Adaptive Coding of Images Using DCT and Classification of Blocks for Segmentation

Johelden Campos Bezerra[1] and Evaldo Gonçalves Pelaes[2]
Centro Federal de Educação Tecnológica do Pará – CEFET/PA[1] / Universidade Federal do Pará - UFPA[2]
joe@nautilus.com.br[1] - pelaes@ufpa.br[2]

Abstract - **This article propose a technique of adaptive coding of images, implemented from the classification of the size blocks MxM, with the aid of the technique of segmentation of images for the detection of discontinuities in the image, generating K classes with different levels of discontinuities, containing, each one, a variable number of blocks by class. The bits allocation table in this method is determined by the log-variance rule, using only one block of size K of sub-blocks MxM, each one belonging to a class. The method is applied using the DCT, with the uniform, laplacian and gaussian quantization of Lloyd-Max, and with blocks of size 8x8. The proposed method was applied in several images, like Lena, Barbara, Baboon, Goldhill, presenting a medium gain of 4 dB in the SNR at rate 1 bpp, and a maximum gain of 9,04 dB at rate 5 bpp, when compared with the method of images coding using DCT without classification.**

I. INTRODUCTION

In the last years, due mainly its great application, the appearance of new techniques and new algorithms in the several areas that involve image processing has been growing significantly, such as, representation and modeling of images, analysis, restoration and texture of images, coding, compression and interpolation of images [3],[4]. That work will concentrate basically in areas that refer to the digital coding of images.

The digital coding of images is a procedure of representation of an image as a sequence of bits to store it or to transmit it through a channel, with the main objective of minimizing the number of necessary bits to represent the image with a certain fidelity criterium, preserving the necessary quality and intelligible levels for a given application. That is due mainly to the storage or transmission cost, that it will grow unavoidably with the increase of the number of bits. [2]-[4]

This article presents in the section II a brief description of the digital image coding process by transform. Section III the segmentation process is described which will be used as tool for the classification of the blocks. In the section IV will present the proposed method of adaptive coding, where the results obtained by the simulation come in the section V and it finishes with the conclusions obtained with the results.

II. TRANSFORM IMAGE CODING

The Fig. 1 shows the diagram of blocks of a typical system of images coding by transform. In this system the original image is transformed to a more appropriate domain in order to explore its characteristics, such as the discorrelation between the neighboring pixels and the property of compactation of energy. This procedure is accomplished through of transform known in the literature, such as the discreet coseno transform (DCT) [7]. In general, the transform coefficients concentrate great part of the energy on some coefficients, what leads to an image recovery with a good quality from a reduced number of coefficients.

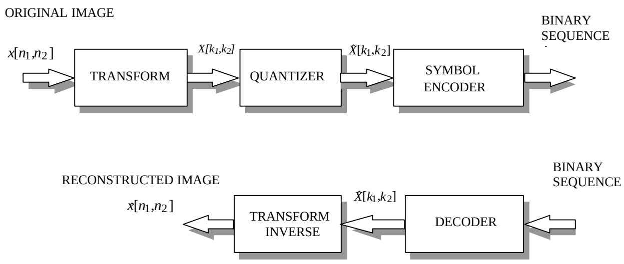


Fig. 1. Diagram of blocks of a coder and decoder of images using transforms.

The reconstructed image is similar to the original one presenting only error due to the quantization, supposing that the channel coding does not to introduce error in the process. This error is usually measured in two ways: a) subjective forms: an observer verifies the original and the reconstructed images after the process, trying to locate possible differences introduced in the image and; b) objectify form: where is calculated the sign-to-noise ratio (SNR) between the original and reconstructed image, that will be used in this work.

III. TECHNIQUE OF IMAGES SEGMENTATION

The techniques of images segmentation, which it is the first step in the image analysis, are used to extraction of information regarding some characteristic of an image subdividing it into its constituent parts or objects. In the present work, a segmentation technique is used as tool to classify the blocks of an image in classes. After the coding of adaptive form is done [4].

Segmentation algorithms for monochrome images generally are based on one of two basic properties of gray levels values: discontinuity and similarity. In this work it was used the principle of the techniques based on discontinuities.

