Analysis and Characterization of JPEG 2000 Standard for Imaging Devices

S. Battiato[1], A. R. Bruna[2], A. Buemi[2] and A. Castorina[2]

Abstract — *This paper presents a series of possible improvement strategies and/or implementations related to the JPEG 2000 standard for imaging devices. A comparative study relative to various coding options is presented together with some brief considerations about rate-ctrl techniques and perceptive improvement strategies.*

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 [6]. 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 [6], [8]. The high rate of innovation in the image processing technology involves changing in the product features. The new Jpeg2000 compression standard [10], [14], [21] 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.

The JPEG2000 standard allows to manage images with functions that were not present in the other formats and this suggest a wide usage of this format in mobile imaging applications. JPEG2000 has a better efficiency than JPEG for high compression rates and currently perform, on more images, better results than any other standard. Moreover using a progressive decoding, the image visualization will be refined at each step adapting itself to the particular resolution display used (i.e resolution and/or quality scalability).

Up to now Digital Still Camera (DSC) [6] market has been focused on quality improvement (obtained by increasing the resolution of the sensor and by using more sophisticated image processing algorithms) and in reduction of processing time. On the other hand Mobile Still Image is focused on the achievement of the best trade off between quality and low power in terms of processing, related hardware and bandwidth occupation. A typical mobile device for image has a VGA (640 x 480 pixels) sensor that captures an image of 0.3 million of pixels doing very few elaborations. The main purpose of the paper is the detailed description of this new different world both in term of complexity (hardware/software) and related capabilities with respect to the JPEG 2000 standard compression.

The rest of the paper is structured as follows. The next section describes how imaging devices acquire images using CCD/CMOS sensors. Section III is devoted to the description of an exhaustive study about the numerous coding options of the new standard. Some considerations about Jpeg2000 rate-control techniques are briefly discussed in Section IV. The next Section shows a smart technique able to improve the perceived quality of images heavily Jpeg2000-compressed modulating the relative bit-rate of adjacent tiles. A conclusions Section closes the paper tracking direction for future works and research.

II. SENSORS AND DATA ACQUISITION

A digital engine uses an electronic sensor to acquire the spatial variations in light intensity and then use image processing algorithms to reconstruct a color picture from the data provided by the sensor (see Figure 1).

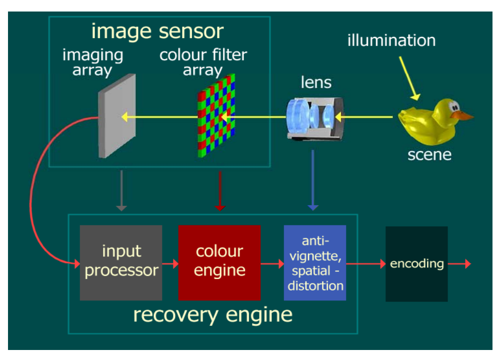


Figure 1: Image acquisition (DSC/Mobile) principle.

Two technologies exist to manufacture imaging sensors: CCD (Charge Coupled Device) or CMOS (Complementary Metal Oxide Semiconductor**)**.

The CMOS approach is more flexible because each pixel can be read individually. CCDs use a special manufacturing process to create the ability to transport charge across the chip without distortion. CMOS chips, on the other hand, use

[1] Sebastiano Battiato works at Catania University – Dipartimento di Matematica ed Informatica – E-mail {battiato@dmi.unict.it};
[2] Antonio Buemi, Arcangelo Ranieri Bruna, Alfio Castorina work for STMicrolectronics – AST Catania Lab – E-mail: {antonio.buemi, arcangelo.bruna, alfio.castorina}@st.com.

Contributed Paper
Manuscript received September 23, 2003

completely normal manufacturing processes to create the chip: the same processes used to make most microprocessors. There are several noticeable differences between CCD and CMOS sensors.

CCD sensors create high-quality, low-noise images. CMOS sensors, traditionally, are more susceptible to noise but are low-power: 100 times less than an equivalent CCD sensor. Based on these differences, CCDs tend to be used in cameras that focus on high-quality images with lots of pixels and excellent light sensitivity. CMOS sensors usually have lower quality, lower resolution and lower sensitivity. However, CMOS cameras are much less expensive and have great battery life.

Usually each sensitive element of the sensor (known as pixel) is sensitive to one color component only. This pattern is known as Bayer pattern [7]. Picture quality is strictly related to the number of pixels composing the sensor: the higher the better. But increasing the sensor resolution without increasing its size reduces the area of individual pixels, and therefore their sensitivity (see Figure 2).

