Hindawi Publishing Corporation
EURASIP Journal on Applied Signal Processing
Volume 2006, Article ID 92734, Pages 1–13
DOI 10.1155/ASP/2006/92734

A Comparison of Set Redundancy Compression Techniques

Samy Ait-Aoudia and Abdelhalim Gabis

Institut National d'Informatique (INI), BP 68M, Oued Smar 16270, Algiers, Algeria

Received 27 February 2005; Revised 30 November 2005; Accepted 21 January 2006

Medical imaging applications produce large sets of similar images. Thus a compression technique is necessary to reduce space storage. Lossless compression methods are necessary in such critical applications. Set redundancy compression (SRC) methods exploit the interimage redundancy and achieve better results than individual image compression techniques when applied to sets of similar images. In this paper, we make a comparative study of SRC methods on sample datasets using various archivers. We also propose a new SRC method and compare it to existing SRC techniques.

1. INTRODUCTION

Medical imaging applications produce a huge amount of similar images. Storing such amount of data needs gigantic disk space. Thus a compression technique is necessary to reduce space storage. In addition, medical images must be stored without any loss of information since the fidelity of images is critical in diagnosis. This requires lossless compression techniques. Lossless compression is an error-free compression. The decompressed image is the same as the original image.

Classical image compression techniques (see [1–5]) concentrate on how to reduce the redundancies presented in an individual image. These compression techniques use the same model of compression as shown in Figure 1. This model ignores an additional type of redundancy that exists in sets of similar images, the "set redundancy."

The term "set redundancy" was introduced for the first time by Karadimitriou [6] and defined as follows: "*Set redundancy is the interimage redundancy that exists in a set of similar images, and refers to the common information found in more than one image in the set.*" The compression techniques based on set redundancy follow the model presented in Figure 2. These methods are referred to as SRC (for set redundancy compression) methods. After extracting the set redundancy, any compression algorithm can be applied to achieve higher compression ratios.

In this paper, we present an evaluation of the set redundancy compression (SRC) methods combined with different archivers. The SRC methods tested are the Min-Max differential method (MMD), the Min-Max predictive (MMP) method, and centroid method. The archivers used for individual compression are RAR compressor which is based on [7–9], Gzip which is a variation of Ziv-Lempel (1977) [9] method, Bzip2 that uses Ziv-Lempel (1978) [10] method, and the ZIP archiver. The Huffman encoder [7] is also used in the evaluation.

This paper is organized as follows. We define, in Section 2, the correlation coefficient to quantify similarity between images. The different SRC methods are explained in Section 3. We present, in Section 4, a new predicting scheme for the Min-Max predictive method. Experimental results on medical CT (computed tomography) and MR (magnetic resonance) brain images are given in Section 5. Section 6 gives conclusions.

2. IMAGES SIMILARITY

The redundancy extraction is a worth operation if the images in the set are similar. The visual impression is not sufficient to state that two or more images are similar. We must have a statistical criterion to test similarity. Two images are said to be similar or statistically correlated if they have similar pixel intensities in the same areas or they have comparable histograms.

The correlation coefficient is used to quantify similarity. For two datasets $X = (x_1, x_2, \ldots, x_N)$ and $Y = (y_1, y_2, \ldots, y_N)$ with mean values x_m and y_m, Neter et al. [11] defined this coefficient as

$$r = \frac{\sum_{i=1}^{N} (x_i - x_m)(y_i - y_m)}{\sqrt{\sum_{i=1}^{N} (x_i - x_m)^2} \sqrt{\sum_{i=1}^{N} (y_i - y_m)^2}}. \quad (1)$$

The correlation coefficient is also called Person's r. To avoid the manipulation of negative values, r^2 is often used

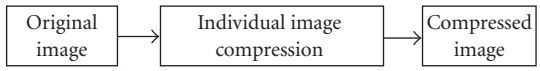


Figure 1: Standard compression model.

instead of r. For two datasets X and Y, a value of r^2 close to 0 means that no correlation exists between them. A value of r^2 close to 1 means that strong correlation exists between the two datasets. X and Y are perfectly correlated if $r^2 = 1$. In context of images, a value r^2 close to 0 means that the two images are totally dissimilar, a value r^2 close to 1 indicates "strong" similarity, and a value $r^2 = 1$ means that the images are identical.

We give two examples to test the existence of correlation among images. Figure 3 shows two successive MRI brain scans of the same patient. The value $r^2 = 0.80$ indicates strong similarity between these two images. Figure 4 depicts two nonsimilar images. The correlation parameter $r^2 = 0.005$ indicates that the two images are noncorrelated.

3. SET REDUNDANCY METHODS

In this section we present four types of SRC methods: the Min-Max differential method [6, 12], the Min-Max predictive method [6, 13], the centroid method [6, 14], and the multilevel centroid method [15]. These methods are fast, lossless, and easy to implement.

3.1. Min-Max differential method

MMD uses, for extracting the "set redundancy" in a set of similar images, two images: a maximum image and a minimum image. To create the minimum (MIN) image, the pixel values across all the images in the set are compared, and for each pixel position the smallest value is chosen. Similarly, the maximum (MAX) image is created by selecting the largest pixel value for each pixel position. Then, the set redundancy can be reduced by replacing every image in the set by its differences from the min or the max image, such that for every pixel position, MMD finds and stores the smallest difference value (see Figure 5). Note that pixel values are indexed with only one subscript, despite corresponding to a two-dimensional array. The image is observed pixel by pixel in a predefined raster scan order.

The algorithms of both encoder and decoder are presented below. For each pixel at position i:

(1) *encoder*:

$$D_i = \begin{cases} \text{value}\,(P_i) - \min_i & \text{if } (\text{value}\,(P_i) - \min_i) \\ & \quad < (\max_i - \text{value}\,(P_i)), \\ \max_i - \text{value}\,(P_i) & \text{otherwise;} \end{cases} \tag{2}$$

(2) *decoder*:

$$\text{value}\,(P_i) = \begin{cases} D_i + \min_i & \text{if } (\text{value}\,(P_i) - \min_i \\ & \quad < \max_i - \text{value}\,(P_i)); \\ \max_i - D_i & \text{otherwise,} \end{cases} \tag{3}$$

where D_i, is the difference value to be stored in the difference image, $\min_i$ is the value at position i in the MIN image and $\max_i$ is the value at position i in the MAX image.

To synchronize encoding and decoding, the encoder uses consistently Min or Max curves until it finds a difference value larger than $(\max - \min)/2$. In that case, it encodes this value and switches to the other curve. The decoder follows the same rule; when it finds a difference larger than $(\max - \min)/2$, it also switches to the other curve.

