JPEG2000 CODED IMAGES OPTIMIZATION USING A CONTENT-DEPENDENT APPROACH

Sebastiano Battiato, Antonio Buemi, Gaetano Impoco and Massimo Mancuso

Abstract — **This paper presents a method to optimize the visual quality of Jpeg2000-compressed images using an adaptive approach to assign more bits to regions in which errors are more visible, maintaining the global bit-rate unchanged. The encoder partitions the image into rectangular blocks called "tiles" to perform the compression. The proposed approach is based on the assumption that, by computing simple energy measures, it is possible to discern the most perceptually significant regions of the image and to code corresponding tiles properly. The effectiveness of the proposed algorithm depends on the accuracy of this classification that, in turn, depends on the tile's size. Experimental results show that the proposed technique yields a perceptive quality improvement of the image with no significant overhead in terms of complexity of the standard Jpeg2000 encoder.**

Index Terms — **Classification, compression standard, content-dependent optimization, EBCOT, HVS, Jpeg2000.**

I. INTRODUCTION

The wide diffusion of digital still and video image capturing devices pointed out new challenges mainly focused in quality improvement and reduction of the processing time [4]. Increasing of digital pictures quality is achieved by reducing noise introduced by the sensor during the acquisition of the data [2], while increasing the resolution of the sensor allows improving compression performances for a more efficient manipulation and storage of images [3, 5]. The high rate of innovation in the image processing technology involves changing in the product features. The new Jpeg2000 compression standard [7, 13, 16] is promising a lot as far as compression efficiency and features to be supported, at the cost of a higher complexity with respect to Jpeg. It has been created to include all functionalities needed in a compression system. These additional functions make it suitable for a wide range of applications and markets: Internet, mobile communication, digital photography, digital library, printing, scanning and so on.

Jpeg2000 works partitioning the input image into rectangular "tiles" that are compressed independently. Differently from standard Jpeg, which decomposes the image into classic 8x8 blocks, in Jpeg2000 standard the tile size is not fixed. Thus, the tile size is arbitrarily chosen by the user, up to include the whole image, while the global compression ratio is fixed assigning the same target bit-rate to every tile.

Several techniques, based on perceptual models exploiting Human Visual System (HVS) properties, were developed to code fewer bits to represent perceptual less important areas using standard compression algorithms [6]. Perkins and Lookabaugh [15] proposed a bit allocation algorithm appropriate for any subband image coding system. This method assigns to each subband a number of bits proportional to the energy content of the subband. It also exploits the available psycophysical data concerning the sensitivity of the human eye to different spatial frequencies. Since the human eye is not equally sensitive to signals at all spatial frequencies, a set of weights is introduced in order to make the algorithm able to distribute the available bits in a perceptually significant manner. A method that fully exploits HVS properties in order to improve Jpeg compression standard performances, has been also proposed in [18] by Tong et alii. Like the algorithms discussed above, this approach uses a perceptual model to capture the perceptually most significant areas in the image, scaling properly the Jpeg quantization table. Wen et alii [19] presented also a processing scheme for image compression based on fuzzy control. This approach takes into account the background illumination levels and spatial sensitivities, using appropriate fuzzy rules.

Masking properties of the HVS are exploited also in Jpeg2000. Daly et alii [8], have developed two techniques that could work together to take advantage from the masking properties of the visual system. Using such properties, the degradation of the image fidelity could be managed as a function of increasing signal energy and an adaptive quantization scheme could be designed in order to allocate more levels to low amplitude coefficients. One approach computes the quantization error for each coefficient, using the information related to the coefficient itself, while the other one uses also coefficients' neighborhood within a wavelet band to obtain an estimation

S. Battiato and A. Buemi are at Catania Advanced System Technology of STMicroelectronics, Stradale Primosole 50, Catania 95121 Italy (e-mail: sebastiano.battiato @st.com, antonio.buemi@st.com).

G. Impoco is at Dipartimento di Informatica, Università di Pisa, Corso Italia 40, Pisa 56100 Italy (e-mail: impoco@di.unipi.it)

M. Mancuso is at Agrate Advanced System Technology of STMicroelectronics, via C. Olivetti 2, Agrate-Brianza Italy (e-mail: massimo.mancuso@st.com).

