

2D-DCT Distance Based Face Recognition Using a Reduced Number of Coefficients

Derzu Omaia
PPGI / DI / UFPB

Brasil

derzu@lavid.ufpb.br

JanKees v. d. Poel
PPGEM / DEM / UFPB

Brasil

jkvdpoel@yahoo.com.br

Leonardo V. Batista
PPGI / DI / UFPB

PPGEM / DEM / UFPB

Brasil

leonardo@di.ufpb.br

ABSTRACT

Automatic face recognition is a challenging problem, since human faces have a complex pattern. This paper presents a technique for recognition of frontal human faces on gray scale images. In this technique, the distance between the Discrete Cosine Transform (DCT) of the face under evaluation and all the DCTs of the faces database are computed. The faces with the shortest distances probably belong to the same person; therefore this evaluating face is attributed to this person. The distance is calculated as the sum of the differences between the modules of DCT coefficients. Only a few coefficients are used in this computation; they are selected from the low frequency of the DCT. Experimental tests on the ORL database reaches a recognition rate of 99.75%, with low computational cost and no preprocessing step. Additionally, the method achieved 100.0% of recognition accuracy when applying a zooming normalization over the ORL database.

KEYWORDS

Face Recognition, Discrete Cosine Transform, Feature Selection, Classification, Performance.

I. INTRODUCTION

Systems for biometric pattern recognition are widely used in security area. These systems use unique human characteristics such as fingerprints, iris, voice, face, enabling differentiation among human beings [14].

Facial research in computer vision can be divided into several areas, such as face recognition, face detection, facial expressions analysis, among others [14]. Face recognition systems have a wide range of application, especially when dealing with security applications, like computer and physical access control, real-time subject identification and authentication, and criminal screening and surveillance. It is also a research topic in several fields, like image processing, neural networks, computational vision, computer graphics and psychology[14].

The biggest difficulty in developing a robust face recognizer is that a human face can undergo several transformations: the same person may use different face accessories like glasses, earrings, piercings and makeup; he

can change his hair into long/short/skin/fringe and even dye it; can wear beard and mustache; can be at many facial expressions like smiling, sad or doing a grimace; and he can be at many different ages. These facial changes make the recognition task very difficult – even human beings do some recognition mistakes from time to time.

This article presents a holistic face recognition method which is based on the magnitude of the Discrete Cosine Transform (DCT) coefficients. The most relevant coefficients for recognition are selected using the technique of low frequencies selection [11]. To classify a face, its DCT coefficients are extracted. Next, the distances between the DCT coefficients of the face to be classified and the DCT coefficients of all faces in the database are calculated. The shortest of all distances will probably be associated with faces of the same person; therefore the face under classification will be classified as belonging to the person whose face in the database has the shortest distance to the face to be classified.

The Order-One Minkowski Metric was used to calculate the distances between the face coefficients. This distance is calculated as the sum of the module of the differences between the absolute values of the DCT coefficients and it turns to be simple and efficient.

Several methods described in literature are also based on mathematical transforms, such as DCT, Karhunen-Loève Transform (KLT) also known as Principal Components Analysis (PCA) and the Wavelet Transform [11]. Many DCT-based face recognition methods are being proposed, mainly because of some DCT properties and the existence of fast algorithms to perform its computation. Those methods are achieving high rates of hits, comparable with those obtained by methods based on KLT and Wavelet [6].

Some of the DCT-based methods are not holistic, as they only use blocks of DCTs on specific face positions, such as eyes, nose and mouth [3][10][12]. Others are holistic and use DCT to select the attributes and then use classifiers based on similarities such as the K-Nearest Neighbor and the Minimum Distance to Prototype [9][11]. Hafed and Levine [9] reached a recognition rate of 93.0%. Podilchuk and Xiaoyu [12] obtained a 94% rate of success, whilst Matos [11] reached 99.25%. Despite Kumar et al. [10] and Bicego et al. [3] reached 100%, they use the DCT

in conjunction with Hidden Markov Models (HMMs), which leads to a system with a high computational cost.

