for a better performance system. Then from the selected bands, we use the discrimination power analysis to search for the coefficients having the highest power to discriminate different classes from each other. Feature selection algorithm is a key issue in all pattern recognition system, in fact this algorithm is utilized to define features vector among several ones, where these features are selected according a specified discrimination criterion. Many classifiers are used to evaluate our approach like, support vector machine and random forests. The proposed approach is validated with Yale and ORL Face databases. Experimental results prove the sufficiency of this method in face and facial expression recognition field.

Corresponding Author:

Mourad Moussa,
NOCCS Laboratory,
Department of Applied Informatics,
National Engineering School of Sousse, Tunisia.
Email: Mourad.enim@yahoo.fr

1. INTRODUCTION

In the last decades, face and facial expressions recognition has received crucial attention from the worldwide research community, and it is a popular scope in data and object recognition and computer application according to its enormous and motivated range of security and criminal applications such as forensic face recognition [1-3], biometric authentication [4, 5], video surveillance [6], information security [7] and edge detection [8]. Face recognition system is mostly used to perform verification of human identity, this way is based on feature extraction and dimensionality reduction, and a number of facial recognition systems have been produced with distinct measure of success. Although, there are various face recognition algorithms, working well in different environments, facial recognition is still a very challenging problem in real applications and, up to now, there is no technique that offers a robust solution to different situations and applications that face recognition may encounter. Therefore, the number of techniques is large and diverse. Face and object recognition problem can be divided into two types: (a) face verification, and (b) face identification. First, a face verification technique is used to involve whether an acquired face image matches

with another one preexisting in a database. Then, the second group tries to verify a human face from a given sample of that face. Among face recognition algorithms, DCT has been employed more recently in face recognition for image compression and dimension reduction [9].

The main contribution of our algorithm is reducing size and cost, while increasing speed. The rest of the paper has been arranged as follows. Section 3 presents the feature extraction and selection techniques. Section 4 explains the assessment of the performance of the proposed approach by using different classifiers. Section 5 details the proposed approach. Experimental results and discussion are found in Section 6. Finally, the conclusions inferred from the results are presented in Section 7.

2. PREVIOUS WORKS

The Due to its promising properties, DCT is an efficient way for parameters extraction and selection, applied in a great deal of occasions in recognition systems such as face recognition [10], and palm print recognition. This approach is also operated in many fields like image coding and compression [11]. After applying the discrete cosine transform, the main necessary information of frame work is focused in the weak frequencies, which are considered as features in the DCT algorithm when it will be used for dimension reduction [12, 13]. Ramasubramanian et al., [14] suggests a face recognition method using DCT combined with LDA. Dabbaghchian et al., [15] and Hongtao et al., [16] respectively chose separability measure to select the DCT coefficients. Yet, the discriminant ability of the feature vector combined with the selected DCT coefficients is not strong necessarily. Dabbaghchian et al., were used the premask to discard the high and low frequencies and reduce the range of the discrimination coefficients. Whereas, Eesa et al., [17] have shown improvement in performance with the use of fractional coefficients to complete conversion of the images at reduced computations resulting in faster image identification, other author present an incorporated method for feature selection.

For the recent decades, random forest is widely used in computer vision application, in fact it is a competitive classification algorithm and received growing interests [18, 19]. In Salhi et al., [20] were dealing with detection and recognition of face and facial expression using the random forest. Cortes et al., [21] presented the foundations for support vector machine (SVM), this technique has become a well-known prominent method to solve pattern classification and regression problems [22]. To deal with the face classification problem, many researchers [23] and Wang et al., [24] have applied SVM in their studies and showed positive experiment results. In S.-K., et al., [25] were performed SVM feature extractions for face recognition.