Manuscript received June 24, 2002

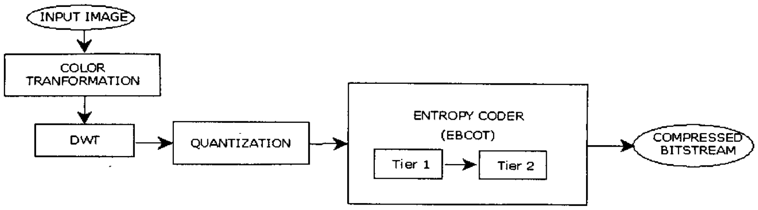


Figure 1: Jpeg2000 encoder basic scheme.

of the quantization error.

This paper presents a method to optimize the visual quality of Jpeg2000 compressed images, leaving the global bit-rate unchanged. Tiles that carry most important perceptual information are found using a "light" but effective classification algorithm. Such technique, based on simple energy measures, is able to distinguish between different areas (texture, homogeneous, edge...), in order to assign a different number of bits to tiles.

The remainder of the paper is organized as follows. Section 2 presents a brief overview of the Jpeg2000 compression standard. Section 3 describes the proposed optimization method while Section 4 presents results obtained applying such approach, in comparison with the standard method. Conclusion remarks are discussed in the last section.

II. JPEG2000, AN OVERVIEW

A brief Jpeg2000 overview is provided in this section. For more accurate review of the standard see [7, 13, 16].

Jpeg2000 compression standard has been developed to address the need of higher performances, as well as the support for new features, required in the area of still and video images encoding. Since each application area imposes different compression requirements, Jpeg2000 was created to provide functionalities that current standards does not address efficiently or does not address at all, in order to cover a wide range of applications. Figure 1 shows a simplified scheme of the Jpeg2000 encoding procedure. A preliminary step of the algorithm consists on partitioning the input image into its color components, which are coded independently. The standard supports two different color transformations:

- **Irreversible Color Transformation (ICT)**: the YCrCb transform commonly used with original Jpeg images;
- **Reversible Component Transformation (RCT)**: provides a similar decorrelation but allows lossless reconstruction of all components.

The encoding procedure starts partitioning the image into rectangular "tiles" on a regular grid. Arbitrary tile sizes are allowed, and the whole image can also be regarded as a single tile. Moreover, is possible to choose to partition the original image into overlapping tiles as to reduce blocking effects. The Discrete Wavelet Transform (DWT) [1] is then performed on each tile in order to decompose the tile into resolution levels. The subbands of the DWT coefficients are quantized and collected into rectangular arrays of "code-blocks". The bit-planes of the coefficients in each code-block are entropy coded. The entropy coding is based on the EBCOT algorithm (Embedded Block Coding with Optimized Truncation) [17], which makes the standard able to provide the main features described below:

- **Good performances at low bit-rates**: significantly improved compression quality at low bit-rates with respect to existing standards is achieved;
- **Scalability:** image can be restored increasing pixel accuracy or spatial resolution, because random access to the bitstream is allowed. Thus, scalability is obtained simply by discarding unwanted portions of the compressed data;
- **Lossless and lossy compression:** both lossy and lossless compression are provided, so scalability from lossy to lossless coding during progressive transmission is also possible;
- **Error resilience:** codestream is designed to preserve code-blocks which store most important information, and to avoid decoding failures;
- **Region Of Interest (ROI)**: most significant regions of the image can be coded and transmitted at higher quality than others.

Since EBCOT sorts coded data into "quality layers", ideally this embedded structure makes the encoder able to assign incremental contribution to quality layers, in such a way that the optimal rate-distortion image representation is built. EBCOT consists of two different coding engines, that is the low-level embedded block coding ("Tier 1") and the layer formation ("Tier 2"). The algorithm applied to find the best distortion for the requested target bit-rate is exactly the same for each tile of the image. Next section describes a method that exploits HVS properties to reduce the number of bits assigned to tiles in which errors are less visible and allocates more bits to the other tiles.

Threshold		class	credits
mean	variance		
<	<	plain	-2
<	>	edge	2
>	>	texture	-1
>	<	unknown	0

Table 1: Tiles classification and credit assignment scheme.

III. CONTENT-DEPENDENT OPTIMIZATION ALGORITHM

Jpeg2000 works partitioning the input image into rectangular tiles that are compressed independently. It is possible to fix the global compression ratio prior to compression. This is usually done assigning the same bit-rate to every tile.