The method proposed in this paper uses the DCT not only to generate but also to select the most important attributes of the faces under analysis, and applies the Nearest Neighbor Classifier to classify the test face among all training faces. The work presented here has results that exceeds those exposed in [9] and in [11] and improve the previously obtained results through some changes in the methods.

Using the Olivetti Research Lab (ORL) database and the leave-one-out training/classification approach, a success rate of 99.75% was achieved. This rate represents 1 error in 400 classifications. These results were obtained at a low computational cost, since DCT computing techniques are very effective and the Minkowski Order-One distance is very fast to calculate.

II. DISCRETE COSINE TRANSFORM

The mathematical theory of linear transforms plays a very important role in the signal and image processing area. They generate a set of coefficients from which it is possible to restore the original samples of the signal.

In many situations, a mathematical operation – generally known as a transform – is applied to a signal that is being processed, converting it to the frequency domain. With the signal in the frequency domain, it is processed and, finally, converted back to the original domain. A mathematical transform has an important property: when applied to a signal, i.e., they have the ability to generate decorrelated coefficients, concentrating most of the signal's energy in a reduced number of coefficients [2].

The Discrete Cosine Transform (DCT) is an invertible linear transform that can express a finite sequence of data points in terms of a sum of cosine functions oscillating at different frequencies. The original signal is converted to the frequency domain by applying the direct DCT transform and it is possible to convert back the transformed signal to the original domain by applying the inverse DCT transform.

After the original signal has been transformed, its DCT coefficients reflect the importance of the frequencies that are present in it. The very first coefficient refers to the signal's lowest frequency, known as the DC-coefficient, and usually carries the majority of the relevant (the most representative) information from the original signal. The last coefficient refers to the signal's higher frequencies. These higher frequencies generally represent more detailed or fine information of signal and probably have been caused by noise [7]. The rest of the coefficients (those between the first and the last coefficients) carry different information levels of the original signal.

In the image processing field, it is interesting to use a two-dimensional DCT (2D-DCT), because images are intrinsically two-dimensional elements. The standard JPEG,

for example, establishes the use a 2D-DCT at the decorrelation step [13].

Figure 1 shows the application of the DCT on one of the face images obtained from the ORL database. Figure 1.a displays the original image, and Figure 1.b displays the result of applying the DCT on the original image. At Figure 1.b, it is possible to verify that most of the image's energy is concentrated in the upper left corner. This is the region that represents the DCT lowest frequency coefficients.

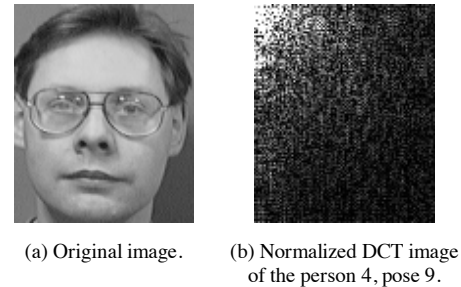


Figure 1: ORL Database: person 4, pose 1, and its DCT.

Figure 2 shows the original face and two reconstructed versions of it after applying the DCT. Figure 2.a corresponds to the original face, with dimensions 92x112, i.e., a matrix with 10,304 values. The next two faces represent the reconstruction of the original image using, respectively, 2,576 e 625 DCT coefficients. To obtain the reconstructed images, the following procedure was adopted: application of the direct DCT transform on the original face, then setting to zero the DCT coefficients to be discarded and, finally, application of the inverse DCT transform on the new matrix of coefficients. Figure 2.b illustrates the reconstruction of the original face considering only the DCT coefficients from the first quadrant, that is, preserving only 25% of the DCT coefficients (coefficients from position [1,1] to [56,46] were retained and the rest of them were put to zero). Figure 2.c illustrates the reconstruction of the original face preserving only 6,07% of the DCT coefficients (coefficients from position [1,1] to [25,25] were retained and the rest of them were put to zero).

