Compressive Data Hiding: An Unconventional Approach for Improved Color Image Coding

Patrizio Campisi

Dipartimento di Ingegneria Elettronica, Universitá degli Studi di Roma "Roma Tre," Via della Vasca Navale 84, 00146 Roma, Italy
Email: campisi@ele.uniroma3.it

Deepa Kundur

Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto,
Toronto, Ontario, Canada M5S 3G4
Email: deepa@comm.toronto.edu

Dimitrios Hatzinakos

Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto,
Toronto, Ontario, Canada M5S 3G4
Email: dimitris@comm.toronto.edu

Alessandro Neri

Dipartimento di Ingegneria Elettronica, Universitá degli Studi di Roma "Roma Tre," Via della Vasca Navale 84, 00146 Roma, Italy
Email: neri@ele.uniroma3.it

Traditionally, data hiding and compression have had contradictory goals. The former problem adds perceptually irrelevant information in order to embed data, while the latter removes this irrelevancy and redundancy to reduce storage requirements. In this paper, we use data hiding to help improve signal compression. We take an unconventional approach and consider "piggy-backing" the color information on the luminance component of an image for improved color image coding. Our new technique essentially transforms a given color image into the YIQ color space where the chrominance information is subsampled and embedded in the wavelet domain of the luminance component. Our technique can be used as preprocessing to improve the performance of popular image compression schemes such as SPIHT that are optimized for grayscale image compression. Simulation results demonstrate the superior performance of the proposed technique in comparison to JPEG and straightforward SPIHT.

Keywords and phrases: color images, data hiding, compression, multiresolution analysis.

1. INTRODUCTION

Data hiding within multimedia has received growing interest in recent years due to its potential for signal captioning, maintaining audit trails in media commerce, and copy protection through the development of digital watermarking technology. By embedding key information with the media itself, it is safe from content separation. Data hiding is the general process by which a discrete information stream is merged within media content by imposing imperceptible changes on the original *host* signal.

One of the main obstacles within the data hiding community has been developing a scheme which is robust to *perceptual coding*. *Perceptual coding* refers to the *lossy* compression of multimedia signals using human perceptual models; the compression mechanism is based on the premise that minor modifications of the signal representation will not be noticeable in the displayed signal content. These modifications are imposed on the signal in such a way as to reduce the number of information bits required for storage of the content. Human perceptual models are often theoretically and experimentally derived to determine the changes on a signal which remain imperceptible. A duality exists between the problems of perceptual coding and data hiding; the former problem attempts to remove irrelevant and redundant information from a signal, while the latter uses the irrelevant information to mask the presence of the hidden data. Thus, the goals of data hiding and perceptual coding can be viewed as being somewhat contradictory. As a result, several papers have dealt with integrating perceptual coding with

data hiding [1, 2, 3, 4, 5, 6, 7], and others have investigated the theoretical relationship between both processes [8, 9]. In [10], data hiding for media compression is investigated. The method operates in the frequency domain and it is based on linear projection, quantization, and perturbation.

The central theme of all the works cited above is that there must be an appropriate *compromise* between data hiding and compression to develop a method which performs both reasonably. It is assumed that each process hinders, not helps, the objective of the other. Specifically, data hiding decreases the overall possible compression ratio and perceptual coding tampers with the hidden information, so that extraction is difficult.

In this paper, we take a different, perhaps even eccentric, perspective. We try to identify how data hiding can be used for *improved* practical compression.

1.1. Objectives of this paper

In this work we present an approach to improve the efficiency of compression by incorporating data hiding principles. On a larger scale, the presented work aims to, in part, investigate the contradictory processes of data hiding and compression in order to derive insights into effective means to merge them.

Specifically, we wish

- to design a compression scheme in which color information is "piggybacked" on the grayscale component to provide the option of viewing the information in color or as a monochrome signal;
- to compare our proposed data hiding-based compression approach practically with JPEG, a popular color image compression format, and SPIHT, an effective wavelet-based compression algorithm.

2. HYPOTHESIS AND INTUITION

Consider the signal $f_0(x)$ which represents an audio signal, image, or video sequence. There exists a family of functions in the set $\mathbf{P}(f_0)$ which are perceptually identical to $f_0(x)$. Thus, if $f_k(x) \in \mathbf{P}(f_0)$ then we know that $f_k(x)$ is perceptually identical to $f_0(x)$.

In ideal compression, every possible perceptually identical signal is mapped to the same representation. Thus, all signals in the set $\mathbf{P}(f_0)$ will be collapsed into one compressed signal. In data hiding, information is embedded into a host signal $f_0(x)$ by modifying it so that the resulting signal is perceptually identical to the original; therefore this new signal is also in the set $\mathbf{P}(f_0)$. It then follows that ideal compression applied to a signal containing hidden information has the effect of annihilating the discrete data.