3.2. Min-Max predictive method

The MMP method also uses the Min and Max images. It is more elaborated than the MMD method but it is also a more powerful method. For each pixel at position i, the MIN image provides the minimal value $\min_i$ of all the images, and the image MAX provides the maximum value $\max_i$. These two values are the limits of the range of the possible values that a pixel at position i can have in each image in the set. After dividing this interval into N levels, a pixel at position i in each image can be represented as a level L_i between its corresponding minimum and maximum values (see Figure 6). The level L_i is given by the equation

$$L_i = N\left(\frac{\text{Value}\,(P_i) - \min_i}{\max_i - \min_i}\right), \tag{4}$$

where L_i is the level of a pixel at position i in a given image, and N is number of levels ($N = 256$).

Neighboring pixels often have similar levels despite having different values. For example, consider the values of the following neighboring pixels given in Table 1.

From (4), a prediction scheme for the value of pixel P_i can be defined as

$$\text{value} - \text{predicted}(P_i) = \min_i + \frac{L'_i}{N}(\max_i - \min_i), \tag{5}$$

where L'_i is the level predicted for a pixel at position i.

The prediction concerns only the element L'_i in the preceding formula. The MMP method predicts the value of a pixel P_i by using the level information from already treated neighboring pixels. Since the levels of neighboring pixels are often similar, this is a good prediction scheme.

Karadimitriou [6, 13] defined three predictors. These predictors determine three variations of Min-Max predictive methods referred to as MMP1, MMP2, and MMP3. The predictions schemes for MMP methods are shown in Table 2.

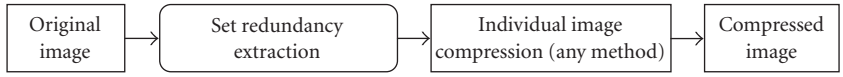


Figure 2: Enhanced compression model.

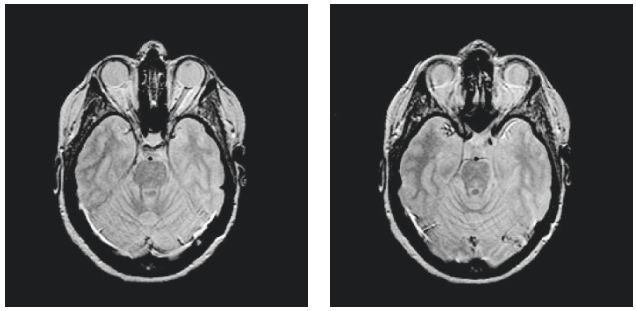


Figure 3: Two successive MRI brain scans.

Figure 4: Two dissimilar images.

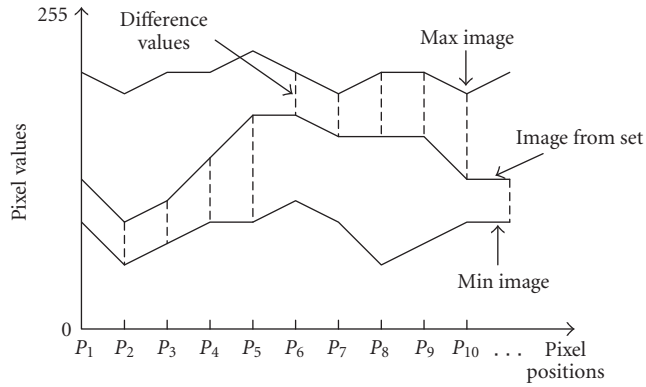


Figure 5: Min-Max differential method.

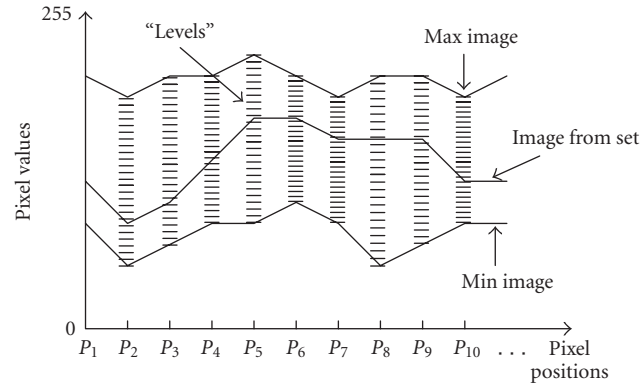


Figure 6: Min-Max predictive method (20 levels).

L_{upper} is the level of the upper neighboring pixel, L_{left} is the level of the left neighbor, and $L_{\text{upperleft}}$ is the level of the upper left neighbor (see Figure 7).

For every image in the set, the encoding process consists of storing the differences between the predicted values and the original values. These differences values replace the original values. To restore the original image from the differences stored, the decoding process calculate the predicted values, and then adds the corresponding differences values.

3.3. Centroid method

The "centroid" method [6, 14] (which is also used in [16]), uses the average image of a set of similar images to predict the values of the difference image. If the prediction is efficient enough, the difference image will contain small values having a Laplacian distribution with most of values very close to zero.

A simple scheme for predicting the pixel value at position i in image j is

$$F_{i,j} = m_i, \tag{6}$$

where m_i is the average value at position i across all images and $F_{i,j}$ is the predicted value. This scheme is not very efficient. A more sophisticated scheme [14] can be expressed as follows:

$$\begin{aligned} F_{i+1,j} &= m_{i+1} + x_{i,j} - m_i, \\ D_{i+1,j} &= x_{i+1,j} - F_{i+1,j}, \end{aligned} \tag{7}$$

where $F_{i+1,j}$ is the predicted value at position $i+1$, $X_{i,j}$ is the pixel value at position i, m_i is the average value of position i across all images, and $D_{i+1,j}$ is the difference value of position $i+1$ in image j between the original and the predicted values. The detailed demonstration of (7) can be found in [6].

Table 1: Example of neighboring pixels levels.

Pixel value	Min value	Maximum value	Level
99	15	197	118
105	21	205	117
112	29	210	117
102	19	199	118

Table 2: Level prediction in MMP methods.

MMP method	Level prediction
MMP1	$L'_i = L_{\text{left}}$
MMP2	$L'_i = (L_{\text{upper}} + L_{\text{left}})/2$
MMP3	$L'_i = L_{\text{upper}} + L_{\text{left}} - L_{\text{upperleft}}$

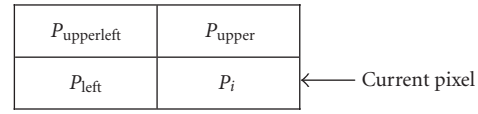


Figure 7: Notation used for specifying neighboring pixels of current pixel *Pi*.