The global compression factor, measured in bits per pixel (bpp), is given by:

$$R_I = \frac{\sum_T R_T}{N_T} \qquad (1)$$

where N_T is the number of tiles and R_T is the compression factor for a single tile *T*. Jpeg2000 reference encoders [10, 11, 12, 14] fix $R_I = R_T$ (all tiles compressed at the same bit-rate). Is implicitly assumed that the important details are uniformly distributed over the image tiles, which is not always the case. In fact, some tiles loose important features due to lossy compression, while other tiles retain scarcely visible details. The proposed approach re-distributes the available bits over the tiles, in order to preserve important data. It reduces the number of bits assigned to tiles in which distortion is less visible and, conversely, allocates more bits to tiles in which errors are more visible. This results in an image of better visual quality, while the total number of bits coded remains unchanged.

Let $R_T = R_I \cdot w_T$, where w_T is a positive weight. We call w_T the *perceptual weight* of the tile T. Thus eq. (1) becomes:

$$R_I = \frac{\sum_T (R_I \cdot w_T)}{N_T} = R_I \frac{\sum_T w_T}{N_T} \qquad (2)$$

with the constraint:

$$\frac{\sum_T w_T}{N_T} = 1 \qquad (3)$$

To fix the weights properly, tiles features should be recognized. Since the standard allows arbitrarily large tile size, each tile may present various morphological characteristics, so the classification is a very hard task. The proposed solution consists of two distinct steps:

1. Classifying each tile with respect to its features to estimate the error-visibility;
2. Assigning of a weight to the tiles on the basis of the prior classification.

To obtain a suitable classification of the tile content, the input image is filtered using a combination of directional edge-detection filters. We use Sobel filters to find horizontal and vertical edges and Roberts filters to detect diagonal edges [9]. To guarantee real time performances, only simple energy measures are computed. The classification is based on two statistical measures over the filtered images: the mean of the filtered coefficients of each tile, which measures the "quantity" of edges in the tile, and the variance of each coefficient with respect to the mean, which measures the mean slope of the luminance values of the pixels close to the edges. A low mean indicates that the tile has a plain background. If the variance is also low, than the tile has no significant details. Conversely, a high variance means that the tile has well visible details. High values for both mean and variance suggest that the tile coefficients exhibit high variability. Applying appropriate thresholds to the mean and variance measures, we are able to classify tiles. A simple classification is shown in Table 1. Here, each tile is classified as *Plain* (if the area presents both low mean and variance values), *Edge* (low mean and high variance), *Texture* (high mean and variance values) or *Unknown* (when the classifier is not able to understand tile features). Figure 2c) presents an example of tile classification. To deal with colour images, the proposed algorithm takes into account only the luminance channel.

The weight assignment is based on "credits". A positive number of credits is assigned to tiles with visually important details and, conversely, a negative number of credits is associated with tiles where distortion is less visible. An example of credit assignment is shown in Table 1. A negative number of credits is assigned to "plain-classified" tiles, because of the low error visibility in plain areas. As distortion is visible near to sharp edges, a positive number of credits is assigned to "edge-classified" tiles. A more subtle problem is the assignment of credits to "texture-classified" tiles. These tiles have a high information content but, due to a *visual masking* effect, the visibility of distortion in these areas is low.

Now that we have assigned a number of credits c_T to each tile *T* in the input image, we have to assign the corresponding weights. Let Δ_{pos} be the weight of a positive credit. Then the weight of a negative credit is defined by the following equation:

$$\Delta_{neg} = \frac{\Delta_{pos} \cdot N_{pos}}{N_{neg}} \qquad (4)$$

where N_{pos} and N_{neg} are the total number of positive and negative credits. The weights are assigned as follows:

$$w_T = \begin{cases} 1 + \Delta_{pos} \cdot c_T & \text{if} \quad c_T > 0 \\ 1 + \Delta_{neg} \cdot c_T & \text{if} \quad c_T < 0 \\ 1 & \text{if} \quad c_T = 0 \end{cases} \qquad (6)$$

Note that, due to the equation (4), the previous assignment does not violate the constraint (3).

There is a problem with the definition of Δ_{neg} in (4). When Δ_{pos} or N_{pos} are large it could happen that $\Delta_{neg} > 1$, thus violating the constraint $w_T > 0$. Moreover, even if every weight is positive, some weights could be too small or too large, resulting in a poor performance of the algorithm. To bound Δ_{pos} and Δ_{neg} to a reasonable range, we enforce $w_T \in [m_T, M_T]$, where $0 < m_T \leq 1 \leq M_T$. The value of Δ_{pos} and Δ_{neg} must be chosen in order to ensure that:

$$m_T < -\min(c_T) \cdot \Delta_{neg} \quad \text{(7a)} \quad \text{and}$$

$$\max(c_T) \cdot \Delta_{pos} < M_T \quad \text{(7b)}$$

The boundaries of m_T and M_T are chosen as a function of the target bit-rate R_I and of the information content of each tile *T*.