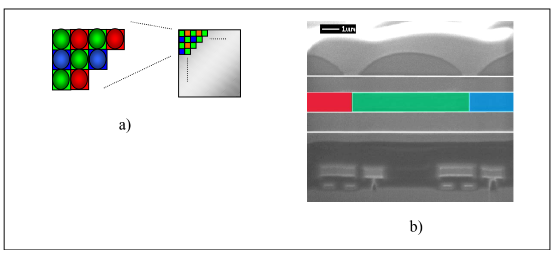


Figure 2: A sensor color is obtained by a mono sensor plus suitable RGB color filters. Each pixel/color filter has a small lens a) placed on top, improving significantly light gathering at pixel site. On the right b) it is showed the typical layout of a single CMOS sensor.

In wireless imaging field, the characteristics of the sensor used for mobile application must take into account primarily dimension and power consumption. For a mobile device working with battery supply is essential to use less charge as possible and not to be big or have an excessive weight. For all these reasons a CMOS sensor is preferred to a CCD sensor and many sensor manufacturers are oriented in this direction. They support both still and video imaging with lower resolution as VGA (640x480) or CIF (352x288) formats. Sensor must be very low power and in less than a cubic centimeter must comprise also optic, lens and mechanical housing. Another issue for the system-level project regarding the sensor is that the preferred area to integrate the sensor in near the handset's antenna and so strong Electro-Magnetic Interference can be registered on the sensor. For this reason sensor packaging has a relevant role in project phase. Another very important point in the wireless imaging architecture is the partitioning of the image processing in the architecture and the interface between communication modules. For these reasons the sensor could be used only as input device demanding properly processing to a remote and distributed environment [19].

Recently we have considered the possibility to compress images coming from CCD/CMOS sensors directly in Bayer format [3], [5] to reduce the amount of switching activity inside the bus-data. Classical coding techniques in this case do not always offer satisfying performances. In particular, a non-standard "vector quantization"-based [5] approach was compared with JPEG and JPEG-LS standard algorithms. All these techniques have high performance in terms of compression quality at high bitrates without perceptual loss. Moreover, compression does not affect further stages of IGP that take compressed data as input in order to obtain a full color image in a typical digital camera architecture. Performances were evaluated also in terms of computational complexity, a basic evaluation criteria for "bayer-oriented" compression algorithms. The Jpeg2000, as described in the next section, for its relative complexity, is not well suited for such kind of application where a low-power consumption solution is needed.

III. Features and Complexity Analysis

A. Coding Options

The Jpeg2000 image compression standard has been designed to address requirements of a large range of applications, so it provides many different features and functionalities. Supporting different functionalities means, of course, realizing a more complicated design of the encoder system that manage these new capabilities. The influence of the main Jpeg2000 coding parameters over the compression quality has been evaluated varying the coding parameters and measuring the resulting PSNR. In particular, the experimental results reported here has been obtained over a database of 30 images coded varying the following parameters:

- Bit rate (minimum = 0.125 bpp, maximum=3 bpp).
- Wavelet transforms.
- Number of transform decomposition levels (varied from 2 to 10).
- Tile size (minimum = 64x64 pixels, maximum = 512x512).
- Overlapped and un-overlapped tiling.
- No-tiling encoding has also been considered.

All the Jpeg2000 coded images have been generated using the Verification Model software [15]. The best compression results were obtained with (9,7) wavelet transform, overlapped tiling with big tile size and 4 decomposition levels.

B. Wavelet transforms comparison

To evaluate how wavelet transforms affect the compression quality in Jpeg2000, our database of testing images has been encoded setting the number of decomposition levels to the default value (five) without using tiling partitions and properly varying the output bit rate. Finally the PSNR of the resulting

images has been computed. (9,7) floating point wavelet application yields better compression than the (5,3) integer wavelet, in terms of PSNR.
The gap between the wavelet transforms grows with the increasing of the bit rate until of 2 bpp, and then it begins to decrease.

On the average, however, there is a PSNR difference of about 0.3 dB in the compression quality obtained using the two DWT different coding options, so there is no a perceptive difference in the visual analysis of two images obtained in this way (see Figure 3 for an example).

Figure 3: input image (top). Image compressed at 0.125 bpp using W5X3 (middle) and W9X7 (bottom) with no tiling, and 5 decomposition levels.

Figure 4: Input test image for Jpeg/Jpeg2000 comparison.