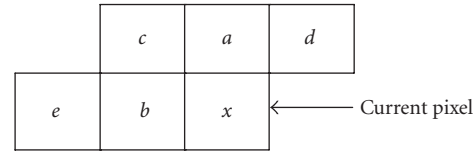


Figure 8: Notation used for specifying neighboring pixels of current pixel *x*.

3.4. Multilevel centroid method

Proposed by El-Sonbaty et al. [15] and derived from the centroid method, this model executes the centroid method N levels times. Given a set of similar images X, the corresponding median image (median_1) is calculated. Applying the centroid method on the given input set, the difference_1 set (difference images at level 1) is obtained. Repeating the process recursively, the median_2 is obtained from the difference_1 set and applying centroid method again, the difference_2 set is also obtained. The process stops when all levels are processed. The first level is the centroid method. The prediction scheme of this method is the same as the centroid method, and is given by

$$\begin{aligned} F_{i+1,j}(n) &= m_{i+1}(n) + x_{i,j}(n) - m_i(n), \\ D_{i+1,j}(n) &= x_{i+1,j}(n) - F_{i+1,j}(n), \end{aligned} \tag{8}$$

where $F_{i+1,j}(n)$ is the estimation of a pixel at position $i+1$ in an image j at level n, $x_{i,j}(n)$ is the value of pixel i of the image j at level n, $m_i(n)$ is the value of pixel i of the median image at level n, and $D_{i+1,j}(n)$ is the value of pixel i of the difference image j at level n.

4. THE NEW MMP PREDICTIVE SCHEME

The three predictors used by Karadimitriou [6, 13] by assigning to L'_i (see Section 3.2) information from previous treated pixels are "not flexible." We propose to use a more elaborated predicting scheme. This scheme is based on the predictor used in Weinberger et al. proposal, LOCO-I (low complexity lossless compression for Images) [17]. LOCO-I uses a nonlinear predictor with edge detecting capability. It guesses the value of the current pixel x based on neighboring pixels (see Figure 8).

The approach in LOCO-I consists in performing a primitive test to detect vertical or horizontal edges. If an edge is not detected, then the guessed value is $a + b - c$. Specifically, the LOCO-I predictor guesses

$$\text{predicted } x = \begin{cases} \min(a,b) & \text{if } c \geq \max(a,b), \\ \max(a,b) & \text{if } c \leq \min(a,b), \\ a+b-c & \text{otherwise.} \end{cases} \tag{9}$$

LOCO-I is the algorithm at the core of the ISO/ITU/14495-1 standard for compression of continuous-tone images, JPEG-LS (see [18]). The guessed value is seen as the *median* of three fixed predictors a, b, and $a+b-c$. The predictor used in LOCO-I was renamed during the standardization process "median edge detector" (MED).

From the MED predictor we derive a new predicting scheme. In (5), the predicted term L'_i will be calculated as follows:

$$L'_i = \begin{cases} \min\left(L_{\text{upper}}, L_{\text{left}}\right) & \text{if } L_{\text{upperleft}} \geq \max\left(L_{\text{upper}}, L_{\text{left}}\right), \\ \max\left(L_{\text{upper}}, L_{\text{left}}\right) & \text{if } L_{\text{upperleft}} \leq \min\left(L_{\text{upper}}, L_{\text{left}}\right), \\ L_{\text{upper}} + L_{\text{left}} & \\ \qquad -L_{\text{upperleft}} & \text{otherwise,} \end{cases} \tag{10}$$

where L_{upper} is the level of the upper neighboring pixel, L_{left} is the level of the left neighbor, and $L_{\text{upperleft}}$ is the level of the upper left neighbor.

Since the image is processed pixel by pixel in a raster scan order, pixels of the first line do not have upper left or upper neighbors. In this case, the value L_{left} will be assigned to L'_i. Similarly, the value L_{upper} will be assigned to L'_i for pixels of the first column in the image. Note that for the first pixel of every image (no processed pixels yet), the value 128 is chosen to be the predicted level.

The idea behind the use of the new predictor is to expect better results than those obtained by using predictors defined

in Section 3.2. We call the new method resulting from this predicting scheme MMPM for MMP MED.

5. EXPERIMENTAL RESULTS

The evaluation of set redundancy method is made on sample medical images. The images were taken from "M.D. Anderson Cancer Center in Houston, Texas" and "Harvard Medical School." All images were gray-level, and were scaled to 8 bits/pixel. All experiments were performed under Windows XP operating system.

To make the evaluation of the SRC methods, we have used the standard compression algorithms RAR, Bzip2, Gzip, ZIP, Huffman. The medical images are compressed by these algorithms with and without using the set redundancy extraction. Each algorithm is tested separately and the attained compression ratios are compared. The compression ratio is given by

$$R = \frac{\text{Size(original image)}}{\text{Size(compressed image)}}. \tag{11}$$

The improvement against standard compression method is also needed in the evaluation. It shows if the use of SRC methods is really effective. The improvement in compression is defined by

$$A = \frac{R_{\text{SRC}} - R}{R}, \tag{12}$$

where R is the compression ratio achieved when using a standard compression method only, and R_{SRC} is the compression ratio achieved when combining SRC with that standard compression method.

5.1. M.D. Anderson Cancer Center images

From M.D. Anderson Cancer Center images, a set of 10 CT (computed tomography) similar images, and another set of 10 MR images are chosen to conduct the first tests. These two sets were selected and used by Karadimitriou [6, 12–14] and also used by Sonbaty et al. [15], so an easy comparison can be made. The resolution is 512×512 for the CT images and 256×256 for the MR images.

5.1.1. CT experiments

The sample set of computed tomography images used in the experiments is shown in Figure 9. The set contains axial CT brain scans where horizontal slices of the brain at the eye-level are depicted. The images were selected from patients of both sexes, various ages, and a variety of pathological conditions.

From the chosen set, the "average," "minimum," and "maximum" images were created to be used in the MMD, MMP, and centroid methods. These three images are shown in Figure 10.