3. FEATURE EXTRACTION AND SELECTION
3.1. Feature extraction using DCT

Feature extraction, which leads to reduce the feature space dimensionality, is the main objective for many authors, indeed data become enough improved and comprehensible. It has been a principal topic of research and development since a few decades. DCT is a mathematic tool widely used on signal and image processing, it is a real number domain transformation that transforms the data from temporal domain to frequency domain. After this transformation, the distribution of the coefficients becomes more concentrated, such that the main data of the frame work focused on the weak amplitude frequencies. Consequently, DCT is an algorithm that contributes to improve results in image and voice compression field. Also it can be used for feature for classification. DCT represents a signal analysis tool and a robust face recognition system in front variations facial geometry and illumination resource. DCT is conceptually similar to discrete Fourier transform (DFT), since it transforms a signal or an image from the spatial domain to the frequency domain. Considering the input sequence $f(i)$ the general equation for 1D (N data items) DCT defined as follows:

$$F(u) = \sqrt{\frac{2}{N}} \sum_{i=0}^{N-1} A(i)^* \cos\left(\frac{u(2i+1)\pi}{2N}\right) * f(i), Where, A(i) = \begin{cases} \frac{1}{\sqrt{2}} \ for \ u = 0 \\ 1 \ otherwise \end{cases} \tag{1}$$

The basic encoding operation of the DCT is as follows:
- The size of the input image is N x M.
- $f(i,j)$ represents the pixel intensity for $x(i,j)$.
- $F(u,v)$ is the corresponding coefficient for pixel $x(i,j)$.
- Much of the energy of the images signal lies at low frequencies, which appear in the upper left DCT corner.
- The lower right DCT coefficients represent higher frequencies, and are often sufficiently small to be discarded with little visible distortion. The DCT coefficients are calculated using (2), and the matrix representing DCT coefficients can be divided into three matrix namely low, middle and high frequencies.

The general equation for 2D (N*M image) DCT is defined as follows:

$$F(u,v) = \sqrt{\frac{2}{N}} \sqrt{\frac{2}{M}} \sum_{i=0}^{N-1} A(i) * \cos\left(\frac{u(2i+1)\pi}{2N}\right) * \sum_{j=0}^{M-1} A(j) * \cos\left(\frac{v(2j+1)\pi}{2M}\right) * f(i,j) \qquad (2)$$

3.2. Fractional coefficients

In this proposed method, DCT is applied on all the images to obtain transformed image content. Then different fractional coefficients of these transformed images are considered to construct the feature vectors. For the 50% reduced fractional coefficients, upper triangular part of the transformed content is retained while for 25% reduced fractional coefficients, only the first quadrant of the transformed image is booked. The process is repeated until reduced fractional coefficients of 1.5625.

The number of DCT coefficients used to construct the feature vectors and the computational complexity are reduced considerably with each different level of fractional coefficients like shown in Table 1, in the case of Yale database. Figure 1 explains the DCT coefficients extraction used for facial recognition using fractional coefficients of transformed images.
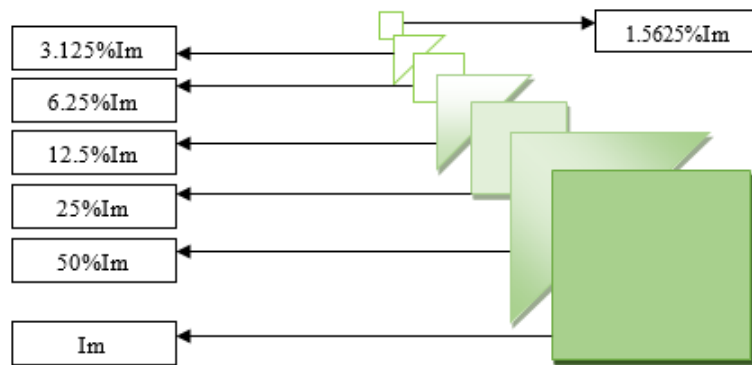

Figure 1. Fractional coefficients of the transformed image 'Im'

Table 1. DCT coefficients number used for different fractional coefficients

Fractional coefficient	Number of used DCT coefficients	Fractional coefficient	Number of used DCT coefficients
100%	50176	12.5%	6272
50%	25088	6.25%	3136
25%	12544	3.125%	1568