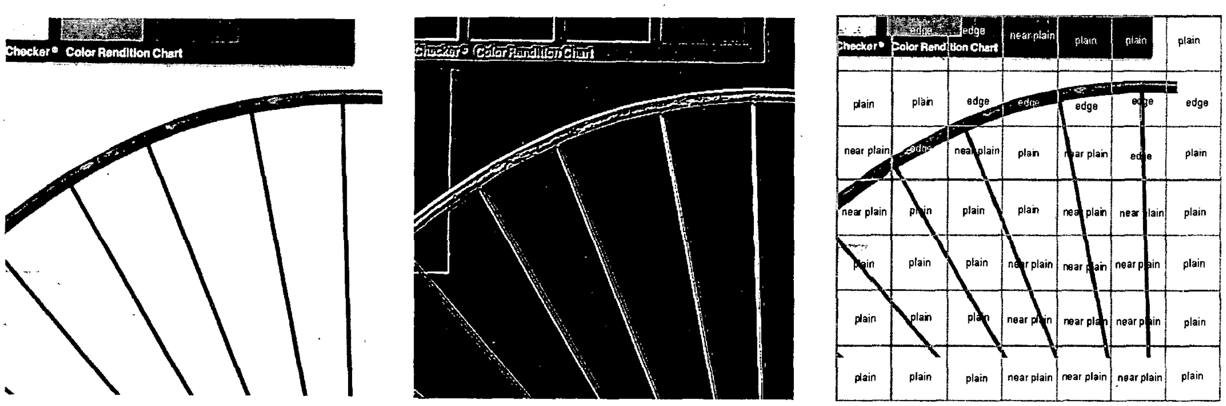
Threshold				class	Credits
mean 1	mean 2	SQM 1	SQM 2		
<	<	<	<	plain	-2
>	<	<	<		
<	<	>	<	near plain	-1
>	<	>	<		
<	<	>	>	edge	2
>	>	>	>	texture	-1
>	<	>	>	unknown	0
>	>	<	<		
>	>	>	<		

Table 2: Improved classification and credit assignment.

IV. Results

Experimental results show that the proposed approach has better performances than classic Jpeg2000 in terms of perceptual quality. Effectiveness of the algorithm depends on the classification and it's strongly related with the tile size and the bit-rate chosen. Of course, at high bit-rates no perceptible improvements could be achieved varying the distribution of available bits among tiles and, on the other hand, when the tile size is too large it carries too much information to build a coherent classification. We report some examples of images optimized with the proposed algorithm, using a slightly more sophisticated classification scheme, as shown in Table 2. Figure 2 c) shows the classification that describes the test image "Bike1" in Figure 2 a). Observe that in the border tiles there is not enough information to get a correct classification. Some tile in the center of the image has been classified in an unexpected way. The behavior of the classifier could be explained as follows. Our method is based on the measurement of the variance that provides an accurate description of the details contained in each tile (Figure 2b). Thus, even a non-uniform colour distribution in an object of the scene has a heavy influence on the measure, and the threshold should be fixed properly to obtain a robust classification. But a high threshold could, in the same way, introduce errors in presence of uniform colour distributions. A useful approach to this problem could be reducing the number of the gray levels, with an appropriate quantization, in order to build clusters of pixels with the same gray value. This should reduce the susceptibility of the edge detection filters to noise peaks in the image, thus resulting in a more robust classification. On the other hand, such a quantization could smooth important edges reducing the effectiveness of the filtering stage. The quantization step should be carefully designed so as to avoid this problem. However, the comparison between the image "Bike1" compressed using the standard Jpeg2000 and the proposed optimization method shows that a perceptible improvement of the quality of the image is achieved (Figure 3). The improvement is clearly visible at 0.125 bpp and at 0.25 bpp, while at 0.5 bpp it's perceptible just in some regions of the picture (e.g. text).

Figure 4 shows another example. In this case, good results were obtained for the tiles containing the center of the racket (Figure 5 a-b), at the cost of a loss of data in the "textured" zones of the image that corresponds to tiles enlarged in Figure 5 c-d. However, the global quality of the image appears improved.



(a) (b) (c)

Figure 2: a)"Bike1" test image. (b) Filtered image describing energy features of the image "Bike1". (c) "Bike1" classification with tiling of size 64X64 pixel.