These techniques are defined by a guided operation by neighborhood. Logical and arithmetic operations guided by neighborhood uses the convolution concept with masks. To each relative position of the mask on the image, the central pixel of the subimage in subject will be substituted. In a matrix denominated image destiny, for the sum of products of the coefficients with the gray levels contained in the region encompassed by the mask. That is, the response of the 3x3 mask at any point in the image is:

$$R_n = W_1Z_1 + W_2Z_2 + ... + W_9Z_9 = \sum_{i=1}^{9} w_i z_i \qquad (1)$$

where n = 1, 2,... is the size of the image, Z_i is the gray level of the pixel associated with mask coefficient W_i.

The Fig. 2 illustrate this procedure, hence to run a mask through the image.

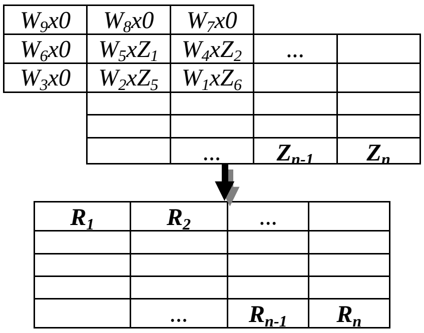


Fig. 2. Convolution process with mask and the result.

To this work was implemented a mask to detection of discontinuities, which is showing in the Table 1.

Is said that a discontinuity has been detected at the location on which the mask is centered if

$$|\boldsymbol{R}| > \boldsymbol{T}$$

where T is a nonnegative threshold, and R is as defined previously. Basically, all that this formulation does is measure the weights differences between the central point and its neighbors. The idea is that the gray level of an point will be quite different from the gray levels of its neighbors.

TABLE 1
MASCK USED TO DETECTION OF DISCONTINUITIES

0	0	0
0	2	-1
0	-1	0

IV. THE PROPOSED METHOD

A. Classification of the blocks using segmentation for the detection of discontinuities

In the technique of transform coding, the images are divided in blocks and from them are applied the transformation, quantization and coding techniques, respectively. To encode an image so that the same presents a better performance, the method of coding adaptive can be used, which can be obtained by the classification of the constituent blocks of the image, so that, the blocks with similar characteristics are classified in a same class, for example, all the blocks of low frequency should belong to a same class [1],[9],[10].

In this work, this method of adaptive classification, it was obtained with the aid of the segmentation technique described in the section III, as will be described to proceed.

Firstly, the original image is segmented, to obtain an image with two gray levels, an equal to zero and another with different value of zero. After, divide the image segmented in size blocks of 8x8 pixels. [8][11]

The classification of the blocks of the original image, is obtained from an analysis in the segmented image. The level of activity of the blocks of the segmented image is verified, through to amounts of pixels different of zero that each block presents, because that amount of pixel different of zero in the segmented image represents an discontinuity existent in the original image, for example, the blocks with high activity (high frequencies) will present several pixels with different value of zero, while the blocks of low activity (low frequency) will present few pixels with different value of zero or none [8][11].

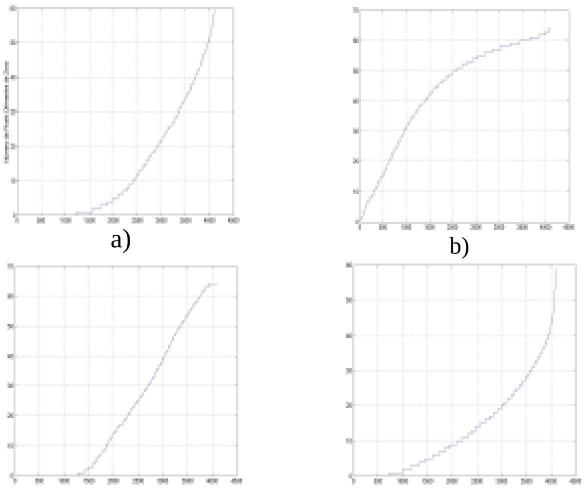
It is known that the size of the original and segmented image is the same, then the classification of the blocks of the images original can be doing, being made an analysis of the corresponding blocks on the segmented image, as illustrated in Fig. 3, that shows the original Lena image and segmented Lena image, which will be used to classify the blocks of the original image.