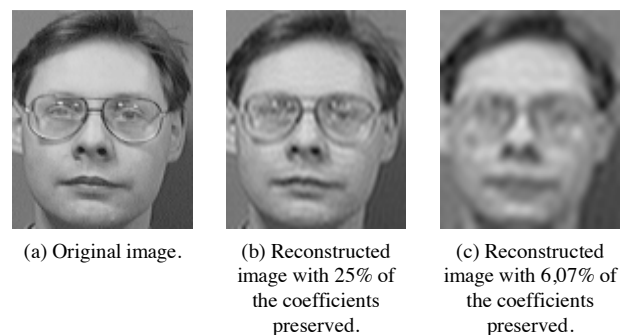


Figure 2: ORL Database: Person 4, pose 1, and its image reconstruction.

From what was shown due to the reconstructed images from Figure 2, it is possible to see that the DCT-based dimensionality reduction is capable of producing good results. The reconstructed images when considering only the low frequency coefficients obtained from applying the DCT do present a reduction in the details, but important information to characterize this images (such as the forehead line, the nose, the mouth, the ears, among others) are preserved. These results do even suggest the complete viability of a face recognition method that uses a DCT-based dimensionality reduction.

The 2D-DCT used in this work is the DCT-II. The DCT-II definition is shown in Equations (1) and (2). In this context, the original image is the gray-scale matrix $x[m,n]$, with dimensions m by n , that represents the image. The DCT-II computation then produces a matrix $X[k,l]$, also with dimensions m by n , of coefficients. The variables m and n are the coordinates in the space domain and k and l are the coordinates in the frequency domain [13].

$$X[k,l] = \frac{2}{N} c_k c_l \sum_{m=0}^{a-1} \sum_{n=0}^{b-1} x[m,n] \cos\left[\frac{(2m+1)k\pi}{2N}\right] \cos\left[\frac{(2n+1)l\pi}{2N}\right] \quad (1)$$

Where, in Equation (1):

$$c_k, c_l = \begin{cases} \left(\frac{1}{2}\right)^{1/2} & \text{to } k=0, l=0 \\ 1 & \text{to } k=1,2,\dots,a-1 \text{ and } l=1,2,\dots,b-1 \end{cases} \quad (2)$$

The first coefficient, $X[0,0]$, is referred as the DC (Direct Current) coefficient and depends only on the average brightness of the image. The other coefficients are known as AC (Alternate Current) factors [13].

In this paper, the DCT used is always the 2D-DCT so, from now on, the term DCT actually means the 2D-DCT.

III. PROPOSED FACE RECOGNITION METHOD

The proposed method is divided into three stages: attribute selection, training and classification. Figure 3 shows a generic face recognition system, where these stages are depicted. The feature selection stage extracts the attributes. The training step chooses the best group of features to perform the identification task, considering all the training samples. The classification stage compares an unknown sample with the ones in the database, using the selected features to determine the best match. In our system no preprocessing stage is performed.

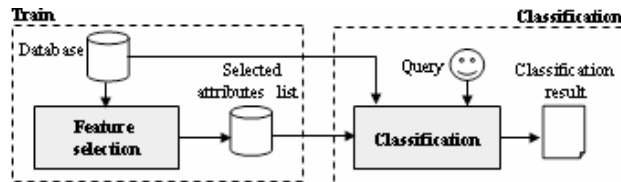


Figure 3: Generic face classifier, adapted from Matos [11].

A. Attribute Selection Stage

Whereas the faces have been previously detected by some other algorithm, the stage of attribute selection specify the list of characteristics that best represents a person, considering the various poses present in the training database. The process of attribute selection extracts the most relevant information to be used to discriminating a person, i.e. the information that is common to a person and is different for the other persons [5].

To perform the attribute extraction, the approach used is the selection of the lowest DCT frequency coefficients. This approach is fast and simple, as it neither evaluates the DCT coefficient of the image, nor performs calculations or any kind of comparison. As the DCT of an image is a 2D signal, the selection of the lower frequency coefficients consists simply in defining a geometric region at the beginning of the 2D signal [11]. An example of such masks for attribute selection using four square regions is shown in Figure 4.a, illustrating this selection approach for low frequencies. It shows the result of applying the 2D-DCT on an image, followed by a normalization stage for easy viewing. As can be seen, the largest energy concentration occurs in the low frequency coefficients region. Thus, this approach is suitable to capture the most important coefficients of the image [11].