Results of tests on CT images (compression ratios and improvement in compression by using SRC methods) are presented in Table 3. The histograms representing

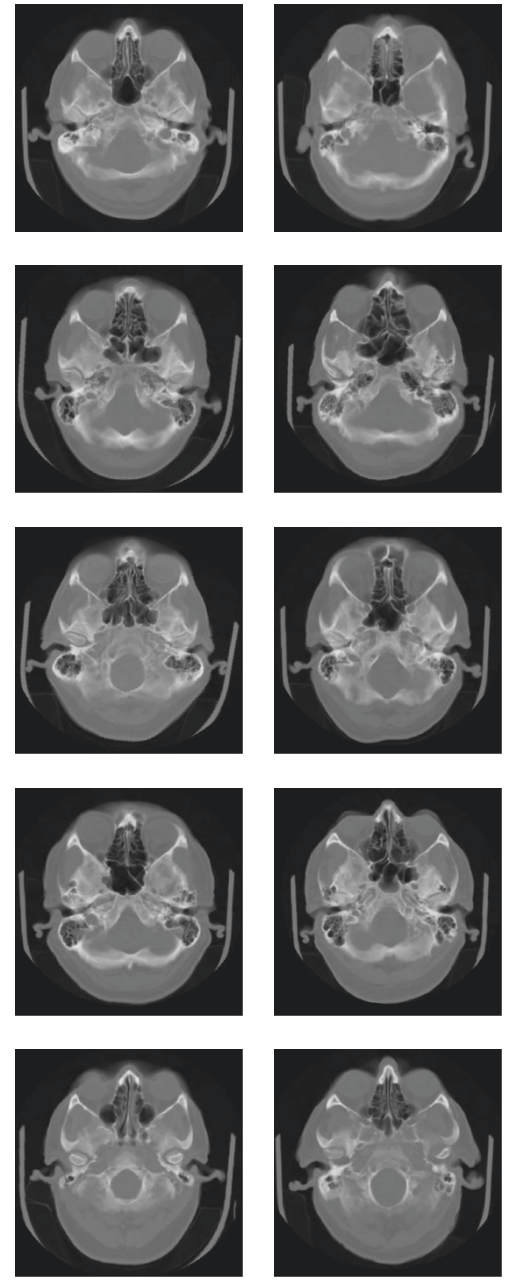
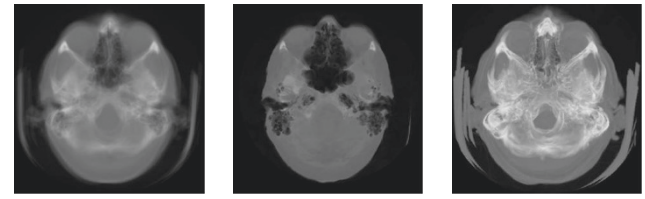


Figure 9: CT test images.



(a) Average CT image. (b) Minimum CT image. (c) Maximum CT image.

Figure 10: CT average, minimum, and maximum images.

Table 3: Experimental results on CT images.

Compression technique	Average size (KO)	Average compression ratio	Improvement %
Original image	256	—	—
Bzip2	74.35	3.44 : 1	—
Centroid + Bzip2	72.55	3.52 : 1	2
MMD + Bzip2	75.78	3.37 : 1	−2
MMP1 + Bzip2	71.71	3.57 : 1	3
MMP2 + Bzip2	64.64	3.96 : 1	15
MMP3 + Bzip2	63.35	4.02 : 1	17
MMPM + Bzip2	61.92	4.13 : 1	20
Mutlilevel centroid (2 levels) + Bzip2	83.32	3.07 : 1	−10
Gzip	100.46	2.54 : 1	—
Centroid + Gzip	82.48	3.10 : 1	22
Gzip + MMD	88.71	2.88 : 1	13
MMP1 + Gzip	78.17	3.27 : 1	28
MMP2 + Gzip	70.92	3.61 : 1	42
MMP3 + Gzip	69.08	3.70 : 1	45
MMPM + Gzip	67.64	3.78 : 1	49
Mutlilevel centroid (2 levels) + Gzip	89.82	2.85 : 1	12
Huffman	193.45	1.32 : 1	—
Centroid + Huffman	98.41	2.60 : 1	96
MMD + Huffman	125.93	2.03 : 1	54
MMP1 + Huffman	84.08	3.04 : 1	130
MMP2 + Huffman	75.35	3.39 : 1	156
MMP3 + Huffman	69.15	3.70 : 1	180
MMPM + Huffman	69.06	3.71 : 1	181
Mutlilevel centroid (2 levels) + Huffman	91.31	2.80 : 1	112
RAR	76.09	3.36 : 1	—
Centroid + RAR	72.60	3.52 : 1	4
MMD + RAR	82.52	3.10 : 1	−7
MMP1 + RAR	67.37	3.8 : 1	13
MMP2 + RAR	62.73	4.08 : 1	21
MMP3 + RAR	57.37	4.46 : 1	32
MMPM + RAR	56.75	4.51 : 1	34
Mutlilevel centroid (2 levels) + RAR	82.57	3.10 : 1	−7
ZIP	99.94	2.56 : 1	—
Centroid + ZIP	80.47	3.18 : 1	24
MMD + ZIP	87.35	2.93 : 1	14
MMP1 + ZIP	75.94	3.37 : 1	31
MMP2 + ZIP	68.36	3.74 : 1	46
MMP3 + ZIP	66.36	3.85 : 1	50
MMPM + ZIP	64.90	3.94 : 1	54
Mutlilevel centroid (2 levels) + ZIP	88.16	2.90 : 1	13

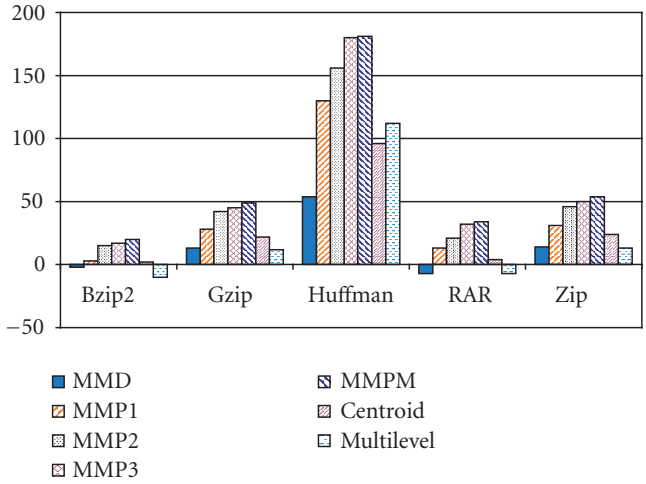


Figure 11: SRC methods improvement on CT images.

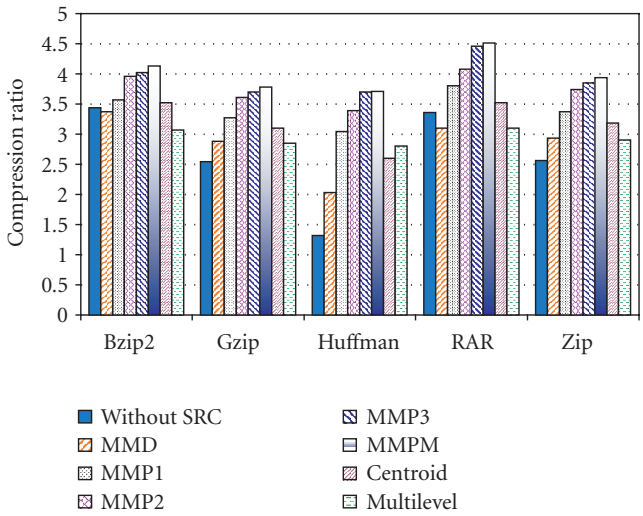


Figure 12: Average compression ratios on CT images.