However, in practice, compression is not completely efficient; that is, there exists some irrelevant information which has not been removed. The nonideality comes from the constraint that the coder have structure, and inadequate perceptual models to account for all masking characteristics [11]. In terms of our description above, for practical compression not all signals in $\mathbf{P}(f_0)$ are mapped to the same representation. Thus, there is a small bandwidth available for data

hiding. If this bandwidth could be used to transmit information about the signal such as chrominance components, then even greater practical compression may be achieved. This approach cannot help to improve the compression ratio in the case of ideal perceptual coding, but can improve the situation in the case of practical compression.

The approach can potentially provide improved performance when we are restricted to use a coding scheme which is not very efficient. For example, one of the most popular wavelet image compression schemes SPIHT [12] is optimized for grayscale compression, and may not perform acceptably for color images. If the technique outlined above could be applied to piggy-back the chrominance components in a compressed version of the luminance image, then some performance advantage may be established. Thus, we propose our scheme as a possible improvement to existing approaches which may not perform optimally. This also has the advantage in that the color chrominance is embedded in the grayscale, so that it may be viewed later even if color is not initially important.

In Section 3, we present our approach for compressive data hiding. Experimental results are provided in Section 4 in which we compare the proposed technique to SPIHT and JPEG to demonstrate its superior performance. Discussion and final remarks conclude the paper.

3. COMPRESSIVE DATA HIDING

3.1. The problem

As discussed in the introduction, we focus on the problem of improving compression through the use of judicious data hiding. Consider the situation in which we are restricted to use a given compression algorithm. For example, we may be using a specific web browser or media player that supports only a limited number of image compression algorithms or an online viewer for dynamic information swapping. It would be valuable to be able to improve the quality of the compressed information without interfering with the compression algorithm. One approach would be to preprocess the signal in order to improve upon potential compression artefacts.

In this paper, we consider using robust data hiding at this preprocessing stage in order to pass or tag signal information reliably through the compression stage. The information then can be used, after compression, in order to improve signal fidelity or provide some other form of added-value. As discussed in the introduction, many watermarking algorithms have been proposed which are robust to specific forms of compression, so we believe that this preprocessing stage is possible to effectively design.

For this work, we specifically consider the case of color image compression; we embed chrominance information into the luminance component of the image to improve the signal fidelity upon post-compression reconstruction for a fixed compression ratio. We make use of the popular wavelet-based compression algorithm SPIHT, so that data hiding can also be effectively accomplished in this multiresolution-like domain.
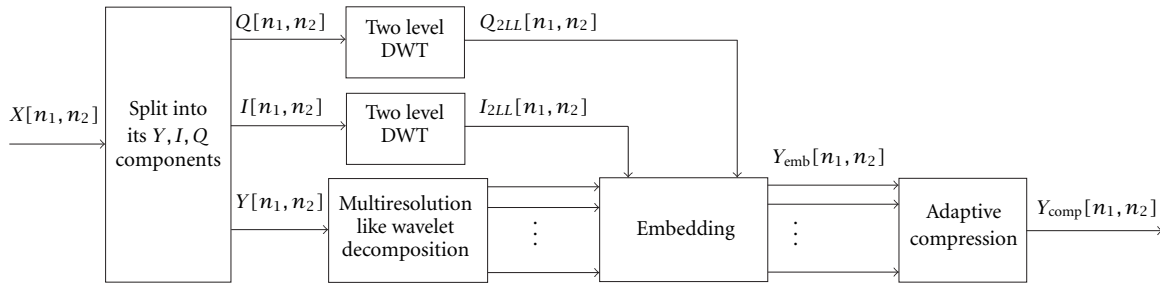
FIGURE 1: Compressive data hiding scheme for color images.

3.2. Overview of the approach

A specific overview of the proposed compressive data hiding scheme for color images is discussed in this section. Figure 1 provides a block diagram representation of the approach. Our original color image to be compressed is denoted $X[n_1, n_2]$.

Since the chrominance components of the signal are piggy-backed on the luminance component of the image using data hiding it is important to judiciously select both the color space and transform domain for this processing. Selection of these two components can affect the data embedding capacity and, hence, the performance of the proposed approach.

As we see in Figure 1, the YIQ color space and the discrete wavelet transform (DWT) domain [13] are incorporated into our approach. Essentially, the color image $X[n_1, n_2]$ is converted to the YIQ color space where the chrominance (or color) components are processed in the DWT domain and are embedded in the wavelet domain of the luminance (or grayscale) component of the original color image. After embedding the information, an adaptive scheme is used to aid in the compression process.

The next two sections discuss why the choice of the color space and transform domain provide a good tradeoff between imperceptibility and robustness of the embedded information. Details of the embedding and adaptive compression technique are also provided in the subsequent sections.