3.2. Featureselection using DPA

The principal objective of feature selection is to specify and to eliminate irrelevant and redundant features if it is possible with respect to the task to be executed. It has the potential to be a fully automatic process, and brings some benefits for data mining, such as: an improved predictive accuracy, more compact and easily understood learned knowledge and reduced execution time for algorithms. In our analysed image, the determination of DCT coefficients follows DP values. All kept parameters number after selection DCT coefficients is denoted by n. Among the selected DCT coefficients only the n DCT coefficients having the larger DP values are choosing. A coefficient DP depends on the inter class variation and the intra class variation. The large DP value is related to large variation between the classes but it is small in the other case (i.e. variation intra-classes). The database contains C clusters and S training images set per class, then totally $C{\times}S$ training images are employed. Estimation of DP values follows many steps described as follows:

a)   Construction of the train set matrix A_{ij}:

$$A = \begin{bmatrix} x_{ij}(1,1) & x_{ij}(1,2) & L & x_{ij}(1,C) \\ x_{ij}(2,1) & x_{ij}(2,2) & L & x_{ij}(2,C) \\ M & M & M & M \\ x_{ij}(S,1) & x_{ij}(S,2) & L & x_{ij}(S,C) \end{bmatrix}_{S*C} \qquad (4)$$

b)   Calculation of average value related to each class:

$$M_{ij}^c = \frac{1}{S}\sum_{s=1}^{S} A_{ij}(s,c), c = 1,2,...,C \tag{5}$$

c)   Calculation of the variance of each class:

$$V_{ij}^c = \sum_{s=1}^{S}(A_{ij}(s,c) - M_{ij}^c)^2, c = 1,2,...,C \tag{6}$$

d)   Calculation of the overage

$$V_{ij}^W : V_{ij}^W = \frac{1}{C}\sum_{c=1}^{C} V_{ij}^c \tag{7}$$

e)   Calculation of the average

$$M_{ij} : M_{ij} = \frac{1}{S*C}\sum_{c=1}^{C}\sum_{s=1}^{S} A_{ij}(s,c) \tag{8}$$

f)   Calculation of the variance

$$V_{ij}^B : V_{ij}^B = \sum_{c=1}^{C}\sum_{s=1}^{S}(A_{ij}(s,c) - M_{ij})^2 \tag{9}$$

g)   Estimation of the DP value in (i,j) position:

$$D(i,j) = \frac{V_{ij}^B}{V_{ij}^W}, 1 \le i \le M, 1 \le j \le N \tag{10}$$

4.   FACE CLASSIFICATION
4.1.   Support vector machine
Facial recognition classification is becoming one of the foremost challenges in the area of computer vision & artificial intelligence. The facial recognition and classification depend upon gesture, pose, facial expression, etc. In our work, we use two best classifiers, which yield the better recognition rate easily and effectively. A support vector machine (SVM) algorithms is widely used in many artificial intelligence application and they have been employed on classification and regression trials. Support vector machines are known as hyperplane classifiers, since they are particularly suited to draw separating lines to distinguish between objects of different class memberships. Thus, the fundamental idea of SVM classifiers is to calculate a hyperplane of the maximal margin. This hyperplane separates the cases of the different class labels. To construct an optimal hyperplane, SVM employs an iterative training algorithm, which is used to minimize an error function. Hence, the operation of SVM algorithm is based on constructing a hyperplane that gives the largest distance to the nearest data of any class's train set. This distance is defined as margin within SVM's theory. So the ideal hyperplane maximizes the margin of the training data, since in general the generalization error of the classifier is lower when the margin is larger. In this algorithm, it should be noticed that a line is bad if it passes very closely to the points since it will be sensitive to the noise and it will not ensure a correct generalization. Therefore, our goal should be to find the line passing as far as possible from all points.

4.2.   Random forests
Significant improvements in classification accuracy have allowed developing an ensemble of trees and letting them vote for the most popular class. Random forests are a combination of predictors of trees such as each tree depends on the random vector values, and each random vector is sampled independently with the same distribution for all the forest trees Breiman. In RF classifier, we follow many steps:

a)   We consider that we have N training cases. One sample among these is taken randomly. Then it will be the training set to grow the tree.

b)   We consider M is the number of the input variables, let the number m<M, where m is selected randomly data out of M at each node.

c)   All these trees are grown in suitable state and there are no prunings.

d)   A new data is predicted by aggregating the predictions of all of the trees.