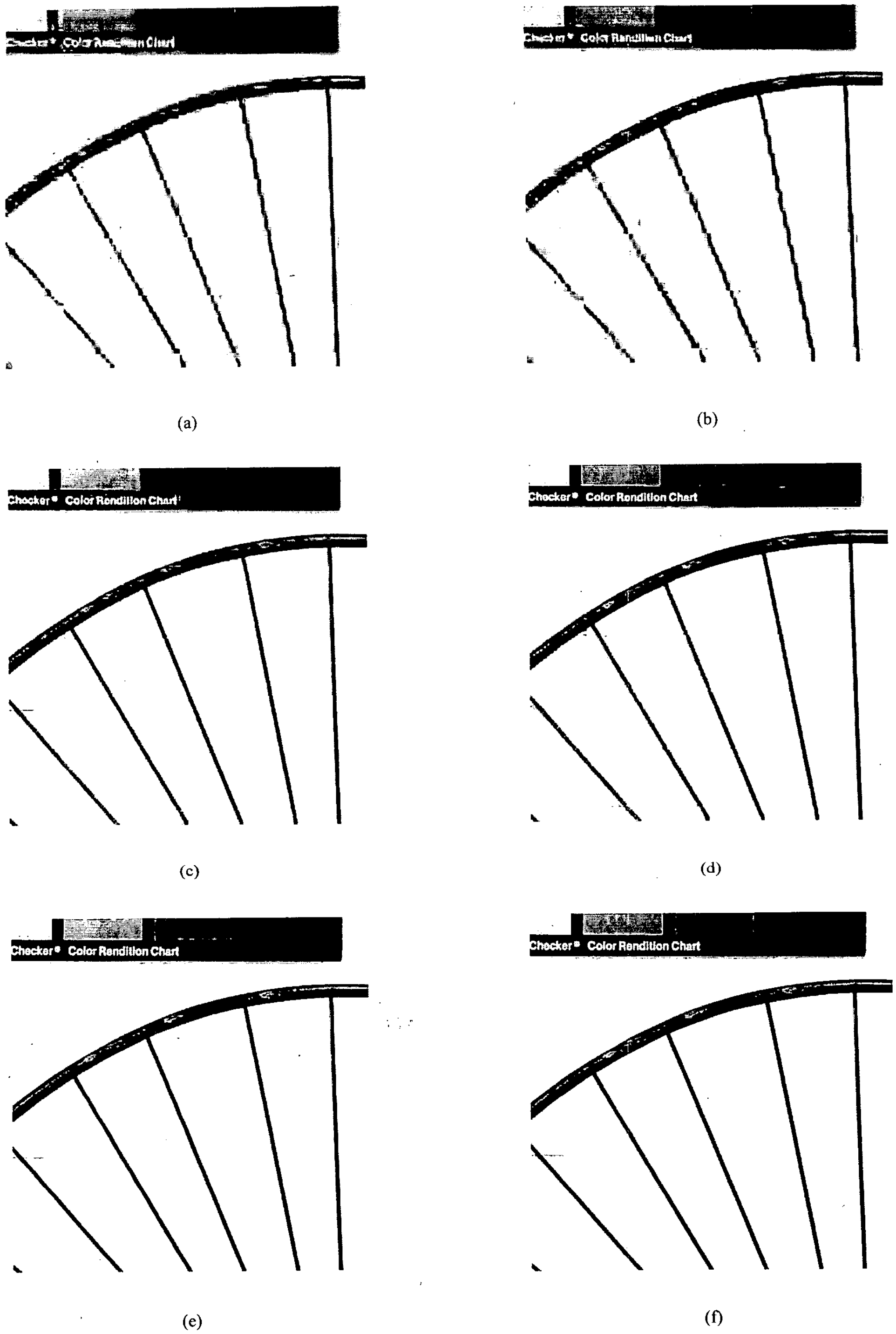


Figure 3: Comparison among Jpeg2000 compressed images obtained using the standard approach (on the left) and the proposed "content-dependent" algorithm (on the right). Compression is performed using tiles of size 64X64 pixels at different bitrates: 0.125 bpp (a-b), 0.25 bpp (c-d), 0.5 bpp (e-f).

We have roughly compared the number of primitive operations per pixel executed by the proposed algorithm and by the Jpeg2000 encoder. We found that our algorithm introduces little computational load in the overall compression algorithm. Moreover, the size of the input in the filtering stage is greater, by some orders of magnitude, than the size of the input of the subsequent steps. Thus, only the filtering phase makes a significant contribution to the complexity of the algorithm. The number of operations per pixel depends highly on the type and number of the edge detection filters used. A reduction of the time complexity in this step would, in turn, lead to an extremely fast algorithm.

