A New Preselection Method for Face Recognition in JPEG Domain Based on Face Segmentation

Alireza Sepas-Moghaddam #1, Mohammad Shahram Moin #2, Hamidreza Rashidy Kanan *3

#Electrical, Computer and IT Engineering Department, Islamic Azad University, Qazvin Branch, Qazvin, Iran
1 sepasmoghaddam@ieee.org
2 moin@qiau.ac.ir
*Electrical Engineering Department, Bu-Ali Sina University, Hamedan, Iran
3 h.rashidykanan@basu.ac.ir

Abstract—Face recognition in JPEG compressed domain has turned into one of the important standpoints for reducing the computational overhead of decompression process without degradation in recognition accuracy. This domain needs some efficient methods for preselecting compressed coefficients and performing recognition process, which leads to improving the recognition rates and decreasing the computational complexities of the feature extraction methods. In this paper, a novel preselection method is proposed for these goals. For the first time a more efficient decompression process is presented, by performing face recognition in zigzag scanned coefficients. In the proposed method, the area of the face is segmented in prominent and non-important regions, and subsequently, the DC and the different number of lower frequency AC coefficients of each block are preselected, regarding the regions used for face recognition process. Experimental results show that the proposed method outperforms existing methods in recognition rate, as well as in time and space complexity aspects.

I. INTRODUCTION

Time and space complexities, which represent an estimation of programs' runtime and the required memory, respectively, as well as disk storage space are amongst the most fundamental challenges in designing biometric systems. Face recognition systems are a type of biometric systems that are widely used in legal and commercial applications with high universality, collectability, acceptability and circumvention [1]. On the other hand, JPEG [2] is considered as a common method for still images’ compression with significant compression ratio, without noticeable degradation in their visual quality. In addition, most of the modern cameras store directly the output images based on this method by embedded software or hardware. Thus, using the JPEG compressed images either mandatory or not, one can markedly solve the storage space problem in face recognition systems.

Block diagrams of JPEG coding and decoding compression method are shown in Fig. 1. In the first step of coding process, i.e. 2-D DCT is applied on 8*8 blocks of image pixels and, subsequently the transformed coefficients are quantized using a quantization table. In the zigzag step, 2-D blocks of image are mapped to a 1-D sequence, through zigzag scan. Then DPCM and RLC methods are applied on DC and AC coefficients. Finally, the coefficients are coded using an entropy coder. The JPEG decoder is composed of the inverse processing steps in reverse order.

Figure 1. Block diagram of JPEG (A) coding (B) decoding procedures.

Traditionally, in face recognition systems, the compressed images must be fully decompressed, which involves a considerable computational overhead in recognition process, caused by the decompression process [3]. To reduce the
impact of this overhead, researchers have proposed to perform face recognition in compressed domain ([4],[5],and [6]). Delac and et al. defined compressed domain as "transform coefficients which are used as input to face recognition systems instead of pixel values"[5]. In this paper, a novel preselection method for face recognition in JPEG compressed domain on zigzag scanned coefficients (Fig. 1.B) has been proposed for improving the recognition rates as well as reducing more the time and space complexities of matching step and decompression process. In addition, the runtime percentages of the JPEG compression/decompression steps are calculated to obtain an accurate estimation of time reduction resulted by applying face recognition in JPEG compressed domain.

The rest of this paper is organized as follows: Section II is dedicated to a review of related works. Section III contains a description of principles and methods used in this paper. Section IV presents experimental simulations, results and comparisons. Section V contains analysis of results and Section VI concludes the paper.

II. RELATED WORKS

In this section, we will introduce the proposed methods and components used in our simulations. The block diagram of the proposed face recognition system is shown in Fig. 2. First, the images (8bit/pixel) in the database are normalized and registered (Section III-B). Then the registered images are compressed by JPEG compression standard and, subsequently their areas are categorized to prominent and non-important regions (Section III-B). In the next step, decoding is done before inverse zigzag scan step and proposed preselection phase (Section III-C) is applied on zigzag scanned coefficients for feature selection. Then, Principle Components Analysis method (section III-D) is applied on the preselected coefficients for feature extraction and subsequently Cityblock and Spearman correlation distance metrics are used for matching the extracted features of probe and gallery images. These distance metrics are selected, according to the results of [5] and [6]. The results are presented by first recognition rank as well as cumulative recognition rank metrics [10].

It is worth taking to consideration, since uncompressed FERET database images are used in this research, we first had to compress them and then partially decompress them to extract quantized coefficients. However, in a real life scenario, in which compressed images are used, we would only have to partially decompress them to extract the appropriate coefficients. The standard quantization table, specified in the JPEG standard [2] (shown in Fig.3) has been used for compressing the images of database, due to this fact that most of the real-world JPEG compressed images are compressed by this table.