3.3. The YIQ color space

The choice of the color space where to perform the decomposition of the given color image is relevant to the problem we have addressed. In fact, since our goal is compression, color spaces, such as the RGB space, where there is a significant correlation between the three color components should be avoided. On the contrary, the color spaces YIQ, YUV, and YC_bC_r nearly provide as much energy compaction as a theoretically optimal decomposition performed using the Karhunen-Loeve transform. Since they are equivalent for our application, we have resorted to split the given color image into its three components in the YIQ color space, where the Y coordinate represents the luminance $Y[n_1, n_2]$ and the I and Q coordinates represent the chrominance components $I[n_1, n_2]$ and $Q[n_1, n_2]$, respectively. The I and Q components are used to jointly represent saturation and hue.

It is worth noting that the luminance component Y contains a large component of the visual content, whereas the two chrominance components, I and Q, contain less perceptual information. Thus, due to the human visual system's lower sensitivity to color information, it is possible to subsample the chrominance information and then integrate it back into the overall color image without any loss of perceptual quality [14].

This color space division is useful as the overall signal can be separated into a high volume luminance component which can act as the host image for data hiding, and two lower volume color information-bearing signals which can act as the payload. In addition, the luminance component is essentially the grayscale component of the signal which allows us to incorporate well-established grayscale image data hiding principles for our embedding procedure.

3.4. The compressive data hiding procedure

3.4.1 The transform domain

The chrominance components $I[n_1, n_2]$ and $Q[n_1, n_2]$ are subsampled and hidden in the luminance component (see Figure 1). For robust grayscale image data hiding, the embedding is performed in a specific transform domain which is selected to be suitable for a specific application or attack on the hidden information. It has been shown [9, 15] that different domains have significantly different data hiding channel capacities. Selection of a suitable transform can significantly improve the robustness and imperceptibility of the hidden information.

In this work, we incorporate the DWT domain. The spatial and frequency localization of this transform suits the behavior of the human visual system (HVS) to visual stimuli [16]. Embedding information in this domain allows the hidden information to be appropriately shaped for reduced distortion. In particular, we have the flexibility to judiciously select DWT subbands of the luminance component in which to hide the chrominance information imperceptibly. In addition, recent work by Fei et al. [9] demonstrates the advantage of this domain for wavelet-based compression. Furthermore, for the adaptive compression approach shown in Figure 1 and discussed in Section 3.4.3, decomposition of the image to compress into multiresolution-like components is convenient and saves unnecessary complexity to transform the

signal into this form for more effective compression.

We next provide the specifics of our proposed data embedding technique. We make use of an *unconventional* two-level multiresolution-like wavelet decomposition on the luminance component. With reference to Figure 2, the first level of the multiresolution decomposition is obtained by performing a DWT on $Y[n_1, n_2]$. In particular, an 8 tap Daubachies filter is used:

$$Y[n_1, n_2] \xrightarrow{\text{DWT}} \left\{ Y_{LL}[n_1, n_2], Y_{HH}[n_1, n_2], \right.$$
$$\left. Y_{HL}[n_1, n_2], Y_{LH}[n_1, n_2] \right\}, \quad (1)$$

thus obtaining the subbands $Y_{LL}[n_1, n_2]$, $Y_{LH}[n_1, n_2]$, $Y_{HL}[n_1, n_2]$, and $Y_{HH}[n_1, n_2]$, which take into account the image at a coarser resolution plus the "horizontal," "vertical," and "diagonal" details of the image also at this resolution, respectively.

The subbands $Y_{LH}[n_1, n_2]$ and $Y_{HL}[n_1, n_2]$ are chosen to host the chrominance information. The rational behind this choice relies on the observation that, in order to obtain a good tradeoff between robustness and transparency, many watermarking techniques (cf. [17, 18] and the references therein) use "middle frequency" coefficients. This makes the subbands $Y_{LH}[n_1, n_2]$ and $Y_{HL}[n_1, n_2]$ suitable to host the data, whereas the subband $Y_{HH}[n_1, n_2]$ is not.

The next step of the method consists in further decomposing the subbands $Y_{LH}[n_1, n_2]$ and $Y_{HL}[n_1, n_2]$ in the wavelet domain thus leading to the subbands:[1]

$$Y_{HL}[n_1, n_2] \xrightarrow{\text{DWT}} \left\{ Y_{ll,HL}[n_1, n_2], Y_{hh,HL}[n_1, n_2], \right.$$
$$\left. Y_{hl,HL}[n_1, n_2], Y_{lh,HL}[n_1, n_2] \right\},$$
$$\quad (2)$$
$$Y_{LH}[n_1, n_2] \xrightarrow{\text{DWT}} \left\{ Y_{ll,LH}[n_1, n_2], Y_{hh,LH}[n_1, n_2], \right.$$
$$\left. Y_{hl,LH}[n_1, n_2], Y_{lh,LH}[n_1, n_2] \right\}.$$

In particular, $Y_{ll,HL}[n_1, n_2]$ and $Y_{ll,LH}[n_1, n_2]$ represent the low-pass subbands, at coarser resolution, obtained from the high frequency subbands $Y_{HL}[n_1, n_2]$ and $Y_{LH}[n_1, n_2]$, respectively. It is expected that their energy contribution is relatively small compared to the energy of the remaining subbands of the set $Y_{\alpha,\beta}[n_1, n_2]$ ($\alpha \in \{ll, hl, lh, hh\}$, $\beta \in \{HL, LH\}$).