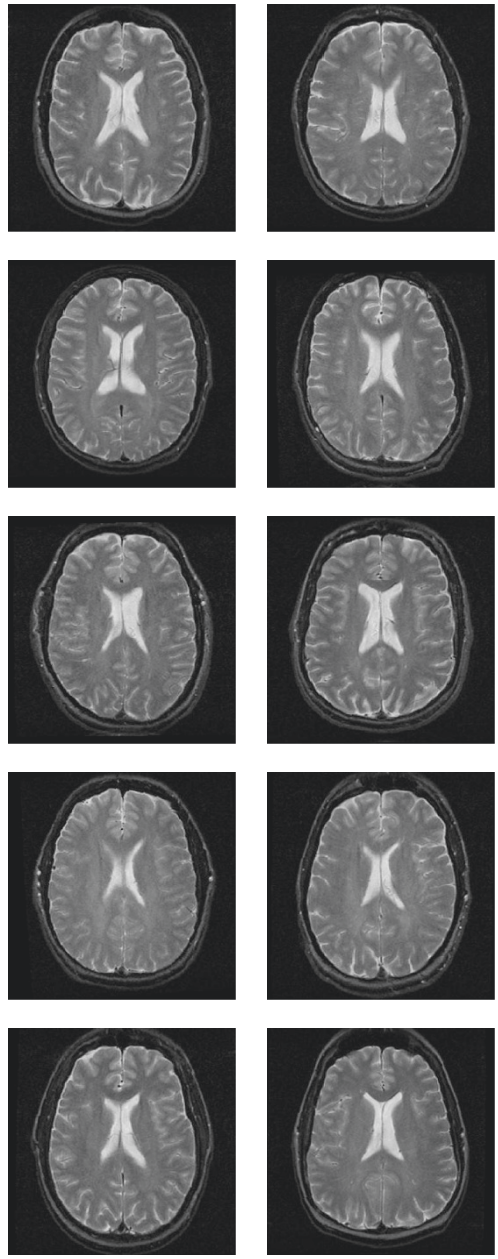


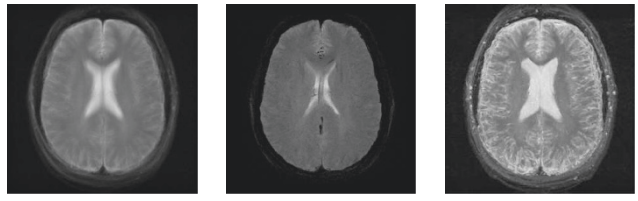
Figure 13: MR test images.

improvements and compression ratios using SRC methods are shown in Figures 11 and 12, respectively.

5.1.2. MR experiments

The set of magnetic resonance images scans depict is horizontal slices about 7-8 cm from the top of the head. These images are shown in Figure 13. From this set, the "average," "minimum," and "maximum" images were created to be used in the MMD, MMP, and centroid methods. These three images are presented in Figure 14.

Results of tests on MR images (compression ratios and improvement in compression by using SRC methods) are presented in Table 4. The histograms representing improvements and compression ratios using SRC methods are shown in Figures 15 and 16, respectively.



(a) Average MR image.

(b) Minimum MR image.

(c) Maximum MR image.

Figure 14: Average, minimum, and maximum MR brain images.

Table 4: Experimental results on MR images.

Compression technique	Average size (KO)	Average compression ratio	Improvement %
Original image	64	—	—
Bzip2	38.25	1.67 : 1	—
Centroid + Bzip2	37.93	1.68 : 1	0.5
MMD + Bzip2	33.13	1.93 : 1	15
MMP1 + Bzip2	33.90	1.88 : 1	12
MMP2 + Bzip2	31.56	2.03 : 1	21
MMP3 + Bzip2	33.59	1.90 : 1	13
MMPM + Bzip2	31.70	2.01 : 1	20
Mutlilevel centroid (2 levels) + Bzip2	41.69	1.53 : 1	−8
Gzip	46.19	1.39 : 1	—
Centroid + Gzip	41.05	1.55 : 1	11
MMD + Gzip	35.31	1.81 : 1	30
MMP1 + Gzip	35.03	1.83 : 1	31
MMP2 + Gzip	33.11	1.93 : 1	39
MMP3 + Gzip	35.02	1.82 : 1	32
MMPM + Gzip	33.10	1.93 : 1	39
Mutlilevel centroid (2 levels) + Gzip	44.09	1.45 : 1	5
Huffman	55.67	1.14 : 1	—
Centroid + Huffman	44.02	1.45 : 1	27
MMD + Huffman	37.12	1.72 : 1	50
MMP1 + Huffman	35.34	1.81 : 1	58
MMP2 + Huffman	32.67	1.95 : 1	71
MMP3 + Huffman	35.17	1.81 : 1	58
MMPM + Huffman	32.48	1.97 : 1	72
Mutlilevel centroid (2 levels) + Huffman	47.48	1.34 : 1	17
RAR	38.22	1.67 : 1	—
Centroid + RAR	36.78	1.74 : 1	4
MMD + RAR	32.10	1.99 : 1	19
MMP1 + RAR	31.94	2.00 : 1	20
MMP2 + RAR	30.52	2.09 : 1	25
MMP3 + RAR	31.65	2.02 : 1	21
MMPM + RAR	29.89	2.14 : 1	28
Mutlilevel centroid (2 levels) + RAR	40.52	1.67 : 1	0
ZIP	46.34	1.38 : 1	—
Centroid + ZIP	41.21	1.55 : 1	12
MMD + ZIP	35.34	1.81 : 1	31
MMP1 + ZIP	35.11	1.82 : 1	32
MMP2 + ZIP	33.21	1.92 : 1	39
MMP3 + ZIP	35.13	1.82 : 1	32
MMPM + ZIP	33.20	1.93 : 1	40
Mutlilevel centroid (2 levels) + ZIP	44.25	1.44 : 1	5

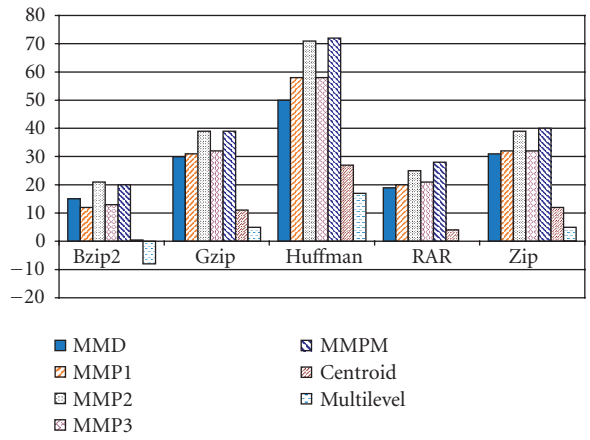


Figure 15: SRC methods improvement on MR images.