C. *Jpeg2000 versus Jpeg*

This section presents a comparison between Jpeg2000 and Jpeg standard. To reduce the amount of data, Jpeg uses a RGB to YCbCr transform followed by subsampling. Even if Jpeg2000 could achieve the same effect using the multi-resolution nature of the wavelet transform [1], our analysis presents results obtained over 4:2:2 Jpeg2000 coded images to make an effective comparison with Jpeg. Multi-components coding is enabled in Jpeg2000 processing each component independently. As show the diagram in Figure 5, which presents the PSNR behavior of the image at different bit rates, in this particular example Jpeg2000 outperforms Jpeg and presents also a reduced blocking effect at 0.25 bpp, but Jpeg yields a better compression at bit rates greater than 1 bpp.

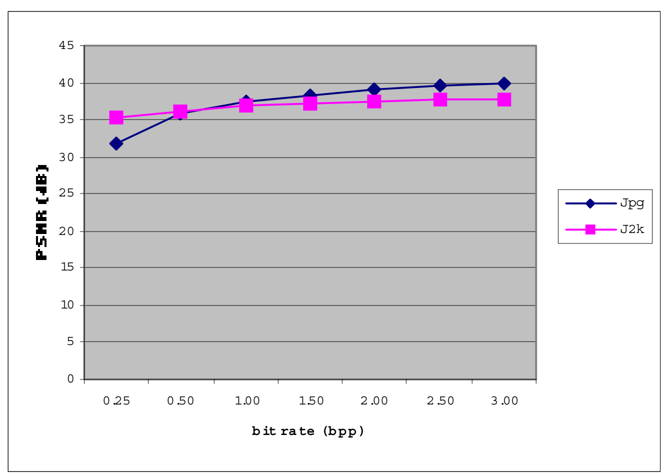


Figure 5: PSNR comparison between Jpeg and Jpeg2000 compression for the picture Figure 4.

However such weakly different results were obtained when different kinds of image are processed. For the textured image in Figure 6, Jpeg2000 outperforms Jpeg at every bit rates (Figure 7). Artifacts are more evident in the Jpeg coded image even at quite high bit rates, as confirmed by a perceptive analysis. The image in Figure 8 shows an example with more colors and details than the previous, and the superiority of Jpeg2000, again, decreases as bit rate increases. Jpeg has better performances since the bit rate is greater or equal to 2 bpp (see Figure 9).

Figure 6: Textured test image for Jpeg/Jpeg2000 comparison.

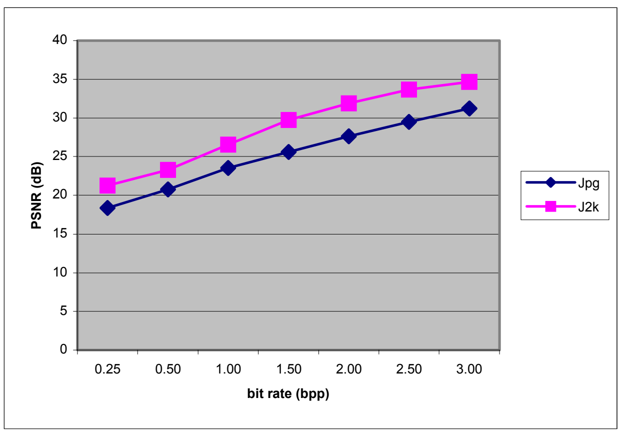


Figure 7: PSNR comparison between Jpeg and Jpeg2000 compression for the picture in Figure 6.

Figure 8: Input test image for Jpeg/Jpeg2000 comparison.

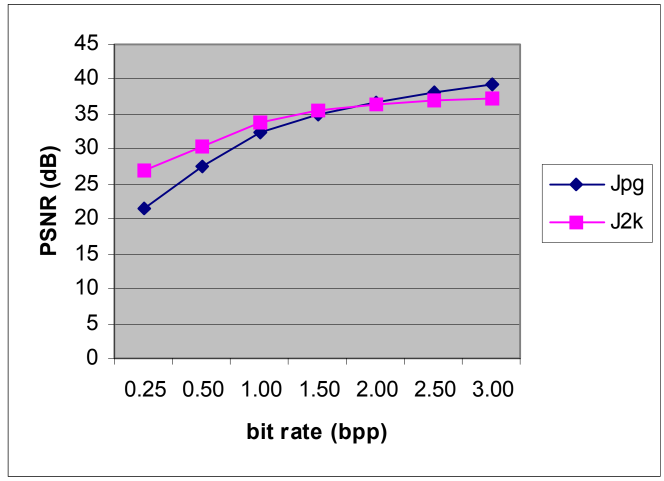


Figure 9: PSNR comparison between Jpeg and Jpeg2000 compression for the picture in Figure 8.