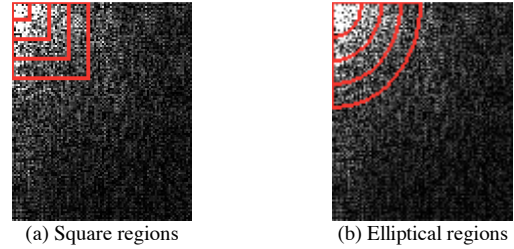


Figure 4: Normalized DCT from person 4, pose 1 of the ORL database, and the regions of low frequency.

One of the major contributions of this work is basically on the way to select the most important of the DCT coefficients. This is done by defining a new geometric shape that is to be used to select the attributes. A previous work [11] used square regions on the coefficient selection stage, as in Figure 4.a. But, in the current study, regions with square, rectangular, triangular, and elliptical shapes were tested.

In Figure 4.b we can see that an elliptical shape selects the coefficients of greater amplitude (white dots of the image) better than a square shape, thus selecting the most relevant coefficients of the DCT 2D signal.

B. Training Stage

The training stage performs various attribute selection steps and decides which one produces the best result. With this in mind, the training algorithm tests various geometric

regions in various scales. Tests were performed on square, rectangular, triangular and elliptical regions.

In this paper, the ORL Database of Faces [1] was used to train and test the proposed method. This database contains photos from faces of 40 people, each person in 10 different poses, totalizing 400 photos. The faces were photographed at different moments, with varying lighting, facial expressions (eyes closed/opened, smiling/not smiling), facial poses and facial details (with/without glasses, with/without beard), among other types of variations. The images are in grayscale, with 256 different levels and a dimension of 92x112 pixels. Some examples are shown in Figure 5.

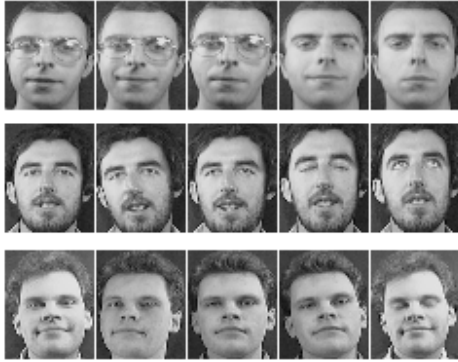


Figure 5: Examples from the ORL database: persons 19, 11 and 36 at their pose 1 to 5.

As the ORL database images are relatively small, tests were performed throughout the first quadrant region of the DCT. This quadrant is the DCT's lower frequency region, and goes from the point (1,1) to (46,56). Over this reduced region of the DCT, a brute force test was used to determine which geometric region and size leads to the best result, which were obtained with elliptical regions of size 6x7 containing 29 coefficients.

These results were obtained over the ORL database. For each different face database, all the training must be repeated, since the choice of the most important group of coefficients is global and not a group for each person. This indicates that, at each insertion of a new person into the database, all the training should be repeated. That happens because the performance of a classifier depends on the interrelationship between the number of selected attributes (space dimension) and the number of samples of class, this problem is known as the Curse of Dimensionality [4].

C. Classification Stage

The classification stage compares a test face with a series of faces from a database, considering the list of attributes selected, and indicates to whom the test face best matches.

Here, the comparison is done using a similarity measure that indicates the resemblance between the test face and all the other faces. The proposed method used, as its similarity

measure, the Order-One Minkowski distance between the amplitude of the coefficients of the training poses and the test face coefficients amplitudes. This distance is calculated as the sum of the differences between the modules of the DCT coefficients.