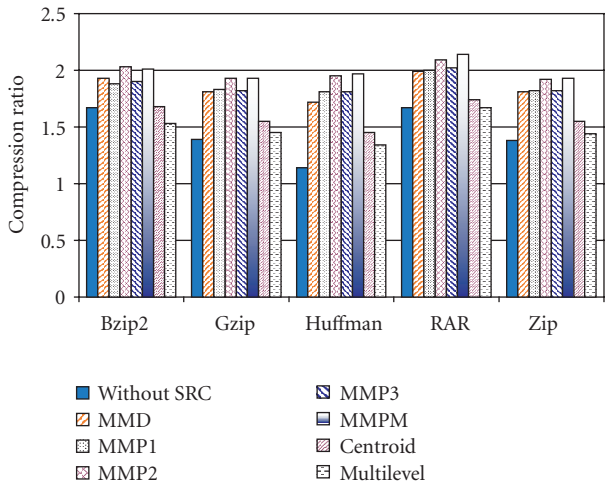


Figure 16: Average compression ratios on MR images.

5.2. Harvard Medical School images

From Harvard Medical School images, two sets of 20 and 30 magnetic resonance images are chosen to make the evaluation. These images are taken from the "whole brain atlas" which depicts various brain diseases. The resolution is 256×256 for all images. The images were converted to PGM format before being processed.

5.2.1. Cerebral edema images

A sample set of medical images is shown in Figure 17. This set contains 20 axial MR brain scans. These images were selected from an MR brain exam of a 51-year old woman. The undertaken exam shows a cerebral edema which corresponds to the high signal extending from the center of the mass through surrounding white matter.

The compression ratios attained on this set by using SRC methods are presented in Table 5. The histogram representing these compression ratios is shown in Figure 18.

5.2.2. Brain tumor images

The set, shown in Figure 19, contains 30 axial MR brain scans. These images were selected from an MR brain exam of a 73-year old right-handed man that sought medical attention because of a grand mal seizure and progressive difficulty with speech. The exam indicates the presence of a brain tumor.

The compression ratios attained on this set by using SRC methods are presented in Table 6. The histogram representing these compression ratios is shown in Figure 20.

5.3. Discussion

From the results shown in the previous tables on sample datasets, we see that the majority of SRC methods carry out an improvement compared to standard compression. This is a good indicator for the effectiveness of using SRC techniques on similar images datasets. The results show that, in most cases, the MMP methods perform better than the other SRC techniques. We also note that the proposed MMPM method attains compression ratios slightly better than the other MMP methods.

The tests have also shown that the centroid and multilevel centroid techniques are not very efficient and that the Huffman encoder gives the worst compression ratios comparatively to other encoders when the number of images in the set grows.

6. CONCLUSION

One of the best application areas for SRC methods is medical imaging. Medical image databases usually store huge amount of similar images (CT, MR, PET, Ultrasound, X-Ray, and Angiography images); therefore, they contain large amounts of set redundancy. This paper attempts to evaluate the performance of various SRC methods on sample datasets of grayscale similar images taken from different sources. An SRC method, called MMPM, is also proposed. It is based on the MED predictor of the JPEG-LS method. In the carried out tests, MMPM performs slightly better than the other MMP methods.

We must mention that, to be effective, the SRC methods impose high similarity in the whole set of images. A preprocessing phase can be done to cluster similar images before launching the compression operation.

In this study, only the effect of compressing sets of grayscale images was evaluated. Further works must consider compressing sets of multispectral or true color images.

SRC methods can also be tested on many other application areas. Satellite image databases, for example, often contain sets of images taken over the same geographical areas, and under similar weather or lighting conditions. They necessarily contain interimage redundancy.