V. Conclusions

A content-dependent method to improve visual quality of Jpeg2000 compressed image has been presented. The proposed approach codes the most visually meaningful regions of the image using more bits then those used to code homogeneous zones, leaving the global bit-rate unchanged. The proposed approach does not modify Jpeg2000 encoder, but processes the bitstream, thus its complexity overhead could be considered negligible. As the proposed algorithm does not modify the encoder, the output bitstream is compliant to the standard.

A promising direction for future work concerns the computational complexity of the proposed algorithm. Since the Jpeg2000 encoder performs a band-pass filtering step during the wavelet transform, the high frequency sub-bands contain some information about the edges in the input image. If we were able to exploit this information, and to obtain a robust classification, we could obtain the filtering step "for free". As a consequence, the overall computational load of the proposed approach would be negligible.

Acknowledgment

We wish to thank Prof. Giovanni Gallo of University of Catania for many helpful suggestions.

References

[1] M. Antonini, M. Barlaud, P. Mathieu, I. Daubechies, *Image Coding Using Wavelet Transform*, IEEE Trans. On Image Processing, Vol.1, n.2, pp. 205-220, April 1992.

[2] S. Battiato, A.Bosco, M. Mancuso, G. Spampinato, *Temporal Noise Reduction of Bayer Matrixed Data,* In Proceedings of IEEE ICME'02 International Conference on Multimedia and Expo 2002, Lausanne, Switzerland, August 2002.

[3] S. Battiato, A. Castorina, M. Mancuso, *High Dynamic Range Imaging for Digital Still Camera*, In Proceedings of SPIE *Electronic Imaging 2002* - Sensors, Cameras, and Applications for Digital Photography IV - San Jose' CA USA, January 2002.

[4] S. Battiato, M. Mancuso, *An Introduction to the Digital Still Camera Technology*, ST Journal of System Research - Special Issue on *Image Processing for Digital Still Camera*, Vol. 2, No.2, December 2001.

[5] A. Bruna, M. Mancuso, *JPEG Compression Factor Control: A New Algorithm*, in Proceedings of IEEE ICCE2001, pp.206-207, Los Angeles, June 2001.

[6] C. Chou, Y. Li, *A Perceptually Tuned Subband Image Coder Based on the Measure of Just-Noticeable-Distortion Profile*, IEEE Trans. On Circuits and Systems for Video Technology, Vol.5, n.6, pp. 467-476, December 1995.

[7] C. Christopulos, A. Skodras, T. Ebrahimi, *The Jpeg2000 Still Image Coding System: an Overview,* IEEE Trans. On Consumer Electronics, Vol. 46, n.4, pp.1103-1127, November 2000.

[8] S. Daly, W. Zeng, J. Li, S. Lei, *Visual Masking in Wavelet Compression for Jpeg2000,* SPIE/IS&T Electronic Imaging: Image and Video Comm. and Processing, San Jose, January 2000.

[9] R.C. Gonzales, R.E. Woods *Digital Image Processing*, Addison Wesley, pp. 197-2001, 1993.

[10] *Jasper Project Home Page*, http://www.ece.ubc.ca/~mdadams/jasper/

[11] *JJ2000: An Implementation of Jpeg2000 Standard in Java*, http://jj2000.epfl.ch/

[12] *Kakadu: A Comprehensive, Heavily Optimized, Fully Compliant Software Toolkit For JPEG2000 Developers*, http://www.kakadusoftware.com/

[13] *JPEG2000 Image Coding System Part I*, ISO/IEC FCD 15444-1. March 2000.

[14] *JPEG2000 Verification Model 9.1,* ISO/IEC JTC 1/SC 29/WG1 N2165, June 2001.