The classification approach used is the Nearest Neighbor (NN). This approach was chosen among others, because it showed the best results in previous works such as Matos [11]. The Nearest Neighbor classifier calculates the distance of the face under evaluation for all other faces (neighbors) in database. The face in question will be attributed to the person belonging to the face with the smallest distance (nearest neighbor). Theoretically, NN can be defined as follows.

Given the m selected coefficients for the person j , $\{y_{1j}, y_{2j}, \dots, y_{mj}\}$, and the amplitudes of the training coefficients from the person j in pose k , $\{w_{1jk}, w_{2jk}, \dots, w_{mjk}\}$, where w_{ijk} being the coefficient of same position as y_{ij} .

Consider now a test face f and its DCT coefficients' amplitudes $\{v_{1f}, v_{2f}, \dots, v_{mf}\}$, with v_{if} a coefficient in the same position as y_{ij} .

The distance between a test face f and a person j in pose k , with $j = 1, 2, \dots, p$ and $k = 1, 2, \dots, q$, is given by:

$$D_{NNfjk} = \sum_{i=1}^m |w_{ijk} - v_{if}| \quad (3)$$

So, according to Equation (4), a test face f is classified as the person j when:

$$D_{NNfjk} \leq D_{NNfgh}, \forall j \neq g, \forall k \neq h \quad (4)$$

The minimum degree of similarity is zero, and that value means that the face being classified is exactly the same of some of the faces in the database. A negative value does not happen because the used distance has absolute values.

IV. TESTS

Tests were conducted to validate the proposed method. The first one was the common leave-one-out approach and the second was the cumulative leave-one-out. Next, a manual zoom test was performed to the single error found in the best result obtained with the method. Then, the procedure involved in this zoom test was automated.

A. Leave-one-out

Cross validation is a technique that analyses how the results of a statistical analysis will generalize to an independent data set. The leave-one-out test rule is a kind of Cross-validation, which leaves one pose out of the training set and do the classification test with it [8]. As the ORL Face Database [1], (which was) used in the tests, has ten poses for each person stored in it, then ten rounds of tests were done. In the first round, the first pose from all

poses is excluded from the training set and used as the test face. In the second round, the second pose is excluded from the training set and used as the test face and so on. This test was performed several times, with different numbers of selected coefficients, selected using different geometric shapes.

B. Cumulative Leave-one-out

The Cumulative Leave-one-out test performed measure the Accumulative Recognition. In this test, the classifier returns a ranked list, in order of increasing distance, of the faces which are most similar to the test face. A common NN classifier would use only the first element of the list (the one with the smaller distance) to perform the classification. The cumulative classifier indicates the minimum size of the list so that the classifier achieves 100%. That is, the minimum size of the list containing a pose of the same person of the test face. This test was also done to different geometric shapes of attribute selection.

C. Manual Zoom

An analysis was performed on the only error found in our best result. In this error, the person 19 in his pose 9 was assigned to person 11, being more like her in pose 5. Figure 6 shows the 10 poses of the person 19 in ascending order. Pose 9 is emphasized and the last image of the second line is the pose 5 of person 11. It can be seen that the pose 9 of person 19 has a slightly smaller face than the nine other poses of this person. Also, it can be seen, that the face of the person 11 in pose 5 has a closest size to person 19 in pose 9 than the rest of the poses of the person 19.



Figure 6: ORL Database: Person 19 in yours 10 poses and person 11 in pose 5.

A test was performed to verify if the recognizer is sensible to the zoom operation. This test manually applied a 5% zoom to the pose 9 of person 19. With this zoom, this pose now has a closer size when compared to the other poses of the person 19, as can be seen in Figure 7. With this manual change, the recognizer reached a hit rate of 100%. That implies that the recognizer is sensible to the zoom operation.



Figure 7: Person 19, in yours 10 poses. The pose 9 has a 5% zoom in.

D. Automatic zooming normalization

Aiming for a face database more robust to small scale changes, the database used was submitted to an automatic pre-processing zoom normalization. This pre-processing applies, on each face, a zoom in and out of 5%, creating one more face for each zoom operation, tripling the size of the database. Thus, each pose will be composed of three images: the original one, the augmented one and the diminished one. Figure 8 illustrates the application of this preprocessing over a sample image of the ORL database. Even with images of slightly different sizes, the attribute selection is performed in the same way, always using regions of the same size regardless of the size of the image.