IV RATE CONTROL

The Jpeg2000 core-coding engine is based on the 2-tiers paradigm of the EBCOT [20], [22] algorithm. The first tier provides a low level coding of block samples belonging to the various image sub-bands by means of a context adaptive BAC (Binary Arithmetic Coding) referred as MQ coder [17]. Once all blocks are compressed they are re-examined to produce a full packed bit-stream, where packets can be described as sub-parts of code-blocks bit-stream that are spread across different quality layers with each quality layer increasingly improving the final image. This framework provides a bit-stream that is both resolution scalable (if DWT is adopted) and quality scalable (if multiple layers are provided). Of course such structured bit-stream is well suited for wireless applications where bandwidth constraints require maximum limited data transmission and/or a limited display resolution is available.

Optimal block truncation is decided within Rate-Distortion optimisation algorithms [16] to produce layers of fixed bit rate or distortion. RD problems can be easily solved relying on the well-known *lagrangian multiplier* [12] method. If, for example, a target rate R_{max} is desired for a given layer the following problem must be solved:

$$Minimize \quad D = \sum_{i=0}^{N-1} D_i^{n_i} \quad such\ that$$
$$R = \sum_{i=0}^{N-1} R_i^{n_i} \le R_{\max} \tag{1}$$

, where N are the number of the blocks, n_i the chosen truncation point for the i-*th* block and $R_i^{n_i}$, $D_i^{n_i}$ respectively distortion and rate for a block at a given truncation point. Possible truncations points (*3*K-1* with *K* the bit depth) are produced by the MQ-coder in 3 different coding steps. Solving (1) equals solving the following:

$$Minimize \quad D(\lambda) + \lambda R(\lambda) = \sum_{i=0}^{N-1} D_i^{n_i} + \lambda R_i^{n_i} \tag{2}$$

, for a $\lambda>0$ found through bisection to achieve the target rate.

The [14] provides a basic distortion metric but it would be appealing to draw a specific metric taking into account specific display limitation (size, bit-depth, viewing angle, …). Such drawbacks are, well known to be relevant for imaging mobile devices.

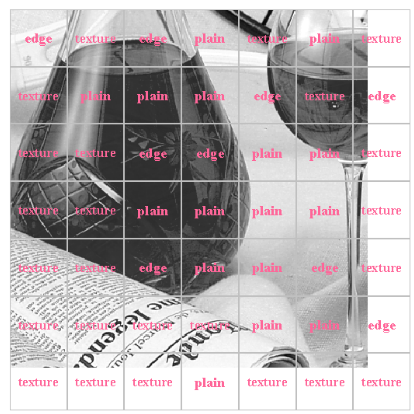
The benefits coming from the above described rate-control mechanism, known in literature as PCRD (Post Compression Rate Distortion Optimization) poses some practical restrictions both in terms of memory allocation and overall computation. Using some adaptive single-pass analysis similar to that described in [8], a more effective strategy could be designed.

V. Perceptive improvement

The algorithms used for compression factor control have the main purpose of assuring a fixed file size. This constraint affects heavily the choice of the trade-off between compression ratio and image quality. In order to improve the visual quality of the image for a fixed bit rate, some basic property of the HVS could be exploited to code fewer bits to represent perceptual less important areas of the image. These properties have been exploited in several methods in order to improve Jpeg compression standard performances [9], [11], [18], [23]. Using similar principles, we developed a method to optimize the visual quality of Jpeg2000 compressed images, leaving the global bit-rate unchanged [4]. Since Jpeg2000 works partitioning the input image into rectangular tiles that are compressed independently, the global compression ratio is fixed assigning the same bit-rate to every tile. 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. 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.

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 [13]. To each tile is then associated a weight 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. 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 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, then 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. Exactly, 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). To deal with colour images, the proposed algorithm takes into account only the luminance channel.

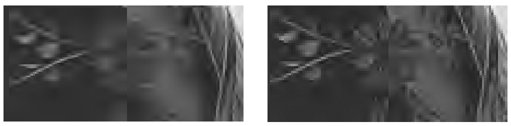
Experimental results show that the proposed approach has better performances than classic Jpeg2000 in terms of perceptual quality. Effectiveness of the algorithm depends on the classification and it's strongly related with the tile size and the bit-rate chosen. Of course, at high bit-rates no perceptible 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.