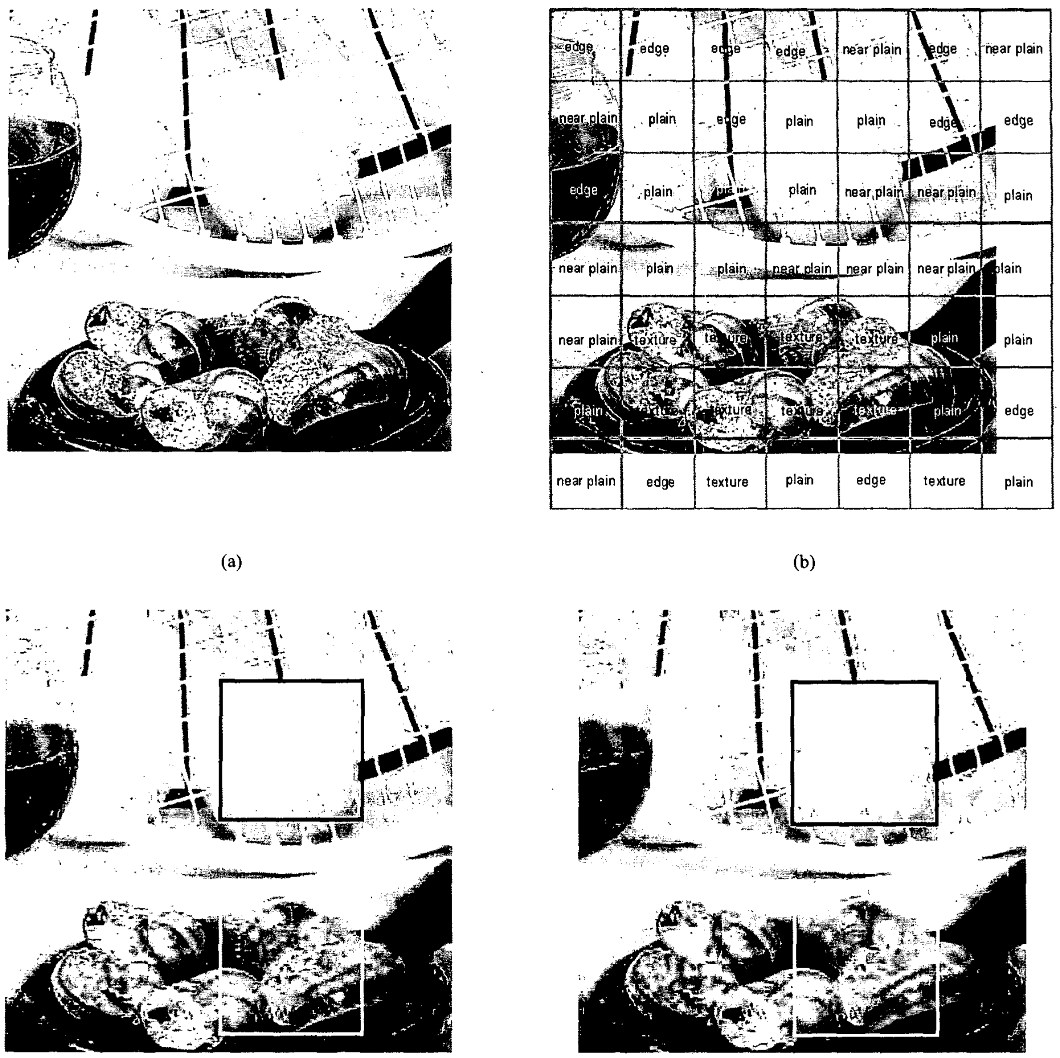
[15] M. G. Perkins, T. Lookabaugh, *A Psychophysically Justified Bit Allocation Algorithm for Subband Image Coding Systems*, pp.1815-1818, in Proceedings of ICASSP, 1989.

[16] D. Taubman, M. Marcellin *JPEG2000 Image Compression Fundamentals*, The Kluwer Int. Series in Eng. and Computer Science Volume 642 Hardbound, ISBN 0-7923-7519-X, Nov. 2001.

[17] D. Taubman, E. Ordentlich, M. Weinberger, G.Seroussi, *Embedded block coding in JPEG2000*, In Proceedings of IEEE ICIP, 2001.

[18] H. H. Y. Tong, A. N. Venetsanopoulos, *A Perceptual Model for Jpeg Applications Based on Block Classification, Texture Masking and Luminance Masking*, 0-8186-8821-1/98. 1998.

[19] K. A. Wen, C.Y. Lu, M. C. Tsai, *The Transform Image Codec Based on Fuzzy Control and Human Visual System*, IEEE Trans. on Fuzzy Systems, Vol. 3, No. 3, 1995.



(a) (b)

(c) (d)

Figure 4: (a) "Bike 2" test image. (b) "Bike2" classification with tiling of size 64X64 pixel. (c) Images compressed at 0.25 bpp using Jpeg2000 and tiles of 64X64 pixels. (d) Optimized image obtained applying the proposed approach.