Each tree for this algorithm gives one clustering pixels for new object that to be classified. Then the forest takes in consideration only the classification having excellent votes. The final classification results are decided by all the experiences. This classification method has two essential parameters, the first one is the number of features used for splitting each node of decision tree (m, m < M where M is the total number of features), the second parameter is the number of trees (k). In this study, m is equal to sqrt (M), k is equal to 100. We note that the larger becomes the number of the forest trees the more the generalization error for forests of tree classifiers converges to a limit. Also the generalization error depends on the individual trees strength as well as the correlation between them.

5.   EXPERIMENTS AND DISCUSSION

Our algorithms are evaluated on the ORL and the Yale benchmark face databases, so that we can evaluate our proposed methods. Several classifiers are used. The developed algorithms are programmed with MATLAB codes. During the simulations phase, the database is composed by ORL and the Yale image and so they are divided into two sets train set and test sets. Experimental results are obtained by calculating the mean of simulations, which present various possible states like presence of shadows, facial expression change or lighting change of the train and test sets. In our simulations, we used one premask as pm [3 15 3 15]. Table 2 illustrates the complexity of our proposed approaches. The false detection rate evaluated in each execution by $FRR = \frac{n_m}{n_t}$. Where n_m is the number of misclassified images and n_t is the number of the test set images.

Table 2. Database and simulations details.

Database	Number of class	samples in each class	Size of the Train set	Size of the test set	Image initial size	Sampled size
Yale	15	11	6	5	243*320	120*120
ORL	40	10	5	5	112*92	112*112

5.1. Experimental results with Yale database

The Yale database contains 165 facial images, with 256 gray levels, of 15 people, each one having 11 images. This database has major variations, such as diverse facial expressions, variation in illumination conditions and subjects wearing eye glasses. Initially, the number of pixel for treated image is 243*320 pixels, at the aim to reduce the complexity of calculation and minimizing the computing time we have operated with image having 120*120 pixels.

Experimental result compares the performance of two approaches, the premask and the fractional coefficients, using three different classifiers: Euclidean distance, SVM and random forests. The proposed premask pm=[3 15 3 15] and the false recognition rate described by Figure 2(a), leads to evaluate the effectiveness according the number of features coefficients; of used classifiers, and Figure 2 (b, c, d) present the false detection rate in with fractional coefficients.