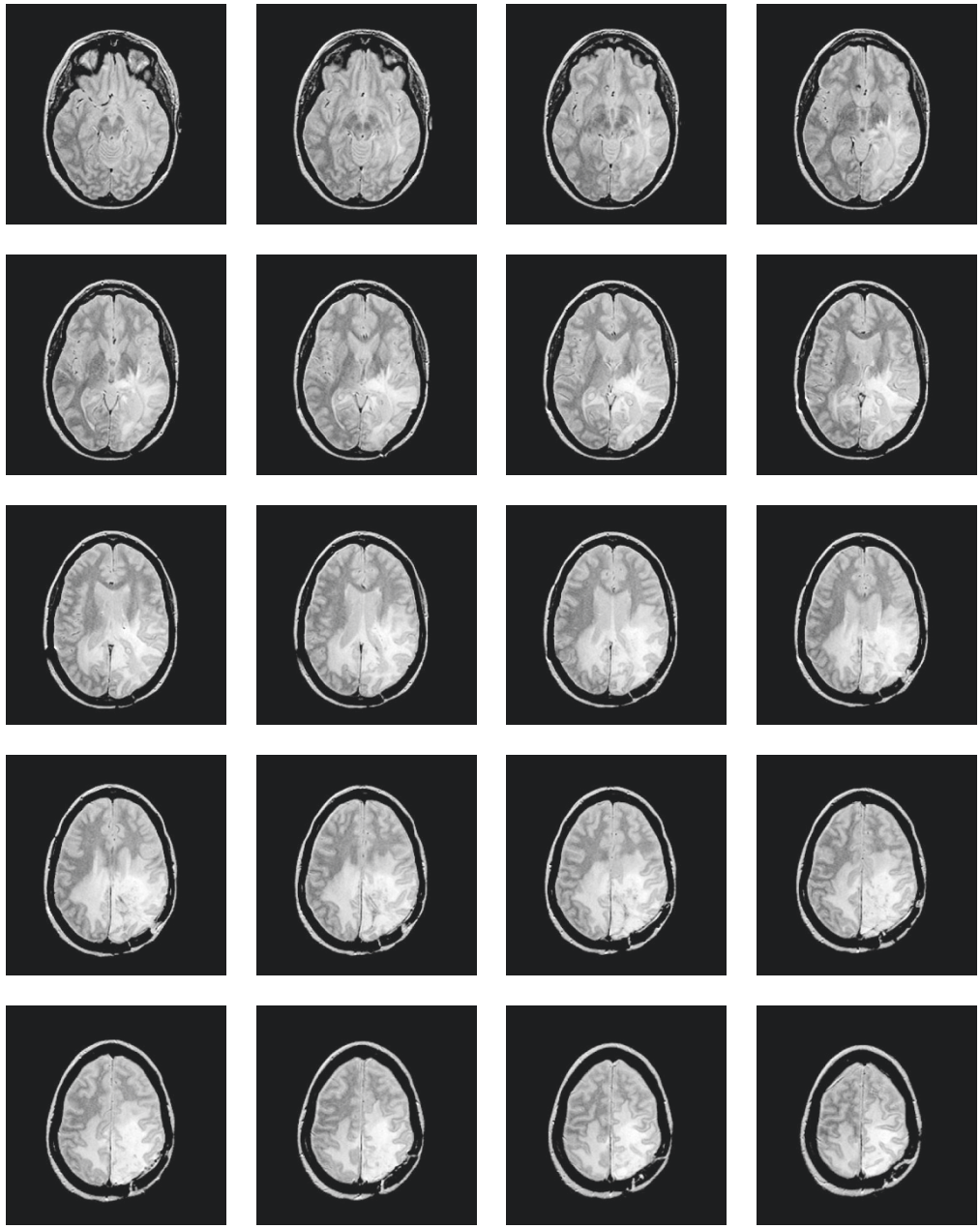


Figure 17: MR brain scans.

Table 5: Average compression ratios on MR images.

	Without SRC	MMD	MMP1	MMP2	MMP3	MMPM	Centroid	Multilevel
Bzip2	4.25	3.95	4.02	4.18	4.08	4.26	3.53	3.38
Gzip	3.63	4.11	4.30	4.38	4.35	4.50	3.45	3.33
Huffman	2.38	2.68	2.86	2.97	2.81	3.02	2.40	2.36
RAR	4.17	4.04	4.20	4.35	4.27	4.47	3.60	3.40
Zip	3.64	3.75	3.99	4.06	4.01	4.15	3.40	3.32

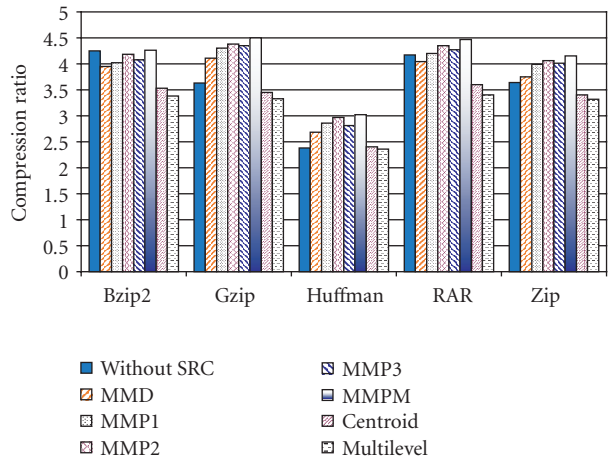


Figure 18: Average compression ratios on MR images.

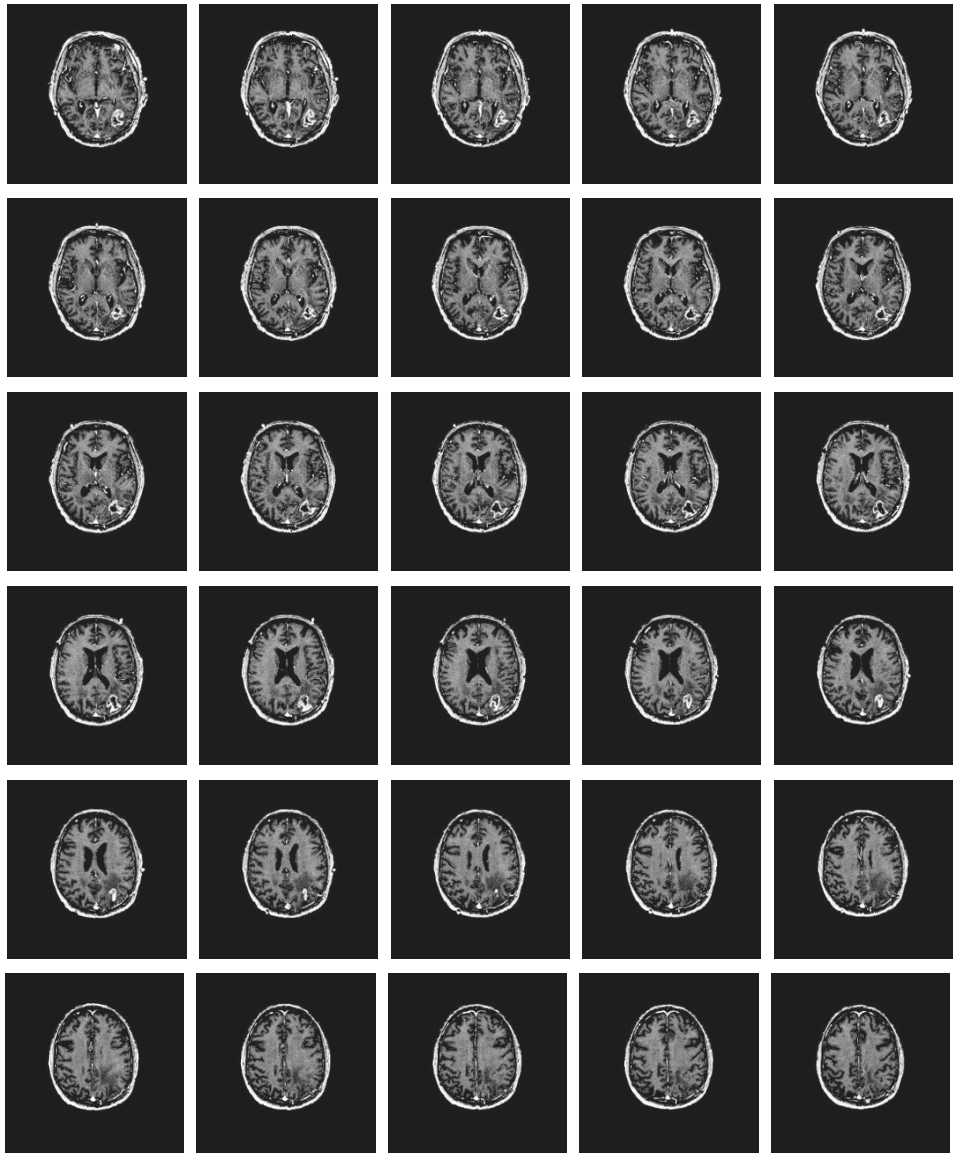


Figure 19: MR brain scans.

Table 6: Average compression ratios on MR images.