Using this new stored database, the test using the common leave-one-out procedure is then executed.



Figure 8: Person 19, on pose 9, and it's zoom application.

V. RESULTS

This topic presents the results of applying the tests described in the previous section on the ORL database and on the zoomed version of this database.

A. Leave-one-out

Figure 9 shows the system performance over the leave-one-out test, and how the number of DCT coefficients used might have an effect on the overall performance of the recognition system when applied to the ORL database. The recognition rate is calculated as the number of correct hits divided by the number of faces present in the database, the result of this division is then multiplied by 100. The results are separated by the four geometric shapes used on the attributes selection, and was used the Nearest Neighbor classifier. It can be observed that there is a decrease in the recognition accuracy rate as the numbers of coefficients increase. That happens because of the Curse of Dimensionality, which states that the performance of a classifier depends on the interrelationship between the

number of selected attributes and the number of samples of classes, and the excessive increase of the number of coefficients causes efficiency decrease [4]. This recognition rate is calculate

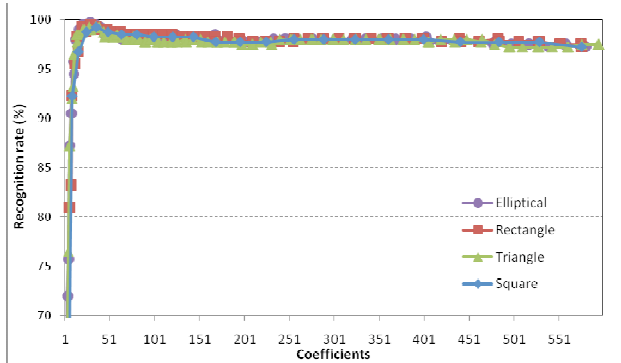


Figure 9: Effect of varying the number of DCT coefficients on recognition accuracy, using geometric regions.

In Figure 10, the curve apex region is highlighted. In this region lies the most significant results. It can be noted that the elliptical format has the best result: only 29 DCT coefficients are enough to achieve an accuracy of 99.75%. This rate represents only one error in 400 classifications. This result was obtained using an elliptical form with a size of 7x6 pixels. If the ORL database was larger, covering more people also would require a larger amount of coefficients for a correct distinction between the classes, due to the dimensionality curse.

The square region always crosses the rectangular region, since a rectangular region can also be a square one. The rectangular region has hit rates greater than the square of the region, reaching 99.50%, while the square only reaches 99.25%. However, the square region was shown for comparison with the previous work of Matos [11], which used only square regions. Using elliptical forms, the detection rate was 99.75%. Comparing to the Matos' work, the error rate was reduced from 0.75% to only 0.25%. This was a 66% substantial reduction in the error rate.

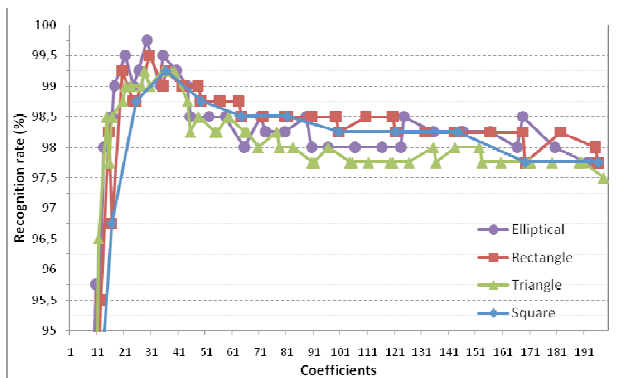


Figure 10: Zoom at the region with best results from Figure 9.

B. Cumulative leave-one-out

Using this approach, the system reaches 100% of recognition accuracy when the correct face is within a range of six returned faces. This range of six faces was obtained with both rectangular and triangular regions. The elliptical region reached 100% only within seven faces returned. This result was not the expected, because the elliptical regions have been more efficient under the common leave-one-out approach. The square region reached the same result only within 9 faces. Figure 11 shows these results.