5.2. Experimental results with the ORL image

The ORL database contains 400 facial images, 40 of 10 people. These images have initially 112*92 pixels with 256 gray levels, and then the size of each image is reduced to 112*112 pixels. The individual images of this database contain a lot of change in expressions, details and scales, but there is no variation in illumination conditions. Figure 3(a) illustrates the variation of false detection rate according of the number of parameters, using many classifiers on the ORL database. As we have already mentioned, there is no change in illumination condition on the ORL database. Reject the three low frequency coefficients increases the false detection rate, since these three coefficients have large discrimination. Due to the same reason, limiting the search area of DPA decreases the correct recognition rate.
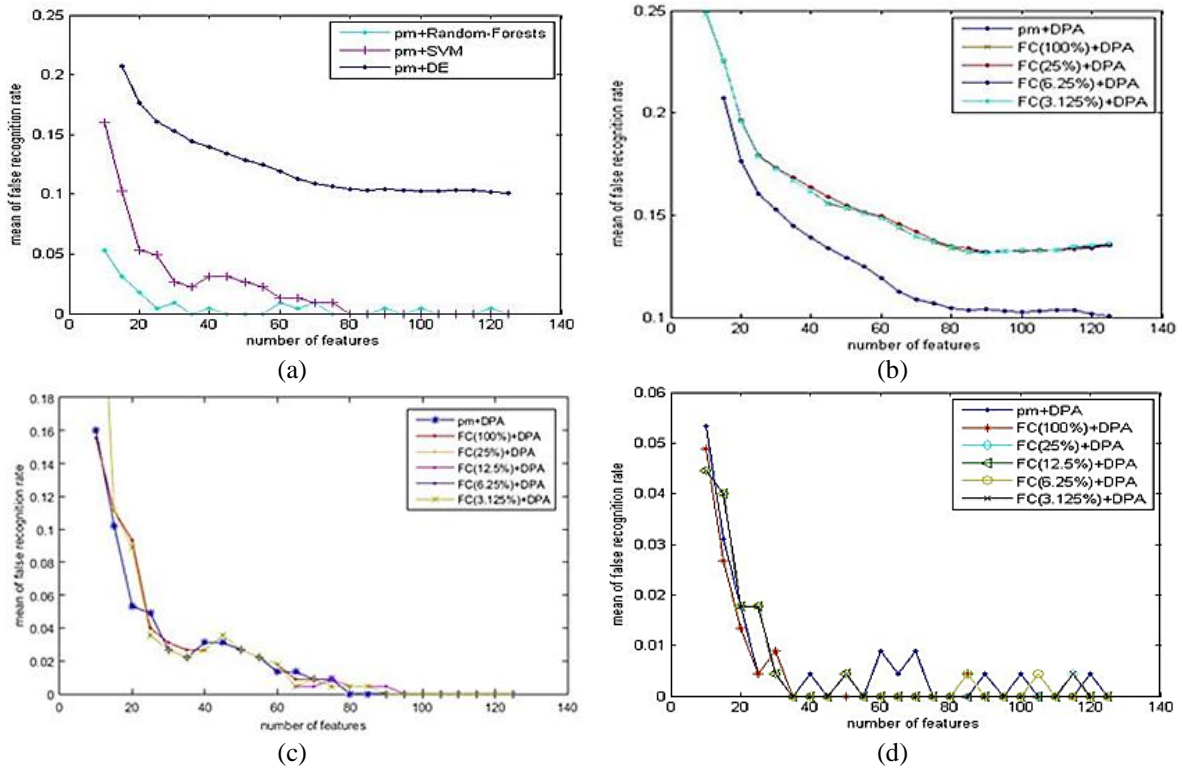
(a)

(b)

(c)

(d)

Figure 2. (a) Mean of false recognition rate with different classifiers Yale database using pm. Mean of false recognition rate with different classifiers on Yale database using fractional coefficients,
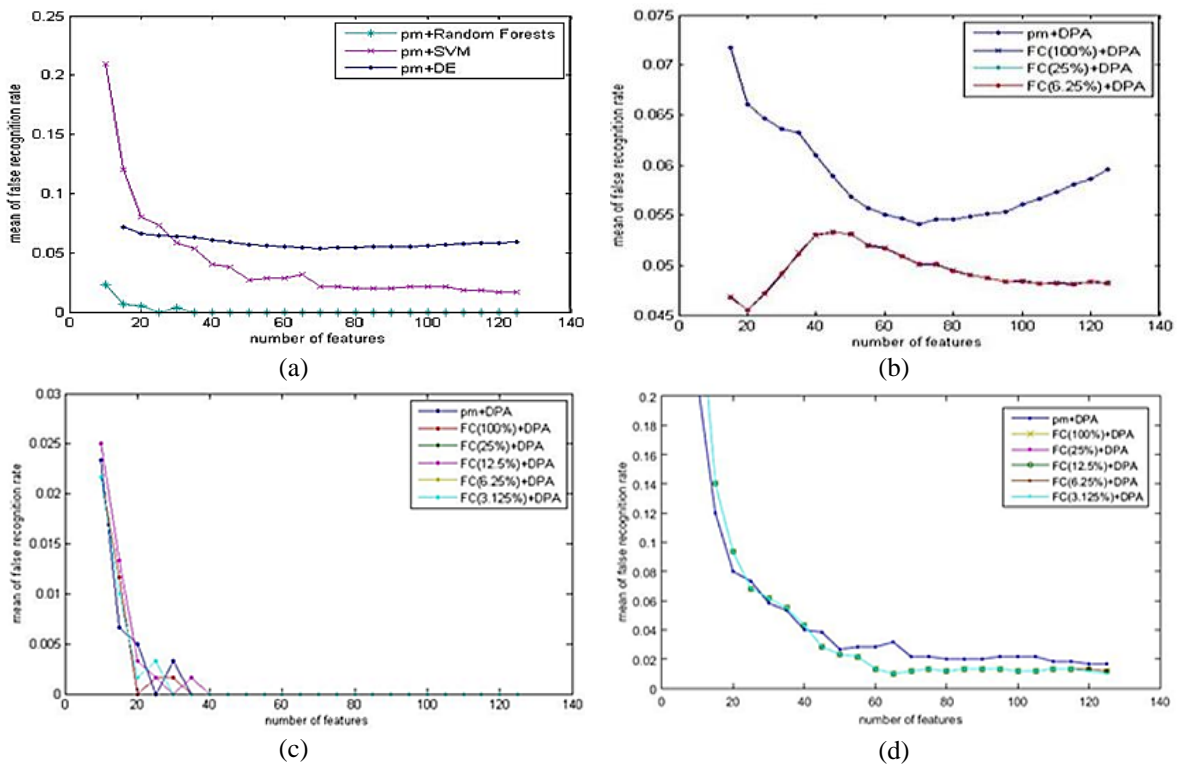(b) DE, (c) SVM, (d) Random forest