Thus, they can be neglected introducing a very low mean square error (MSE) that takes into account the variations occurred in the image details, which however do not affect the image in a perceptual sense.

This conjecture has been experimentally verified upon a wide range of images' typology using both subjective evaluation criteria such as visual perceptibility measures and objective ones like the Peak Signal to Noise Ratio (PSNR), given by

TABLE 1: PSNR.

	PSNR$_{\text{zeroed}}$	PSNR$_{\text{embed}}$
Baboon	33.4	32.1
Biked	35.3	33.4
Lena	36.2	35.3
GoldHill	34.6	32

$$\text{PSNR} = 10 \log_{10} \left(\frac{255^2}{\text{MSE}} \right) [dB], \quad (3)$$

where

$$\text{MSE} = \frac{1}{N^2} \sum_{n_1=0}^{N-1} \sum_{n_2=0}^{N-1} \left(Y[n_1, n_2] - Y_{\text{mod}}[n_1, n_2] \right)^2 \quad (4)$$

is the mean square error between the original image $Y[n_1, n_2]$, of $N \times N$ pixels, and a modified replica $Y_{\text{mod}}[n_1, n_2]$. To assess the general perceptual irrelevance of the $Y_{ll,HL}[n_1, n_2]$ and $Y_{ll,LH}[n_1, n_2]$ bands, we consider a series of test images with widely varying characteristics. For each image, a modified replica is produced by zeroing only the subbands $Y_{ll,HL}[n_1, n_2]$ and $Y_{ll,LH}[n_1, n_2]$ and keeping the remaining perfectly intact. The PSNRs of the resulting images (denoted by PSNR$_{\text{zeroed}}$) are presented in the first column of Table 1. The values are reasonably high; in addition, there is no perceptual change in the quality of the modified signal $Y_{\text{mod}}[n_1, n_2]$ in each test case.

3.4.2 Subsampling of the chrominance information for embedding

After having thus verified that the subbands $Y_{ll,HL}[n_1, n_2]$ and $Y_{ll,LH}[n_1, n_2]$ are perceptually negligible, we use a very straightforward and computationally simple approach to embed the chrominance components: direct replacement of the $Y_{ll,HL}[n_1, n_2]$ and $Y_{ll,LH}[n_1, n_2]$ information with the chrominance content. However, the chrominances need to be preprocessed in order to obtain a perceptually lossless parsimonious representation. This can be performed through subsampling of the color information.

Specifically, the components undergo a two-level wavelet decomposition using, as for the luminance analysis, an 8 tap Daubachies filter. As already outlined, their perceptual contribution is much smaller than the luminance's and thus, their size may be reduced without significant distortion [19]. The subbands, $I_{2LL}[n_1, n_2]$ and $Q_{2LL}[n_1, n_2]$, which are the low-pass chrominance replicas at the coarsest resolution of the performed pyramidal decomposition, can be used to reconstruct the color image without any perceptual quality loss. Experiments have verified this conjecture.

Therefore, we can replace the subbands $Y_{ll,HL}[n_1, n_2]$ and $Y_{ll,LH}[n_1, n_2]$ with information from $I_{2LL}[n_1, n_2]$ and $Q_{2LL}[n_1, n_2]$ to obtain $Y_{\text{emb}}[n_1, n_2]$ as shown in Figure 2. However, before the embedding, the energies of $I_{2LL}[n_1, n_2]$ and $Q_{2LL}[n_1, n_2]$ have to be normalized to the values of the corresponding host subbands, as not to impair the perceptual appearance of the reconstructed image. It should be