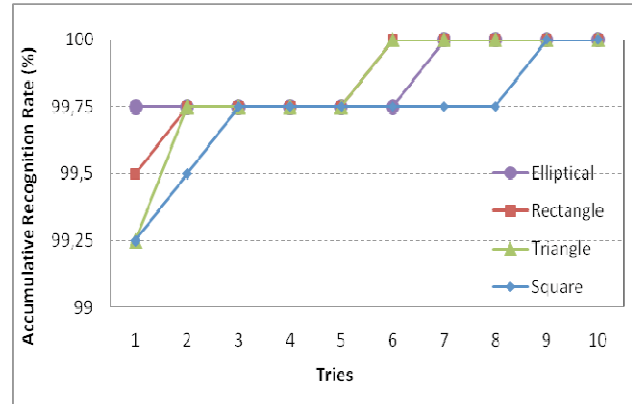


Figure 11: Accumulative Recognition

It is interesting to see that all the four shapes reached a recognition rate higher than 99% when returning only the first face. The results with just one returned face can also be perceived in Figure 10, by selecting only the best results for each geometric shape.

C. Automatic zooming normalization

Using the Automatic Zoom Normalization approach on the ORL database, a recognition rate of 100% was achieved for all the geometric shapes used in the attribute selection stage. This shows that the zooming normalization is a very important step in order to make the system more robust to small scale variations. This sensitivity to zoom is due to the intrinsic properties of the DCT. However, by applying a zoom in and out of only 5% on a face database where the faces size are slightly normalized, it is sufficient to overcome the sensitivity of the DCT.

Table 1 shows the regions, their size, and the number of coefficients used to achieve 100% of recognition. It can be noted that, within very small regions, it is possible to obtain 100%. In the common leave-one-out approach, the best result was reached by using elliptical regions of size 6x7, leading to only 29 coefficients. Using the same approach, but now applying the Zooming Normalization procedure, the system achieves the optimal result with regions of size 3x4, with only 12 coefficients.

Shape	Size	coeffients
square	4x4	16
rectangle	3x4	12
triangular	4x5	14
elliptical	4x5	13

Table 1: Shape, size and coefficients that hit 100%.

D. Processing time

The Face Recognition System proposed in this paper was tested on a 2.0 GHz AMD Turion64 single core computer with 1GB of RAM, running Ubuntu Linux, a common PC of nowadays. Using this platform, the execution time with two different implementations was verified: one developed in Java and the other developed in C++, both implemented during the development of this work.

Most of the execution time is spent with the generation of the DCTs of the face database – approximately 99% of the execution time. However, this step can be performed only once and its result stored for a future use. Table 2 shows the execution time of the leave-one-out test with the ORL database using the C++ implementation, and Table 3 shows the execution time using the Java implementation. The second column of these tables shows the execution time when the DCTs are fully generated, and the third column shows the time when the DCTs were pre-computed and are only loaded from the disk. The leave-one-out round (second row) has 400 classifications, and the time to classify only one face (third row) is the time of the leave-one-out round divided by 400.

	Execution time (seconds)	
	Generating DCTs	Loading DCTs
Obtain DCTs	15.660030	0.076801
Leave-one-out round	0.119613	0.119613
1 face classification	0.000299	0.000299
Total time	15.779643	0.196414

Table 2: C++ execution time.

	Execution time (seconds)	
	Generating DCTs	Loading DCTs
Obtain DCTs	42,400000	0,789000
Leave-one-out round	0,202000	0,202000
1 face classification	0,000505	0,000505
Total time	42,602000	0,991000

Table 3: Java execution time.

It can be seen that the C++ version is faster in all aspects, and it is 10 times faster on doing the DCT loading operation. Thus, the C++ implementation is more suitable for real-time applications.

VI. DISCUSSION AND CONCLUSIONS

The proposed face recognition method has proved to be valid in the performed tests, achieving good results in several evaluation categories as recognition accuracy, robustness and computational cost.