![Block diagram of the proposed face recognition system.](Image)

A. Face Image Database

The grayscale images of FERET database [9] is used in this research with size of 384x256 pixels and 256 gray levels/pixel, i.e. 8 bit/pixel. Experiments have been done on four datasets...
in a standard test space, specified in [10]. These datasets are 
Fb (different expression test), Fc (different illumination), 
Dup1 (images taken anywhere between one minute and 1,031 
days after the gallery image) and Dup2 (images taken at least 
18 months after the gallery image was taken). These datasets 
contain respectively 1195, 194, 722 and 234 images that are 
compared with 1196 reference images in Fa dataset.

\[
Q = \begin{bmatrix}
16 & 11 & 10 & 16 & 24 & 40 & 61 \\
12 & 12 & 14 & 19 & 26 & 58 & 60 \\
13 & 13 & 16 & 24 & 40 & 57 & 69 \\
14 & 17 & 22 & 29 & 51 & 87 & 80 \\
18 & 22 & 37 & 56 & 68 & 109 & 103 \\
24 & 35 & 55 & 64 & 81 & 104 & 113 \\
49 & 64 & 78 & 87 & 103 & 121 & 120 \\
72 & 92 & 95 & 98 & 112 & 100 & 103 \\
\end{bmatrix}
\]

Figure 3. Standard quantization table [2].

B. Image Registration, Normalization and Segmentation

In this research, we have used the approaches presented in 
[10] for registering and normalizing the images. For this 
purpose, after aligning eyes, the distance between eyes is 
converted to 70 pixels, and then the image size is modified to 
150 x 130 pixels after moving the eyes to fixed coordinates. 
Finally, the histogram equalization is used for images contrast 
adjustment. It is worth mentioning that the registration and 
normalization method used in this work are the same used in 
[3], [5], [6], which are the references used as our comparative 
works.

For segmentation process, FERET provides the initial 
coordinates of the eyes, nose and mouths in unregistered 
images. For selecting the prominent region, after registration 
and normalization process, the new coordinates of these 
components are calculated in the registered images as indexes 
of the region. The size of the eyes, nose and mouth regions are 
40*25, 35*55 and 75*35, respectively. We observed that by 
selecting these sizes the desired areas are correctly determined 
for different subjects. These steps are shown in Fig. 4. It is 
worth taking into consideration the image registration and 
normalization are performed in linear time for each image. In 
addition, by adding segmentation step, the time complexity 
still remains linear. Thus, the segmentation method has no 
significant overhead on the proposed system and it has less 
computational complexity than variance analysis method.

Finally, using the method presented in [11], detection, 
registration, normalization and segmentation processes could 
be performed in JPEG compressed domain. Thus, in real-
world applications, JPEG compressed images do not need to 
be fully extracted for this issue.

C. Feature Selection Method

Feature selection methods in JPEG compressed domain are 
studied for improving the recognition accuracy as well as 
reducing the computational complexity of feature extraction 
step. The hypothesis based on which the proposed feature 
selection method is introduced is that all of the details of a 
given face are not useful in recognition process. Indeed, only 
the details in the prominent parts of a face, i.e. eyes, nose and 
mouth contain useful information for the recognition process. 
Thus, in this work, in order to increase and evaluate the 
contribution of the details of these parts in recognition phase, 
different numbers of low frequency coefficients are preselected from these parts of faces.

To perform the preselection process, DC coefficient and 
AC coefficients with lowest frequencies are selected in 1-D 
zigzag scanned sequence. Before zigzag scan step these 
coefficients are located in the first diagonals in 8 x 8 blocks 
of image and after mapping the 8 x 8 blocks through zigzag scan 
to 1-D sequence, they are moved into the first locations 
starting after 64 x i , i = 0 ,..., addresses. Therefore they could be 
selected by Eq.(1) in zigzag scanned sequence. This issue is 
shown in Fig 5.

\[ P = \{X_k ; k = 64 \times i + j, 0 \leq i < n - 1 , 1 \leq j \leq m\} \]  

where n is the number of blocks in image space, P is the 
Preselected set, m is the number of coefficients that a 
preselected in each block and X is the buffer of coefficients. 
In our experiments, 3 and 10 coefficients have been 
considered for m in non-important and prominent areas, 
respectively. The reason for selecting these values will be 
explained in Section 5.