(a)

(b)

(c)

(d)

Figure 3. (a) Mean of false detection rate with various classifiers on ORL database using pm. Mean of false recognition rate with different classifiers on ORL database using fractional coefficients (b) DE,
(c) SVM, (d) Random forest

5.3. Discussion

Some results from our experiments are discussed and summarized as follows. As shown in all of the curves, the false detection rate, obtained with the proposed approaches, is under the influence of feature numbers on the considered database. Therefore, the improvements of our approaches depend on the number of coefficients. Among these approaches, pm+DPA has been compared once with different classifiers and once with the Fractional Coefficients approach. As it can be seen from the curves, the use of the premask window in ORL database generates poor results. In fact, on this database, there is no illumination variation, we notice that the suppression of low frequencies, brings up a compromise between good localization of DP values and increasing of false detection rate, indeed the low-frequencies coefficients have large discrimination, and rejecting them by limiting the search area of DPA also decreases the right recognition rate increases the false detection rate. Therefore, using DCT coefficients increases the true detection rate, and the results obtained with the fractional coefficients are more satisfying, especially with the percentages of 6.25%, 3.125% and 3.125% using the Euclidean distance, SVM and random forests respectively, since we keep all of the low-frequency coefficients.

On the other side, pm+DPA in the Yale database reaches the best result, due to the presence of the illumination effects in this database. Thus, for Yale database, premasking the DPA search area allows to decrease cost of computations and to obtain the minimum of false recognition values. Consequently, using the whole DCT coefficients with this database decreases the true recognition rate. However, the fractional coefficients improve the true recognition rate comparing to the whole DCT coefficients, especially with the percentages of 6.25%, 3.125% and 12.5% using the Euclidean distance, SVM and random forests respectively. Concerning the classification, we firstly note that random forest has a randomized nature, so every time we run the code of this former, we get a different result. Yet, the results show its efficiency at any rate, its advantageous classification performance and lower generalization error compared with SVM and the Euclidean distance, on different feature sets due to its desirable characteristics. First, the trees constructed by the random forest, are not pruned. Thus, this algorithm is more robust to noised data. Otherwise, the objective of random forest is to make the trees independent between them.

This independence will allow the vote of the trees be more efficient.Another advantage is that this algorithm is very simple to implement and its digital cost is very small with regard to the performance obtained. Thus, although its irregular results due to its randomized nature, the random forest method used in this work illustrates an enhancement in face recognition by constructing 100 trees, where further such methodology will incorporate with application model for human recognition application.