Without any pre-processing step, the proposed method achieves a high recognition accuracy of 99.75% when using only 29 coefficients using the leave-one-out approach. On the cumulative leave-one-out the system reaches 100% of correct classification when the six most similar faces are returned. When applying the automatic zooming normalization the system also reaches 100% of correct detection rate.

The proposed method is also suitable for real time applications: in the experimental tests, the classification processing time for only one face, using 29 coefficients, is near 0.0002 seconds, in a common PC, without taking into account the prior feature selection processing time. Due to energy compaction property of DCT it is possible to reduce the processing of 10.304 (112x92) pixels to 29 DCT coefficients.

Testing the proposed algorithm using only the ORL database is not enough to fully validate the method, but it's important to compare methods. Despite the fact that in literature this face database is always being referred, it does not include all variations needed for evaluating the robustness of the method.

Future developments to the method will focus on improving the processing time on very large databases, and integrate the system with a face detector method. Additionally, tests must be done on other face databases.

VII. REFERENCES

- [1] AT&T Laboratories, Cambridge, UK, "The ORL Database of Faces" (Now At&T "The Database of Faces"), Available [Online]: http://www.cl.cam.ac.uk/research/dtg/attarchive/pub/data/att_faces.zip [November, 16, 2008], 1994.
- [2] Batista, L. V. Compressão de Sinais Eletrocardiográficos Baseada na Transformada Cosseno Discreta. D.Sc. Thesis in Electrical Engineering, PPGEE /DEE/UFPB, Campina Grande, Brazil, 2002.
- [3] Bicego, M., Castellani, U, And Murino V, Using HMM and Wavelets for Face Recognition, Proceedings of the 12th International Conference on Image Analysis and Processing, 2003 IEEE.
- [4] Campos, T. E., Técnicas de Seleção de Características com aplicação em Reconhecimento de Faces, Dissertação de Mestrado, USP, São Paulo, Brasil, 2001.
- [5] Duda, R. O.; Hart, P. E.; and Stork, D. G. Pattern Classification. John Wiley & Sons, Inc., 2nd edition, 2003.

- [6] Ekenel, H. K.; Goa, S. H.; Fischerm, M.; and Stiefelhagen, R. Face Recognition for Smart Interactions, IEEE ICME, 2007.
- [7] Gonzalez, Rafael C.; Woods, Richard R.. Digital Image Processing Using Matlab. 3ED. PRENTICE HALL, 2007.
- [8] Haykin, Simon. Neural Networks and Learning Machines. 3rd Edition. Prentice Hall, 2008.
- [9] Hafed, Ziad M and Levine, Marin D. Face Recognition Using Discrete Cosine Transform. International Journal of Computer Vision, v. 43(3), p. 167-188. 2001.
- [10] Kumar, S, A, S; Deepti, D, R, And Prabhakar, B, Face Recognition Using Pseudo-2d Ergodic HMM, IEEE, ICASSP 2006.
- [11] Matos, F. M. de S. Reconhecimento de Faces Utilizando Seleção de Coeficientes da Transformada Cosseno Discreta. M.Sc. Dissertation in Computer Science, PPGI/DI/UFPB, João Pessoa, Brazil, 2008.
- [12] Podilchuk, C. and Zhang, X. 1996. Face recognition using DCT-based feature vectors. In Proceedings of the Acoustics, Speech, and Signal Processing, 1996. on Conference Proceedings., 1996 IEEE international Conference – Volume 04 (May 07 - 10, 1996). ICASSP. IEEE Computer Society, Washington, DC, 2144-2147.
- [13] Rao, k. R.; YIP, P.. Discrete Cosine Transform: Algorithms, Advantages, Applications. Academic Press, Inc, 1990.
- [14] Zhao, W.; Chellappa, R.; Phillips, P. J., Rosenfeld, A.; Face Recognition: A Literature Survey, ACM Computing Surveys, V, 35, N, 4, P, 399-458, 2003.