---

[1]This is the *unconventional* part of the scheme. Usually the $Y_{LL}[n_1, n_2]$ is further decomposed instead of its details. However, this decomposition is necessary for the imperceptibility of our embedding stage as we will discuss later in this section.
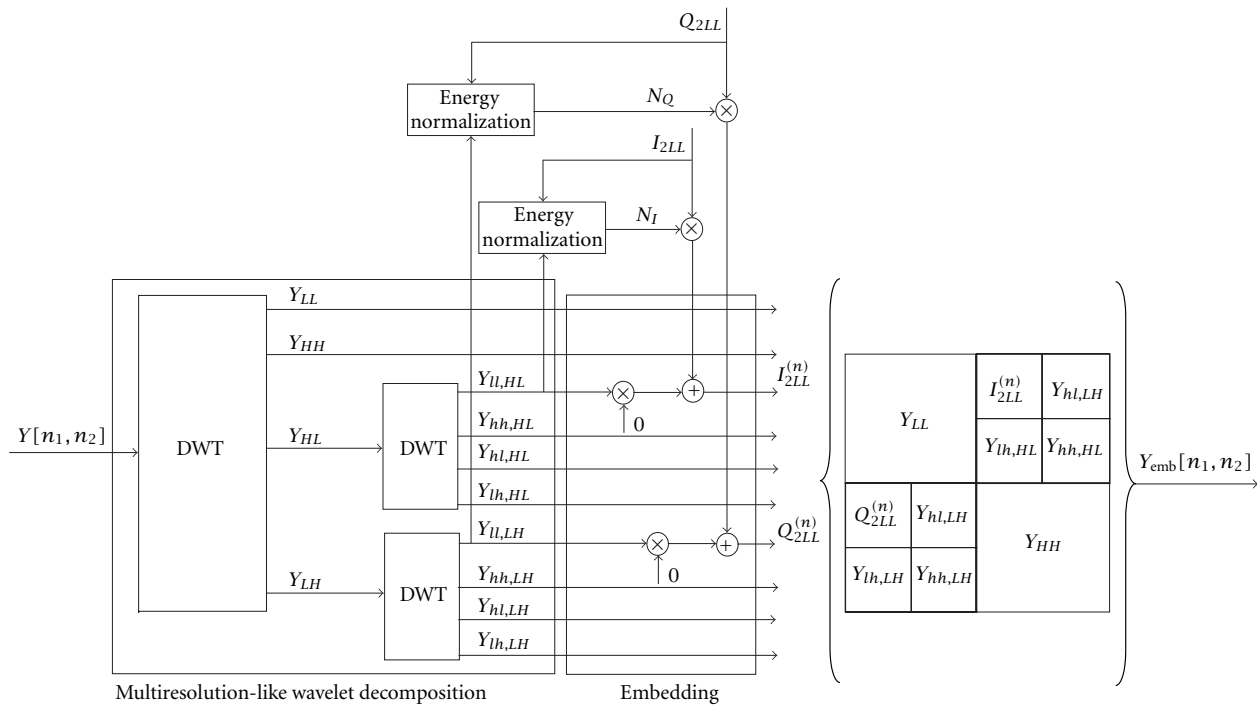
FIGURE 2: Multiresolution-like wavelet decomposition and data embedding scheme.

noted that the normalization values, say N_I and N_Q, have to be transmitted to the decoder since they are necessary to properly reconstruct the color information. To this end they can be embedded in the header of the image.

We have experimentally verified that the aforementioned embedding procedure causes no perceptual degradation to the luminance component of the signal. Some of these results are reported in Figure 3 where $Y_{emb}[n_1, n_2]$, the luminance image component with the embedded chrominance information, is shown for different test cases.

It is worth noting that, the chrominance information $(2 * (64 * 64)$ bytes) represented by $I_{2LL}[n_1, n_2]$ and $Q_{2LL}[n_1, n_2]$ is embedded into the luminance $((256 * 256)$ bytes), having assumed to use a *true color* 24-bit representation for the color images used in our experiments (i.e., 8-bit representation for each component). Nevertheless, the luminance after the embedding appears perceptually indistinguishable from the host image. Also a quantitative evaluation is performed by calculating the PSNR, denoted PSNR$_{embed}$, whose values for different test images, are presented in the second column of Table 1.

The reader should note that the embedding procedure proposed for our implementation is based on replacement of perceptually irrelevant bands which is distinct from popular methods such as Spread Spectrum (SS) watermarking proposed by Cox et al. [17] and Quantization Index Modulation (QIM) proposed by Chen and Wornell [20]. SS watermarking suffers from host signal interference which limits the data hiding capacity necessary for our application. Replacement embedding as performed in the paper is a form of information hiding termed "Low Bit Modulation (LBM)" in [20].

Chen and Wornell show that using coding theory measures such as minimum distance and information capacity that LBM and SS are inferior to QIM in certain contexts.

However, the QIM method is not appropriate for the proposed data hiding application presented in the paper. One problem is that robust QIM implementations require error correction coding which increases bandwidth and makes it impractical to embed the necessary volume of color information in the luminance component. A second problem is that the QIM method is designed to reliably embed data bits, not perceptually viewable information such as logos or chrominance image bands. With perceptual information, some errors can be tolerated unlike with data and hence our data hiding problem is less restricted than that formulated in [20], and hence their solution is not as appropriate. We attempt to clarify in the next paragraph.