	Without SRC	MMD	MMP1	MMP2	MMP3	MMPM	Centroid	Multilevel
Bzip2	5.37	5.12	5.34	5.51	5.39	5.56	4.85	4.66
Gzip	4.97	5.12	5.57	5.66	5.60	5.78	4.52	4.75
Huffman	2.99	3.14	3.47	3.55	3.44	3.56	3.21	3.15
RAR	4.98	5.02	5.23	5.52	5.41	5.60	4.88	4.73
Zip	4.97	5.05	5.48	5.60	5.46	5.68	4.91	4.74

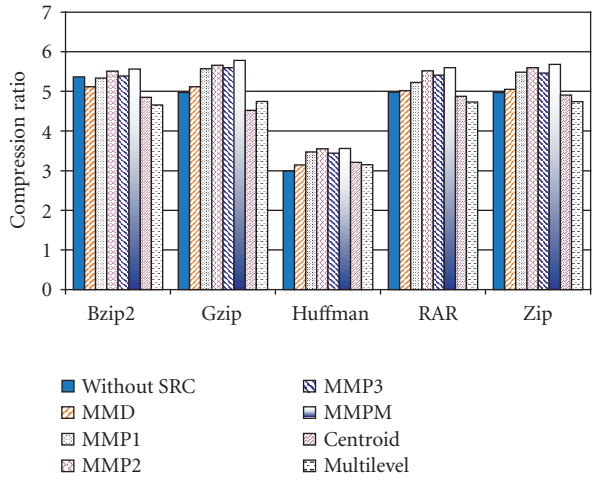


Figure 20: Average compression ratios on MR images.

ACKNOWLEDGMENT

We would like to thank Kosmas Karadimiriou for the helpful discussions and Keith A. Johnson from Harvard Medical School for granting the use of the "whole brain atlas" medical images in the tests.

REFERENCES

[1] H. Bekkouche and M. Barret, "Adaptive multiresolution decomposition: application to lossless image compression," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '02)*, Orlando, Fla, USA, May 2002.

[2] M. U. Celik, G. Sharma, and A. M. Tekalp, "Gray-level embedded lossless image compression," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '03)*, pp. III-245–III-248, Hong Kong, April 2003.

[3] C. C. Chang and G. I. Chen, "Enhancement algorithm for nonlinear context-based predictors," *IEE Proceedings - Vision, Image, and Signal Processing*, vol. 150, no. 1, pp. 15–19, 2003.

[4] D. A. Clunie, "Lossless compression of grayscale medical images: effectiveness of traditional and state-of-the-art approaches," in *Medical Imaging 2000: PACS Design and Evaluation: Engineering and Clinical Issues*, vol. 3980 of *Proceedings of SPIE*, pp. 74–84, San Diego, Calif, USA, February 2000.

[5] J. Jiang, B. Guo, and S. Y. Yang, "Revisiting the JPEG-LS prediction scheme," *IEE Proceedings: Vision, Image and Signal Processing*, vol. 147, no. 6, pp. 575–580, 2000.

[6] K. Karadimitriou, *Set redundancy, the enhanced compression model, and methods for compressing sets of similar images*, Ph.D. thesis, Department of Computer Science, Louisiana State University, Baton Rouge, La, USA, August 1996.

[7] D. A. Huffman, "A method for the construction of minimum redundancy codes," *Proceedings of IRE*, vol. 40, no. 9, pp. 1098–1101, 1952.

[8] D. Shkarin, "Improving the efficiency of PPM algorithm," *Problems of Information Transmission*, vol. 37, no. 3, pp. 226–235(10), 2001.

[9] J. Ziv and A. Lempel, "A universal algorithm for sequential data compression," *IEEE Transactions on Information Theory*, vol. 23, no. 3, pp. 337–343, 1977.

[10] J. Ziv and A. Lempel, "Compression of individual sequences via variable-rate coding," *IEEE Transactions on Information Theory*, vol. 24, no. 5, pp. 530–536, 1978.

[11] J. Neter, W. Wasserman, and M. H. Kutner, *Applied Linear Regression Models*, IRWIN, Burr Ridge, Ill, USA, 1989.

[12] K. Karadimitriou and J. M. Tyler, "The min-max differential method for large-scale storage and compression of medical images," in *Proceedings of of Annual Molecular Biology and Biotechnology Conference*, Baton Rouge, La, USA, 1996.

[13] K. Karadimitriou and J. M. Tyler, "Min-max compression methods for medical image databases," *ACM SIGMOD Record*, vol. 26, no. 1, pp. 47–52, 1997.

[14] K. Karadimitriou and J. M. Tyler, "The Centroid method for compressing sets of similar images," *Pattern Recognition Letters*, vol. 19, no. 7, pp. 585–593, 1998.

[15] Y. El-Sonbaty, M. Hamza, and G. Basily, "Compressing sets of similar medical images using multilevel centroid technique," in *Processing of the 7th Conference on Digital Image Computing, Techniques and Applications*, C. Sun, H. Talbot, S. Ourselin, and T. Adriaansen, Eds., Sydney, Australia, December 2003.

[16] J. D. Lee, S. Y. Wan, and R. F. Wu, "A hybrid compression model for clusters of similar medical images," *Biomedical Engineering - Aplications, Basis & Communications*, vol. 16, no. 1, 2003.

[17] M. J. Weinberger, G. Seroussi, and G. Sapiro, "LOCO-I: a low complexity, context-based, lossless image compression algorithm," in *Proceedings of the IEEE Data Compression Conference*, Snowbird, Utah, USA, April 1996, ISO Working Document ISO/IEC JTC1/SC29/WG1 N203.

[18] M. J. Weinberger, G. Seroussi, and G. Sapiro, "The LOCO-I lossless image compression algorithm: principles and standardization into JPEG-LS," *IEEE Transactions on Image Processing*, vol. 9, no. 8, pp. 1309–1324, 2000.

Samy Ait-Aoudia received a DEA "Diplôme d'Etudes Approfondies" in image processing from Saint-Etienne University, France, in 1990. He had a Ph.D. degree in computer science from Ecole des Mines, Saint-Etienne, France, in 1994. He is currently "Maître de Conférences" at the National Computer Science Institute in Algeria. He teaches different modules at both B.S. and M.S. levels in computer science and

software engineering. His areas of research include image processing, CAD/CAM, and constraints management in solid modeling.

Abdelhalim Gabis received the B.S. degree (Ingénieur d'État en Informatique) from the National Computer Science Institute, Algiers, Algeria, in 2002. He had an M.S. degree from the same institute in 2005. He is currently working toward the Ph.D. degree at the National Computer Science Institute in Algeria. His research interests include data compression as well as image/video coding and processing. He is a Member of the "Perceptions" Research Group supported by the Research Ministry in Algeria.