D. Principle Component Analysis (PCA)

Principle component analysis (PCA) [7] is a subspace 
projection technique that is widely used in face recognition. 
Given an s-dimensional vector representation of each face in a 
training set of images, PCA tends to find a t-dimensional 
subspace whose basis vectors correspond to the maximum 
variance direction in the original image space. This new 
subspace has smaller dimensions compared to primary space, 
due to omitting the principle components with smaller 
eigenvalues. The vectors obtained in this subspace should have 
the most information from the primary image. So the energy of 
image remains almost unchanged and the subspace of the 
images can be compared effectively in matching stage.

The most important difference between using PCA in this 
research and in other works is that in the proposed method 
PCA extraction method is applied on compressed coefficients, 
instead of pixel values. In [12], it was proved that PCA 
spaces projection results are stable, when the original data 
have been transformed using an orthogonal transformation e.g. 
DCT. In addition, in the proposed method, the first eigenvector 
is omitted, for reducing the effects of illumination variants. 
Furthermore, final 40% of the eigenvectors are omitted as well, 
due to this fact that few energy are involved amongst them 
[13].
Figure 4. Image Registration, Normalization and Categorization Processes

Figure 5. Mapping 8x8 blocks to zigzag scanned sequence and then to preselected set. In this figure only three coefficients from each block are selected.

IV. Experimental Results and Comparisons

A. Computational Complexities of Compression / Decompression Steps in JPEG

Computational complexity orders of compression/decompression steps are shown in the first row of Table 1, where N is the number of pixels or transformed coefficients and A and D are the number of AC and DC coefficients, respectively. We have used 500 images from the training set to calculate the runtime percentage of different compression/decompression steps. The runtime percentage and cumulative runtime percentage are shown in the second and third rows of Table 1, respectively.

B. The results of Proposed Method

The experiments results are obtained by applying proposed preselection method and two different distance metrics on four datasets Fb, Fc, Dup1 and Dup2. The first rank recognition and the cumulative rank recognition to 50th rank are shown in Table 2. Results show that using Spearman correlation distance metric leads to a slightly better recognition ranks, compared to Cityblock one.

C. A Comparison Between Recognition Rates of the Proposed Method and Other Related Works of Proposed Method

The results of the proposed method in face recognition are compared with those of related works and showed in Table 3. We can see that the proposed method outperforms three linear subspace methods, i.e. PCA, LDA and ICA in pixel domain using uncompressed and fully decompressed images [14]. In addition, the proposed method outperforms results in compressed domain using all DCT and quantized coefficients in compressed domain, mentioned in [5] and [6] and DCT and quantized preselected coefficients by variance analysis method in [5] and [6], where our method has less computational complexity than variance analysis method.

Computational Complexity of the proposed method

Computational complexities of PCA extraction method is $O(N^3)$, in which $N$ indicates the number of input coefficients. If we use the coefficients preselecting methods with $P$ coefficient ($P < N$), only $P$ coefficients will be input for feature extraction methods instead of $N$ coefficients which it causes a $N/P$ speedup ratio in time complexity and a ratio of...
P/N reduction in space complexity of the feature extraction phase

Using 3/64 coefficients in non-important regions together with 10/64 coefficients in prominent regions instead of all coefficients, will speed up the rate of approximate about 13 times. In addition, referring to Table 1, by omitting the inverse DCT, quantization and zigzag scan steps, the proposed method's complexity and runtime, in term of decompression overhead, will be respectively, O(N² + N + N) and 49.62% lower than that of the recognition in fully decompressed domain. These reduction rates in complexity and time will be O(N) and 4.89%, respectively, compared to the best compressed domain recognition results reported in[6].

V. ANALYSIS OF RESULTS

be Regarding the quantized coefficients of the registered and compressed face images, it is observed that about 82% of the coefficients are zero, due to the properties of DCT and the quantization table. To be more precise, almost all the high frequency AC coefficients which are located in the lower triangular matrix as well as the 7th diagonal of each 8*8 block are zero. Removing these coefficients significantly reduces (more than 67%) the image space, leading to a reduction in time and space complexities of matching stage, without any negative or positive impact on the recognition results.

On the other hand, only a portion of the remaining coefficients in the upper triangular matrix is used in the recognition process. There are two reasons for this fact. First, most (in average 75%) of the high frequency AC coefficients located in the 6th diagonal together with those of the 5th diagonal are zero and non-zero values are located in different locations even for different vector obtained from the same subject. Therefore, eliminating them, not only reduce the computational and space complexity, but also can lead to an improvement in the recognition results. In addition, using the AC coefficients located in the 4th and 3rd ones which present details of the face are not necessary for all areas of the face image. That is to say, the details in the prominent components of the face contain profitable information for the recognition process. On the other hand, these coefficients which mean the detail on the other areas e.g. bread and wrinkle on the face's skin could not be useful in this regard.