Besides, we notice that random forest was considered faster in classification than other conventional classifiers cited above, while the random forest doesn't face this problem. Thus when operating with high dimensional data as face identification in large databases, a large time for both training and testing steps are required when working with Euclidean distance and SVM classifiers. Furthermore, the speedup came also from the fact that each tree is grown independently of the other trees grown. However, it should be mentioned that, although the augmentation of tree's number makes the forest more accuracy, it makes it slower. Concerning SVM, it works better on smaller datasets than in high dimensional spaces, as opposed to random forest, where the power of handling large data set with higher dimensionality exists. Moreover, SVM uses a subset of training samples in the decision function. Concerning the random Forest technique it isn't enough sufficient especially when the data set has more noise and overlapping between classes, as it can face the curse of overfitting. Now arriving to the Euclidean distance, we note that it is the most commonly used and most well-known measure of similarity for its simplicity. However, is very sensitive to small deformations or variable variations, namely, it gives more weight to large variations. That is why we found it outperformed by SVM and random forests. The Euclidean distance is not always appropriate because of its lack of robustness against distortions; it is not able to achieve a very good accuracy and does not allow a good separation of classes.

6.   CONCLUSION

Many contribution are developed in this paper, the first one is the use of the fractional coefficients approach for facial recognition field. Also, we have applied SVM and random forest in face recognition, along with the Euclidean distance. The performances of random forest and SVM to classify the feature vectors were compared and analyzed, and the results of classification issued from diverse feature sets recommended that our approach for selecting the DCT coefficients had a distinct advantage. In fact, this paper demonstrates that random forests gives good results in term of accuracies when compared to Euclidean distance and SVM while testing our algorithm of face recognition. In fact, unlike SVM, the random forest trees use revealing features automatically and more frequently in training process, as they reach independent forecasts, which were combined to gain accurate prediction of the forest. Therefore, due to its flexibility

nature and hierarchical structure, the random forest generalizes well and robust, and presents significant superiority in failure tolerances, which plausibly explain its outperformance compared to other approaches significantly for almost all the feature sets in our study. In fact, unlike SVM, individual decision trees in the random forest informative functions automatically use more frequently in the training process and achieve independent forecasts that were combined to gain accurate prediction of the forest.