The natural redundancy of many forms of perceptual information, such as images, often makes it robust to errors. Specifically, when viewed, the received signal is of better perceptual quality if it is transmitted in raw form through certain nonideal channels than it would be if the information was first source and channel coded [21]. Intuitively, the degradation on the embedded information which characterize our effective attack channel are from quantization for compression. This process is designed to be applied on raw data and leave perceptually salient information in tact. Thus, sophisticated high bandwidth error correction codes are not necessarily required. Our embedding scheme of simple replacement is analogous to transmitting the chrominance bands in the raw.

Overall, the low bandwidth requirements, natural robustness to quantization distortions, and low complexity makes

FIGURE 3: Left column: original grayscale images. Right column: images with the chrominance components I and Q embedded.

our embedding technique more appropriate than other popular schemes such as SS and QIM.

3.4.3 Adaptive compression

A compression algorithm such as SPIHT optimized for grayscale image compression may be applied to the resulting signal in order to produce a coded image. However, we find that further preprocessing can be applied for improved compression by taking into account the diverse nature of the luminance signal component at the different subbands. Our approach follows.

The first step consists of obtaining the subbands $Y_{HL}^{(e)}[n_1, n_2]$ and $Y_{LH}^{(e)}[n_1, n_2]$, after having performed the embedding, by calculating the inverse discrete wavelet trans-
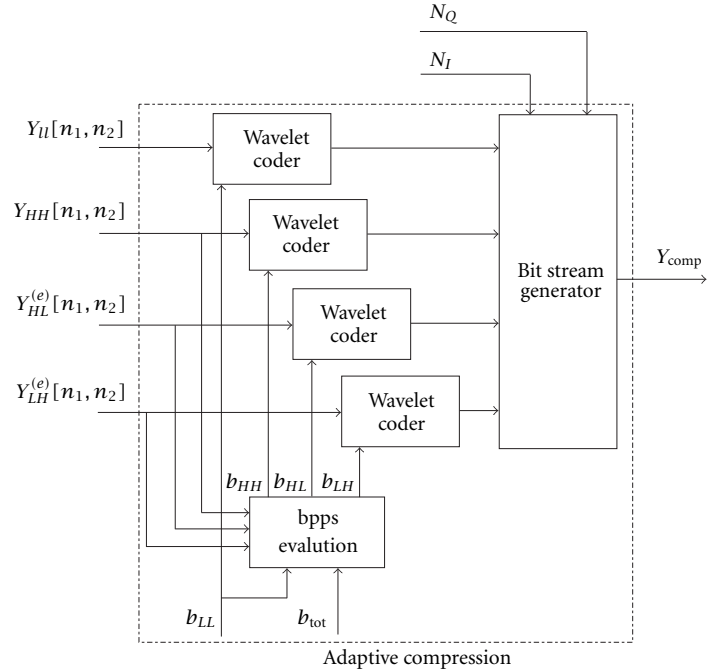

FIGURE 4: Adaptive compression scheme.

form (IDWT) on the subbands of (5):

$$
\begin{aligned}
&\left\{ I_{2LL}^{(n)}[n_1, n_2], Y_{hh,HL}[n_1, n_2], \right.\\
&\left. Y_{hl,HL}[n_1, n_2], Y_{lh,HL}[n_1, n_2] \right\} \xrightarrow{\text{IDWT}} Y_{HL}^{(e)}[n_1, n_2],\\
&\left\{ Q_{2LL}^{(n)}[n_1, n_2], Y_{hh,LH}[n_1, n_2], \right.\\
&\left. Y_{hl,LH}[n_1, n_2], Y_{lh,LH}[n_1, n_2] \right\} \xrightarrow{\text{IDWT}} Y_{LH}^{(e)}[n_1, n_2].
\end{aligned}
\tag{5}
$$

Therefore, the luminance and the embedded chrominance components are represented, in a perceptually lossless manner, by the subbands:

$$
\left\{ Y_{LL}[n_1, n_2], Y_{HH}[n_1, n_2], Y_{HL}^{(e)}[n_1, n_2], Y_{LH}^{(e)}[n_1, n_2] \right\}.
\tag{6}
$$

In order to preserve the information embedded, each subband of (6) is coded, according to the scheme shown in Figure 4, separately. Specifically, a wavelet-based coder is used instead of a DCT-based coder; it has been proven that wavelet based coders provide better rate-distortion performance than the DCT-based JPEG and also allow a progressive coding approach [22]. The method of set partitioning in hierarchical tree (SPIHT) [12] has been employed in the proposed scheme. It is well known that the magnitude of the wavelet coefficients varies from band to band; in particular, lower frequency subbands usually have a higher magnitude than the coefficients in the higher frequency subbands. This suggests that different bit rates must be used according to the specific subband. Consider,

$$
\{ b_{LL}, b_{HH}, b_{HL}, b_{LH} \}
\tag{7}
$$

TABLE 2: Compressive data hiding procedure.