The blocks of the original image will be classified in four different classes, hence each class has different level of activity, such as, the blocks with similar activities will belong at a same classes, it is possible to have classes with different sizes.

To facilitate the classification of the blocks in classes, the blocks of the segmented image are ordered in growing order, in agreement with the number of pixels with different value of zero [11]. The Fig. 4 shows the graphs of the classified blocks in growing order, for Lena (512x512), Baboon (512x512), Barbara (512x512) and Goldhill (512x512), the graphs represent each block of the image segmented in agreement with the amount of pixels different of zero that the same present, and from them the classification of the blocks can be made, as mentioned previously.

Fig. 3. Original and segmented Lena Image, 512x512 pixels.

The Fig. 5 shows blocks of two different classes obtained from the classification executed in the Lena image (512x512 pixels). The Fig. 5a shows a block belonging to the first class, where are classified the blocks of lowest frequency, and can be observed that the same present little or none variation of gray levels. In Fig. 5b is shown a belonging block the class of the blocks of highest frequency, it means that, to the group of blocks with several variations of the gray levels, hence they are exactly those that store to largest amounts of details of the image, for example, borders or contours.

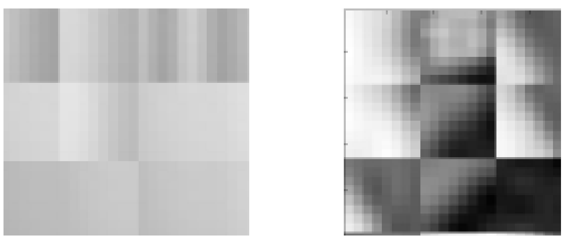


a) b) c) d)

Fig. 4. Representative graph of the relationship between the blocks (horizontal axis) and the number of pixels with different value of zero present in each one (vertical axis), to the a) Lena (512x512), b) Baboon (512x512), c) Barbara (512x512) and d) Goldhill (512x512).

B. Quantization and Bits Allocation

When an image is transformed, the generated coefficients are real numbers, which can assuming an infinity of values inside of a certain interval. Those coefficients need to be represented by a finite number of bits, it means that, they need to be quantized.



a) b)

Fig. 5. examples of blocks: the) low frequency, b) high frequency.

The quantization can be of two types: vector and scalar. The vector quantization was not used in this work.

In the scalar quantization each transform coefficient is represented independently by the number of bits in agreement with the allocation table, that can be established previously, or determined with base in the statistics of the transform coefficients. In this work was used to uniform, laplacian and gaussian scalar quantization.

In general is used as distortion criterion introduced by the quantizador the mean-square error. The quantizador based on the distortion minimum criterion is referred many times how Lloyd-Max Quantizer [3]-[6].

With the transform coefficients of an image of size NxL pixels, we group their in size blocks MxM and classifying their in four classes. The variance of the transformed coefficients is determined for each class. The Fig. 6 illustrate the procedure of grouping of the blocks of a given class for the determination of the variance matrix.

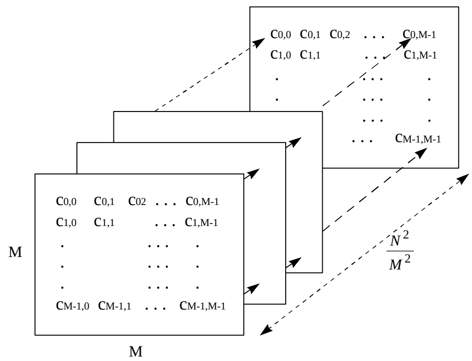


Fig 6. Illustration of the procedure of grouping of the size blocks MxM of a class with Nb blocks.

Each one of the four classes will go present a matrix of size variance MxM representative of the transform coefficients, which will be grouped forming a variance matrix of size 2Mx2M to the image, that will be used to calculate the bits allocation table of the imagems, using the log-variance rule.