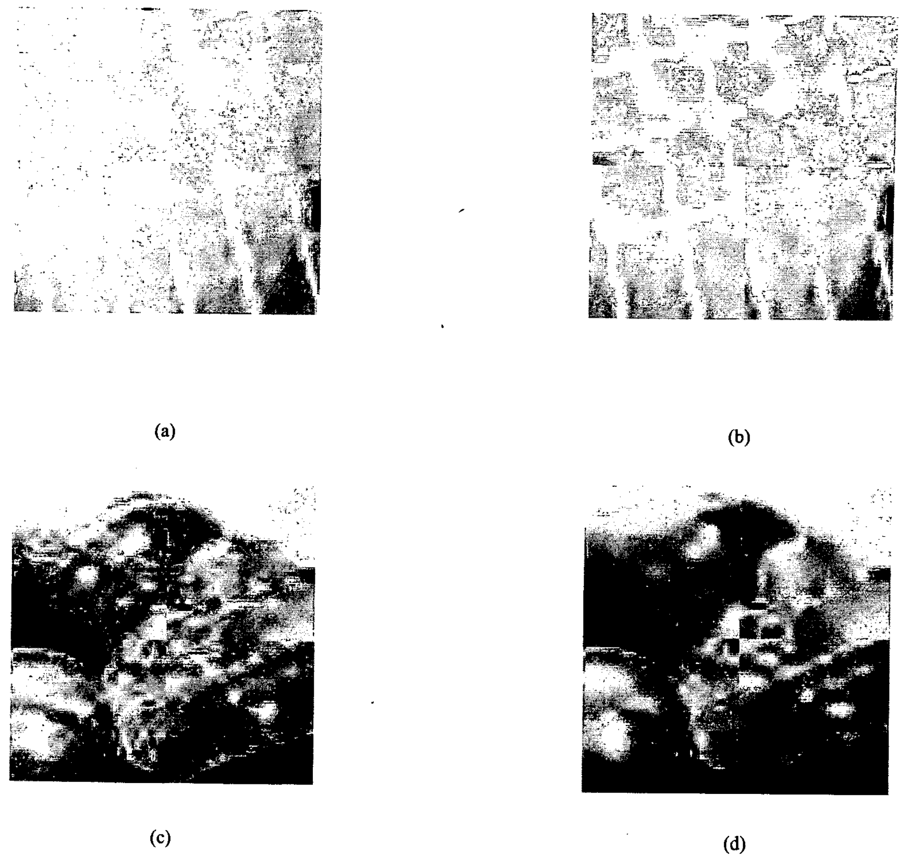


Figure 5: (a-b) Enlarged view of tiles black-highlighted in Figure 6 c) using standard Jpeg2000 and the optimization method: more details are visible in the optimized picture. (c-d) Enlarged view of tiles white-highlighted in Figure 6 d): a loss of data is perceptible in the output of the proposed method.

S. Battiato received his Italian degree in Computer Science in 1995 and PhD in Applied Mathematics and Computer Science in 1999 both at Catania University. He attended several international workshops/conferences in the imaging field where some results of his research has been presented and then published in the relative Proceedings. A few papers have been also published in International Journals. Since November 1999 has been working at STMicroelectronics in AST Catania Lab in the Digital Still Camera Group as Algorithm Dev. Team Leader. His current research interests are in the area of Image Processing, Compression (Jpeg, Jpeg2000), Watermarking, and Analysis/Synthesis of Texture.

A. Buemi received his degree in Computer Science at Catania University in 1998. In 1999 he won a study grant announced by Italian National Research Council and he spent it at Catania University, developing new image processing algorithms based on fuzzy logic. Since the end of 2000 has been working at STMicroelectronics in AST Catania Lab in the Digital Still Camera Group as System Engineer. His current research activity concerns the following areas of Image Processing: compression of digital image (Jpeg2000), textures classification and fuzzy algorithms for medical applications.

G. Impoco graduated in Computer Science at the University of Catania (Italy) in 2001. Since the beginning of 2002 he is a PhD student in Computer Science at the University of Pisa. He is part of the Visual Computing Group (a departement of ISTI-CNR), which is involved in topics in the area of 3D Graphics and Visualization. His current research interests concern image compression, motion tracking and 3D graphics.

M. Mancuso received the Degree in Electronics Engineering (summa cum laude) from the University of Palermo (Italy) in 1990. He cooperated with the University of Genova for the development of CAL system for DSP. After a research period at the University of Palermo in the Image Processing field, he joined the Co.Ri.M.Me (a consortium between SGS-Thomson and the University of Catania) and he was involved in image processing using Fuzzy Logic based techniques. At the present he works as Digital Still Camera & Multimedia Mobile Program Manager at AST Catania Lab of STMicroelectronics. He is author of several papers and patents in Image Processing field.