REFERENCES

[1]   P. Tome, R. Vera-Rodriguez, J. Fierrez and J. Ortega-Garcia, "Facial soft biometric features for forensic face recognition," *Forensic Science International*, vol. 257, pp. 271-284, 2015.
[2]   E. Winarno, et al., "Asymmetrical Half-Join Method on Dual Vision Face Recognition," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 7, no. 6, pp. 3411-3420, 2017.
[3]   M. Moussa, M. Hamila, A. Douik, "A Novel Face Recognition Approach Based on Genetic Algorithm Optimization," *Studies in Informatics and Control*, vol. 27, no. 1, pp. 127-134, 2018.
[4]   S. Kannan and V. Seenivasagam, "Enhanced Biometric Authentication Using Multi Feature Signature Resemblance and Multi View Edge Sectional Similarity," *Asian Journal of Information Technology*, vol. 15, pp. 2542-2548, 2016.
[5]   B. Belean, M. Streza, S. Crisan, S. Emerich, "Dorsal Hand Vein Pattern Analysis and Neural Networks for Biometric Authentication," *Studies in Informatics and Control*, vol. 26, no. 3, pp. 305-314, 2017.
[6]   M. De-la-Torre, E. Granger, R. Sabourin and D. O. Gorodnichy, "Adaptive skewsensitive ensembles for face recognition in video surveillance," *Pattern Recognition*, vol. 48, no. 11, pp. 3385-3406, 2015.
[7]   M. Karovaliyaa et al., "Enhanced security for ATM machine with OTP and Facial recognition features," *Procedia Computer Science*, vol. 45, pp. 390-396, 2015.
[8]   M. Moussa, H. El Ouni, A. Douik, "Edge Detection Based on Fuzzy Logic and Hybrid Types of Shannon Entropy," *Journal of Circuits, Systems, and Computers*, vol. 29, no. 14, pp. 2050227-1 – 2050227-20, 2020.
[9]   C. Sun and E. Yang, "An Efficient DCT-Based Image Compression System Based on Laplacian Transparent Composite Model," *2014 IEEE International Conference on Image Processing (ICIP)*, Paris, pp. 5611-5615, 2015.
[10] N. L. A. Krisshna, V. K. Deepak, K. Manikantan and S. Ramachandran, "Face recognition using transform domain feature extraction and PSO-based feature selection," *Applied Soft Computing*, vol. 22, pp. 141-161, 2014.
[11] M. De-la-Torre, E. Granger, R. Sabourin and D. O. Gorodnichy, "A Novel Preprocessing Technique for DCT-domain Palm-print Recognition," *International Journal of Scientific & Technology Research*, vol. 1, no. 3, pp. 31-35, 2012.
[12] T. L.T. da Silveira, R. S. Oliveira, F. M. Bayer, R. J. Cintra, A. Madanayake, "Multiplierless 16-point DCT approximation for low-complexity image and video coding," *Signal, Image and Video Processing*, vol. 11, pp. 227-233, 2017.
[13] T. L. T. da Silveira, F. M. Bayer, R. J. Cintra, S. Kulasekera, A. Madanayake and A. J. Kozakevicius, "An orthogonal 16-point approximate DCT for image and video compression," *Multidimensional Systems and Signal Processing*, vol. 27, pp. 87-104, 2016.
[14] D. Ramasubramanian and Y. V. Venkatesh, "Encoding and recognition of faces based on the human visual model and DCT," *Pattern Recognition*, vol. 34, no. 12, pp. 2447-2458, 2001.
[15] S. Dabbaghchiana et al., "Feature extraction using discrete cosine transform and discrimination power analysis with a face recognition technology," *Pattern Recognition*, vol. 43, no. 4, pp. 1431-1440, 2010.
[16] Y. Hong-tao et al., "Face Recognition Based on DCT and LDA," *Acta Electronica Sinica*, vol. 37, no. 10, pp. 2211-2214, 2010.
[17] A.S. Eesa, Z. Orman and A.M.A. Brifcani, "A new feature selection model based on ID3 and bees algorithm for intrusion detection system," *Turkish Journal of Electrical Engineering & Computer Sciences*, vol. 23, no. 2, pp. 615-622, 2015.
[18] R. Maheshwari, M. Kumar and S. Kumar, "Optimization of Feature Selection in Face Recognition System Using Differential Evolution and Genetic Algorithm," *Proceedings of Fifth International Conference on Soft Computing for Problem Solving*, vol. 437, pp. 363-374, 2016.
[19] H. Abbas, et al., "Biological landmark vs quasi-landmarks for 3D face recognition and gender classification," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 5, pp. 4069-4076, 2019.
[20] A. I. Salhi, M. Kardouchi and N. Belacel, "Fast and efficient face recognition system using random forest and histograms of oriented gradients," 2012 BIOSIG-Proceedings of the International Conference of Biometrics Special Interest Group (BIOSIG), Darmstadt, pp. 1-11, 2012.
[21] C. Cortes and V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, pp. 273-297, 1995.
[22] K. Padmavathi, M. Bhat and M.V Karki, "Feature Selection based on PCA and PSO for Multimodal Medical Image Fusion using DTCWT," *Computer Vision and Pattern Recognition,* pp. 1-8, 2017.
[23] T.-H. Le and L. Bui, "Face Recognition Based on SVM and 2DPCA," *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 4, no. 3, pp. 85-94, 2011.
[24] Y. Wang and B. Hu, "A More Efficient Face Recognition Framework Based on Illumination Compensation, Kernel PCA and SVM," *2014 Seventh International Symposium on Computational Intelligence and Design*, Hangzhou, pp. 124-129, 2014.
[25] S. Kim, Y. J. Park, K. Toh and S. Lee, "SVM-based feature extraction for face recognition," *Pattern Recognition*, vol. 43, no. 8, pp. 2871-2881, 2010.