After the obtaining of the bits allocation table of size 2Mx2M, the alocation table for the four size classes MxM are then separate and then is made the quantization of the transform coefficients. In this work we used the scalar quantization, then each coefficient is individually quantized in agreement with the number of bits allocated to that position, being obtained the transform and quantized coefficients. The medium coding rate of R, can be expressed as

$$R = (R_1M_1N_1 + R_2M_2N_2 + R_3M_3N_3 + R_4M_4N_4) / NL \quad (8)$$

where R_i, i = 1, 2, 3, 4 are the rates in bpp for each class and M_iN_j, i, j = 1, 2, 3, 4 represents the number of pixels of each class and NL the number of pixels of the original image.

To reconstitute the image, the inverse process is realize, that is, the quantized coefficients are reconstituted by the inverse procedures of quantization and the inverse transform is applied on the reconstituted coefficients, and rearranging the blocks belonging the classes obtained a replies of the original image, that certainly will present some difference due to the coding errors. This difference is measured, in general, by the signal-to-noise ratio (SNR) between the original image and the reconstituted image. The Equation 9 shown as the SNR is calculated.

The Table 2 shown the allocation table obtained using DCT applied to the Lena image, with a medium rate of 1 bpp.

From the Table 2 can be observed that most of the bits allocated is concentrated in the classes in that the blocks present largest variations, it means, the of biggest frequency.

TABLE 2
BITS ALLOCATION TABLE

Class 1	7	3	1	1	0	0	0	0	7	4	2	2	1	1	0	0	Class 2
	2	1	0	0	0	0	0	0	3	2	2	1	1	1	0	0	
	1	0	0	0	0	0	0	0	2	2	1	1	1	0	0	0	
	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	
	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Class 3	7	5	3	3	2	1	1	1	7	6	5	4	3	3	2	2	Class 4
	4	4	3	2	2	1	1	0	5	5	4	4	3	2	2	1	
	3	3	2	2	1	1	1	0	4	4	4	3	3	2	2	1	
	2	2	2	2	1	1	0	0	3	3	3	3	2	2	1	1	
	1	1	1	1	1	1	0	0	2	2	2	2	2	1	1	1	
	1	1	1	1	0	0	0	0	1	1	1	1	1	1	1	1	
	1	0	0	0	1	0	0	0	1	1	1	1	1	1	0	0	
	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	

V. SIMULATION RESULTS

In this section will be presented the results of the application of the method of images coding proposed in the section IV.

In the Tables 3, 4, 5 and 6 are shown the number of blocks by class, the amount of bits per block and the medium rate of bits per pixel, in each class, for the images: Lena (512x512), Baboon (512x512), Barbara (512x512) and Goldhill (512x512), for a medium rate of 1 bpp and using uniform quantization.

Shown in the Fig. 7 the original Lena image (512x512 pixels), that was encoded by the proposed method of adaptive coding, where the results were compared with the results of the coding without the classification. In the Fig. 8 are shown details of the reconstituted image after coding and in the Fig. 9 are shown the graphs for SNR obtained by the coding with classification and without classification.

The method was tested encoding Baboon image (512x512), that is shown in the Fig. 10, which is an image of high frequency, and the results obtained to the SNR are presented in Fig. 11.

TABLE 3

Lena 512X512			
	Number of blocks	Bits per block	Bits per pixel
Class 1	1227	16.00	0.25
Class 2	1218	40.00	0.625
Class 3	396	76.99	1.203
Class 4	1255	129.98	2.031

TABLE 4

Baboon 512x512			
	Number of blocks	Bits per block	Bits per pixel
Class 1	583	11.97	0.187
Class 2	386	24.00	0.375
Class 3	1877	62.98	0.984
Class 4	1250	102.98	1.609

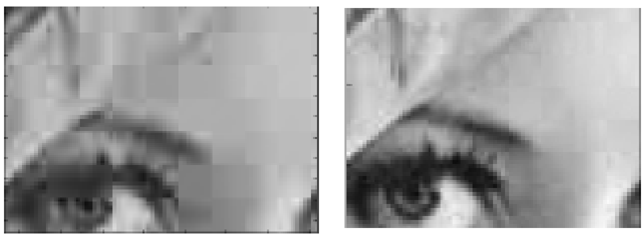
TABLE 5

Barbara 512X512			
	Number of blocks	Bits per block	Bits per pixel
Class 1	845	20.99	0.328
Class 2	1878	65.98	1.031
Class 3	844	108.99	1.703
Class 4	529	139.97	2.187

TABLE 6

Goldhill 512X512			
	Number of blocks	Bits per block	Bits per pixel
Class 1	1176	22.40	0.359
Class 2	1017	49.98	0.781
Class 3	1327	83.97	1.312
Class 4	576	124.99	1.953

Fig. 7. Original Lena image, 512x512 pixels.