Considering all the facts which were mentioned, 3 and 10 coefficients are selected for non-important and prominent regions of the face, respectively. Concerning the datasets, our method works efficiently in Fb dataset, due to the fact that the coefficients that carry the information of a subject’s identity are not discarded and the negative effects of facial expressions which appear in high frequency coefficients are significantly reduced. For Fc dataset as mentioned in [6], the differences arise from changing the overall illumination together with the direction of the light source of the image. When such images are compressed, the effect of smearing decreases the differences between the images of the subject caused by such illumination changes and, subsequently, using the proposed method causes the improvement in recognition results. Prominent areas of the face, which carry the information of a subject’s identity, will not face considerable changes during the time, contrary to other areas. Thus, The most important reason for the effectiveness of the proposed method in dup1 and dup2 sets, which are the images taken at different times is its improvement in contribution to the prominent areas as well as the reduction of the non-important areas' role, especially their details, in recognition process.

TABLE I. TIME/COMPLEXITY OF COMPRESSION/DECOMPRESSION STEPS IN JPEG METHOD

<table>
<thead>
<tr>
<th>Time Complexity</th>
<th>Compression/Decompression Steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complexity order</td>
<td>DC Entropy / Decod./Cod.</td>
</tr>
<tr>
<td>Runtime percentage</td>
<td>O(D log D)</td>
</tr>
<tr>
<td>Cumulative Runtime percentage</td>
<td>33.16</td>
</tr>
</tbody>
</table>

TABLE II. RECOGNITION RATES OF THE PROPOSED METHOD, FOR DIFFERENT DATASETS AND METRICS

<table>
<thead>
<tr>
<th>Distance Metrics</th>
<th>Fb</th>
<th>Fc</th>
<th>Dup1</th>
<th>Dup2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman</td>
<td>RR₁</td>
<td>CRR₁₀</td>
<td>RR₁</td>
<td>CRR₁₀</td>
</tr>
<tr>
<td>Cityblock</td>
<td>86.5</td>
<td>97.4</td>
<td>72.2</td>
<td>95.4</td>
</tr>
<tr>
<td></td>
<td>87.1</td>
<td>97.8</td>
<td>77.9</td>
<td>97</td>
</tr>
</tbody>
</table>
### TABLE III. FIRST RANK RECOGNITION RATES OF THE PROPOSED METHOD AND OTHER METHODS

<table>
<thead>
<tr>
<th>Dataset</th>
<th>PCA Decom</th>
<th>LDA Decom</th>
<th>ICA Decom</th>
<th>DCT-PCA Var Anal</th>
<th>DCT-ICA Var Anal</th>
<th>Quantized Var Anal</th>
<th>Proposed method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fb</td>
<td>79.4</td>
<td>75.5</td>
<td>83.1</td>
<td>83.2</td>
<td>82.5</td>
<td>84.4</td>
<td>83.5</td>
</tr>
<tr>
<td>Fc</td>
<td>50</td>
<td>11.3</td>
<td>68</td>
<td>71.6</td>
<td>77.3</td>
<td>73.2</td>
<td>72.2</td>
</tr>
<tr>
<td>Dup1</td>
<td>39.2</td>
<td>35.6</td>
<td>45.3</td>
<td>46.7</td>
<td>39.1</td>
<td>43.5</td>
<td>46.1</td>
</tr>
<tr>
<td>Dup2</td>
<td>20.5</td>
<td>13.2</td>
<td>31.2</td>
<td>21.4</td>
<td>34.2</td>
<td>22.2</td>
<td>29.1</td>
</tr>
</tbody>
</table>

## VI. CONCLUSION

In this paper, a novel method for preselecting lower frequencies coefficients in JPEG domain face recognition has been proposed, considering the regions of a face. The experiments were performed on FERET database and the results were presented by first recognition rank, as well as cumulative recognition rank metrics.

The simulations show that our proposed preselection method outperforms other methods in term of recognition rate. We also reached a considerable reduction in time and space complexity aspects of the matching phase and that of the decompression process, by performing face recognition in zigzag scanned coefficients.

As future works, to reduce more the computational complexity of decompression process, the proposed preselection method will be investigated to be performed on entropy compressed coefficients.

## ACKNOWLEDGMENT

Portions of the research in this paper use the FERET database of facial images collected under the FERET program, sponsored by the DOD Counterdrug Technology Development Program Office.

## REFERENCES