(a)

(b)

©



(d)

Figure 10: (a) an example of tile classification. (b) Compression obtained by standard Jpeg2000. (c) Compression obtained using the proposed approach. (d) Tiles highlighted in (b) and (c) enlarged to show the different performances.

VI. Conclusion

The Jpeg 2000 compression standard allows working with image compression with new and powerful advanced features. The overall performances are sensibly higher than classical JPEG, especially at low bit-rates. At first glance such behavior coupled with native scalability seems devoted to wireless applications. Nevertheless in real world the j2k format has not widely adopted yet. Classical Jpeg is still nowadays the first choice in imaging devices maybe for two main reasons:

1 Manufacturer industry re-conversion problem (DCT to DWT based processors);

2 Is the overhead due to the numerous coding options and related coding complexity justified for actual device limits?

The paper has shown some initial studies addressing the main basic characteristics. Of course further investigation is needed to find optimal trade-offs.

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.Bruna, A. Buemi, F. Naccari – *Coding Techniques for CFA Data Images* – To appear in Proceedings of IEEE *ICIAP'03 International Conference on Image Analysis and Processing* – Mantova, Italy – September 2003.

[4] S. Battiato, A. Buemi, G. Impoco, M. Mancuso, *Jpeg2000 Coded Images Optimisation Using a Content-dependent Approach*, IEEE Transaction on Consumer Electronics - Vol. 48, Issue 3, pp. 400-408, August 2002.

[5] S. Battiato, A. Buemi, L. Della Torre, A. Vitali, *A Fast Vector Quantization Engine for CFA Data Compression,* To appear in Proceedings of IEEE-EURASIP *Workshop on Nonlinear Signal and Image Processing* – NSIP 2003 – Grado, Italy - June 2003.

[6] S. 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.

[7] B.E Bayer, *Color imaging array*, US Patent 3,971,065 1976.

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

[9] 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.

[10] 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.

[11] 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.

[12] H. Everett, *Generalized Lagrange multiplier method for solving problems of optimum allocation of resources,* Operations Research, vol. 11, pp.399-417, 1963.

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

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

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

[16] A. Ortega, K. Ramchandran, *Rate-Distortion Methods for Image and Videocompression,* IEEE Signal Processing Magazine, November 1998.

[17] W. B. Pennebaker, J. L. Mitchell, *JPEG: Still Image Compression Standard,* Van Nostrand Reinhold, New York, 1993.

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

[19] A. Puliafito, M. Mancuso, M. Villari, M. Guarnera, *Digital still cameras and mobile agents: How to create a distributed service for image processing,* In proc. of EW2002;

[20] D. Taubman, *High performance scalable image compression with EBCOT,* IEEE Trans. Image Processing, vol. 9, pp. 1158-1170, July 2000.

[21] 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.

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

[23] 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.

Sebastiano Battiato received his Italian degree in Computer Science in 1995 and PhD in Applied Mathematics and Computer Science in 1999 both at Catania University. From November 1999 to August 2003 has worked at STMicroelectronics in AST Catania Lab in the Digital Still Camera Group as Algorithm Dev. Team Leader. Actually he works as contract professor at Catania University. His current research interests are in the area of Image Processing, Compression (Jpeg, Jpeg2000), Watermarking, and Analysis/Synthesis of Texture.

Arcangelo Bruna received his Italian degree in Electronic Engineer in 1998 at the University of Palermo. His thesis was on video indexing database and features extraction based on MPEG sequences and was published by IEEE on ICMCS-1999. Then he worked in Rome for Telesoft firm on telecommunication area as consultant for about 9 months. Since July 1999 he is working at STMicroelectronics in AST Catania Lab in the Digital Still Camera Group. During the first year his main activity was the research of new algorithms and architectures for the JPEG compression factor control and his results allowed to get 3 patents. Nowadays he is working on MOW (Multimedia Over Wireless), a project joined with Agrate and Edimburgh Labs.

Antonio Buemi received his Italian degree in Computer Science at the University of Catania in 1998. In 1999 he won a study grant announced by Italian National Research Council and he spent it at University of Catania, 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.

Alfio Castorina received his Italian degree in Computer Science in 2000 at the University of Catania doing a thesis about watermarking algorithms for digital images. Since September 2000 he has been working in STMicroelectonics in the AST Digital Still Camera Group as System Engineer. His current activities include high dynamic range imaging, post-processing image enhancement techniques and Jpeg 2000.