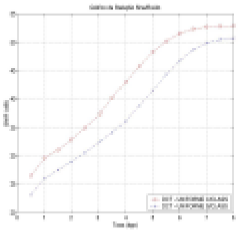
a) b)

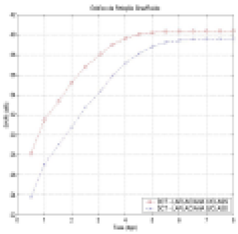
Fig. 8. Details of the Lena encoded image a) without the classification and b) with the classification, by the use of DCT to a medium rate of 1 bpp.

The performance of the coding method with classification was analyzed using Barbara image

(512x512), which is shown in the Fig. 12, where the results obtained for SNR are presented in the Fig. 13.

The classification method was applied in the Goldhill image (512x512), which shown in its original form in the illustration 5.14. The results obtained for SNR are presented in Fig. 15.


a)


b)

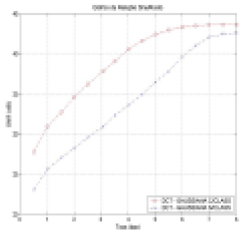

c)

Fig. 9. Graphs of SNR obtained by the Lena image encoded, 512x512 pixels, using DCT with a) uniform, b) laplacian and c) gaussian quantization. " DCT S/CLAS " refers the code without classification.

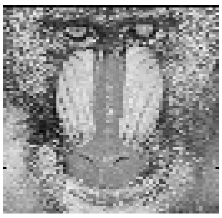
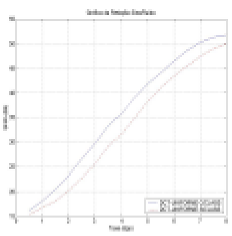
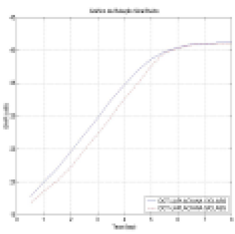


Fig. 10. Original Baboon image, 512x512 pixels.


a)


b)

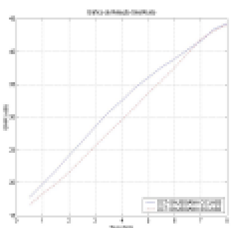
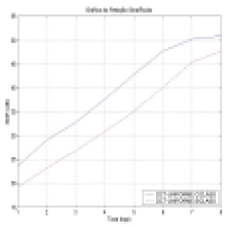
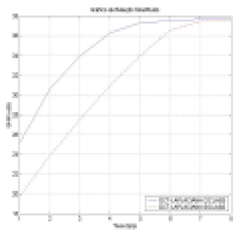
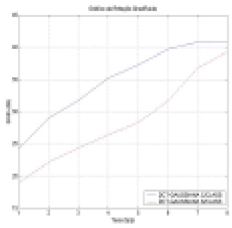

c)

Fig. 11. Graphs of SNR obtained by the Baboon image encoded, 512x512 pixels DCT with a) uniform, b) laplacian and c) gaussian quantization. " DCT S/CLAS " refers the code without classification.

Fig. 12. Original Barbara image, 512x512 pixels.


a)


b)


,

c)

Fig. 13. Graphs of SNR obtained by the Barbara image encoded, 512x512 pixels, using DCT with a) uniform, b) laplacian and c) gaussian quantization. " DCT S/CLAS " refers the code without classification.

Fig. 14. Original Goldhill image, 512x512 pixels.

VI. CONCLUSION

In these work a method of adaptive coding was proposed using a segmentation technique for the discontinuity detection between adjacent pixels belonging to an image to classify the transformed blocks in four different classes, according to the level of activity of the each block.

Was used DCT and scalar quantization. The bits allocation table is determined by the log-variance rule applied to size matrix 16x16, composed by 4 size variance matrix 8x8, belonging to each one of the classes.

For confirmation of the coding method 512 x 512 size images were used and size blocks of 8x8 pixel. The proposed coding method always presented a better peformance when compared with the coding without the classification, in agrement with the results presented in the section V.

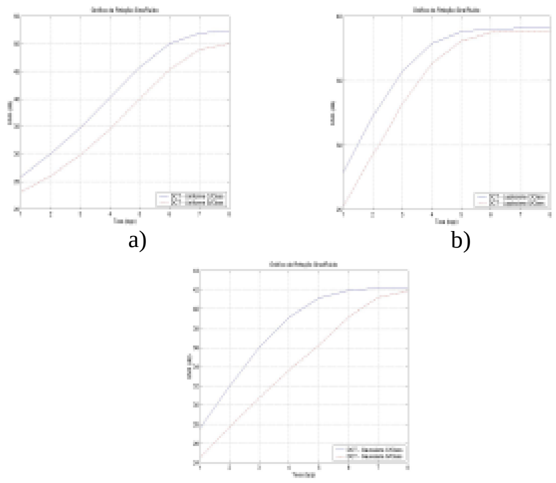


Fig. 15. Graphs of SNR obtained by the Goldhill image encoded, 512x512 pixels, using DCT with a) uniform, b) laplacian and c) gaussian quantization. " DCT S/CLAS " refers the code without classification.

For all the simulated images with the rate of 1bpp, the proposed coding method presented gain superiors in the SNR compared with the method without classification, according to Table 7. The Table 8 showing the maximum gain obtained in the SNR, for the encoded images in different rates.

TABLE 6
GAIN FOR RATE OF 1 BPP WITH GAUSSIAN QUANTIZATION

Image	Gain (dB)
Lena	5.03
Baboon	1.05
Barbara	6.72
Goldhill	4.40

TABLE 6
MAXIMUM GAIN OBTAINED IN THE SNR

Image	Quantization	Rate	Gain (dB)
Lena	uniform	4.5	7.04
Baboon	uniform	4	4.21
Barbara	gaussian	5	9.04
Goldhill	uniform	5	5.76

Therefore, for all the tested images using DCT and uniform, laplacian and gaussian quantization, the proposed method of adaptive coding always presented the best peformance when compared with the method without the classification.

REFERÊNCIAS

[1] W. Chen e C. Smith, “Adaptive Coding of Monochrome and Color Images”, *IEEE Transactions on Communications*, vol. 25, pp. 1285-1292, Nov. 1977.

[2] Ogê Marques e Hugo Vieira. *“Processamento Digital de Imagens”*, Editora Brasport, Rio de Janeiro - RJ, 1999.

[3] RAFAEL GONZALES *“Digital Image Processing”*, Editora Addison-Wesley, 1993.

[4] Jae Lim, *“Two-dimensional Signal Processing and Image Processing”*, Prentice-Hall, 1990.

[5] Evaldo Pelaes, *“Transformada Seno Discreta com Rotação de Eixo Bidimensional (DSTr-2D): Aplicações na Codificação e Interpolação de Imagens”*, Tese de doutorado, Universidade Estadual de Campinas, 1998.

[6] Joel Max, “Quantizing for Minimum Distortion”, *IRE Trans. Inform. Theory*, vol. IT-6, pp. 267-272, March 1960.

[7] R. RAO and P. Yip, *“Discrete Cosine Trasnform”*, Academic Press, Inc, 1990.

[8] Johelden Bezerra, Evaldo Pelaes e Samuel Neves, “Codificação Adaptativa de Imagens usando Classificação por Segmentação”, *XIX SBrT*, Fortaleza-CE, Setembro 2001.

[9] Jainping Pan, “Vector-Scalar Classification for Transform Image Coding”, *IEEE Transactions on Image Processing*, Vol. 8, Nº. 9, Setember 1999.

[10] Yo-Sung Ho and Alien Gersho, “Classified Transform Coding of Images Using Vector Quantization”, *M9.22 IEEE*, pp. 1890-1893, 1989.Wesley, 1993.

[11] Johelden Campos Bezerra, *“Codificação Codificação Adaptativa de Imagens usando Classificação por Segmentação”*, Dissertação de Mestrado, Universidade Federal do Pará, Fevereiro 2001.