Compressive Data Hiding

Embedding steps

(1) The color image $X[n_1, n_2]$ is split into its three color components in the YIQ color space.

(2) The luminance Y undergoes a multiresolution-like wavelet decomposition:

$$Y \xrightarrow{\text{DWT}} (Y_{LL}, Y_{HH}, Y_{HL}, Y_{LH}),$$

$$Y_{HL} \xrightarrow{\text{DWT}} (Y_{ll,HL}, Y_{hh,HL}, Y_{hl,HL}, Y_{lh,HL}),$$

$$Y_{LH} \xrightarrow{\text{DWT}} (Y_{ll,LH}, Y_{hh,LH}, Y_{hl,LH}, Y_{lh,LH}).$$

(3) The chrominance components I and Q undergo a two-level wavelet decomposition and only the "lowpass" subbands at the coarsest resolution I_{2LL} and Q_{2LL} are kept.

(4) I_{2LL} and Q_{2LL} are normalized to the energy of $Y_{ll,HL}$ and $Y_{ll,LH}$, respectively, thus obtaining $I_{2LL}^{(e)}$ and $Q_{2LL}^{(e)}$.

(5) The subbands $Y_{HL}^{(e)}$ and $Y_{LH}^{(e)}$ are obtained as follows:

$$(I_{2LL}^{(n)}, Y_{hh,HL}, Y_{hl,HL}, Y_{lh,HL}) \xrightarrow{\text{IDWT}} Y_{HL}^{(e)},$$

$$(Q_{2LL}^{(n)}, Y_{hh,LH}, Y_{hl,LH}, Y_{lh,LH}) \xrightarrow{\text{IDWT}} Y_{LH}^{(e)}.$$

Compression steps

(6) The global bit rate b_{tot} and the bit rate b_{LL}, for the subband Y_{LL}, are chosen by the user.

(7) The bit rates b_{HH}, b_{HL}, b_{LH} corresponding to the remaining subbands are evaluated according to (8) and (9).

(8) Finally each subband is compressed using the SPIHT coder and the bit stream Y_{comp} is generated.

---

the bit per pixel (bpp) for each of the subbands in (6), respectively. Moreover, let b_{tot} be the desired bpp for the compressed color image, which is related to the bpps in (7) as follows:

$$b_{\text{tot}} = b_{LL} + b_{HH} + b_{HL} + b_{LH}. \tag{8}$$

The proposed criterion for adaptive compression consists in specifying the global bit rate b_{tot} and the bit rate b_{LL} for the subband $Y_{LL}[n_1, n_2]$, since this latter value plays the most significant role in the decoded image appearance. The remaining bpps b_{HH}, b_{HL}, b_{LH} are automatically assigned by the coder in such a way that a higher bit rate is assured to the subbands having higher energy. Therefore,

$$b_{LH} = \frac{\mathcal{E}_{LH}}{\mathcal{E}_{HH}} \cdot b_{HH}, \qquad b_{HL} = \frac{\mathcal{E}_{HL}}{\mathcal{E}_{HH}} \cdot b_{HH}, \tag{9}$$

where $\mathcal{E}_\gamma$ ($\gamma \in (LH, HL, HH)$) are the energy of the different subbands. After having chosen b_{tot} and b_{LL}, according to the user's needs, the bit rates for each subband are obtained from (8) and (9).

As a rule of thumb we have chosen for our experiments a bit rate b_{LL} that is half the global bit rate b_{tot}. As pointed out

in Section 4, this choice, with no claim of optimality, shows the potentialities of our approach.

Finally each subband is compressed, at the rates previously evaluated, using the SPIHT coder. The so obtained compressed subbands along with the energy normalization values generate the bit stream Y_{comp}. The compressive data hiding procedure is summarized in Table 2.

3.5. *Color information retrieval*

At the receiver, the color image is reconstructed, from the compressed bit stream $Y_{\text{comp}}[n_1, n_2]$, by performing dual operations of those of the coding stage. The first step consists of extracting the single subbands from the bit stream and then decoded, thus obtaining an estimate (denoted by ^)

$$\left\{ \hat{Y}_{LL}[n_1, n_2], \hat{Y}_{HH}[n_1, n_2], \hat{Y}_{HL}^{(e)}[n_1, n_2], \hat{Y}_{LH}^{(e)}[n_1, n_2] \right\} \tag{10}$$

of the corresponding quantities in (6). The estimated chrominance information $\hat{I}_{2LL}[n_1, n_2]$ and $\hat{Q}_{2LL}[n_1, n_2]$ is extracted from $\hat{Y}_{HL}^{(e)}[n_1, n_2]$, $\hat{Y}_{LH}^{(e)}[n_1, n_2]$ by performing the DWT as follows:

$$\hat{Y}_{HL}^{(e)}[n_1, n_2]$$
$$\xrightarrow{\text{DWT}} \left\{ \hat{Y}_{ll,HL}[n_1, n_2] = \hat{I}_{2LL}[n_1, n_2], \hat{Y}_{hh,HL}[n_1, n_2], \right.$$
$$\left. \hat{Y}_{hl,HL}[n_1, n_2], \hat{Y}_{lh,HL}[n_1, n_2] \right\},$$

$$\hat{Y}_{LH}^{(e)}[n_1, n_2]$$
$$\xrightarrow{\text{DWT}} \left\{ \hat{Y}_{ll,LH}[n_1, n_2] = \hat{Q}_{2LL}[n_1, n_2], \hat{Y}_{hh,LH}[n_1, n_2], \right.$$
$$\left. \hat{Y}_{hl,LH}[n_1, n_2], \hat{Y}_{lh,LH}[n_1, n_2] \right\}. \tag{11}$$

After having zeroed $\hat{Y}_{ll,HL}[n_1, n_2]$ and $\hat{Y}_{ll,LH}[n_1, n_2]$, the subbands $\hat{Y}_{HL}[n_1, n_2]$ and $\hat{Y}_{LH}[n_1, n_2]$ are reconstructed by performing a one level IDWT

$$\left\{ \hat{Y}_{ll,HL}[n_1, n_2] = 0, \hat{Y}_{hh,HL}[n_1, n_2], \right.$$
$$\left. \hat{Y}_{hl,HL}[n_1, n_2], \hat{Y}_{lh,HL}[n_1, n_2] \right\} \xrightarrow{\text{IDWT}} \hat{Y}_{HL}[n_1, n_2],$$
$$\left\{ \hat{Y}_{ll,LH}[n_1, n_2] = 0, \hat{Y}_{hh,LH}[n_1, n_2], \right.$$
$$\left. \hat{Y}_{hl,LH}[n_1, n_2], \hat{Y}_{lh,LH}[n_1, n_2] \right\} \xrightarrow{\text{IDWT}} \hat{Y}_{LH}[n_1, n_2]. \tag{12}$$

An estimate of the luminance $\hat{Y}[n_1, n_2]$ is achieved according to the following formula:

$$\left\{ \hat{Y}_{LL}[n_1, n_2], \hat{Y}_{HH}[n_1, n_2], \right.$$
$$\left. \hat{Y}_{HL}[n_1, n_2], \hat{Y}_{LH}[n_1, n_2] \right\} \xrightarrow{\text{IDWT}} \hat{Y}[n_1, n_2]. \tag{13}$$

Finally, the chrominance components are upsampled to the image dimension and combined with the estimated luminance thus obtaining the color image. Experimental results are presented in the next section.

FIGURE 5: For each row from left to right: original color image (24 bits/pixel), compressed image using the proposed approach (0.15 bits/pixel), compressed image using SPIHT (0.15 bits/pixel), and compressed image using the JPEG method (0.25 bits/pixel (maximum compression rate allowed by the JPEG coder)). First row: "Baboon," second row: "Biked," third row: "Lena," fourth row: "GoldHill."

TABLE 3: Bit rates employed for the different subbands with a global compression rate of 0.30 bpp.

	b_{tot}	b_{LL}	b_{HH}	b_{HL}	b_{LH}
Baboon	0.30	0.15	0.5	0.5	0.5
Biked	0.30	0.1625	0.055	0.0425	0.04
Lena	0.30	0.15	0.0375	0.0575	0.055
GoldHill	0.30	0.15	0.05	0.04	0.06

4. EXPERIMENTAL RESULTS AND CONCLUSIONS

In this section the effectiveness of the proposed method is discussed.

In Figures 5, 6, and 7, the compressed color images, obtained using the proposed approach are shown. The employed bpps values for the different subbands of the images under examination are reported in Table 3 for the case of compression at 0.30 bpp. For the sake of comparison, in Figures 5, 6, and 7, along with the original color images and their compressed replicas obtained using our approach, their JPEG and SPIHT compressed versions at different bit rates (0.15 bpp, 0.30 bpp, and 0.45 bpp), are also provided. It is worth pointing out that in Figure 5 the JPEG displayed images have been compressed at 0.25 bpp, instead of 0.15 bpp like the others in the same figure, since the JPEG coder does not allow further compression.

The assessment of the performance of our method requires the quantification of the perceptual error between two color images. To this end, it is crucial to adopt color spaces which are related to the perceptual characteristic of the human visual system thus allowing the definition of simple

FIGURE 6: For each row from left to right: original color image (24 bits/pixel), compressed image using the proposed approach (0.30 bits/pixel), compressed image using SPIHT (0.30 bits/pixel), and compressed image using the JPEG method (0.30 bits/pixel). First row: "Baboon," second row: "Biked," third row: "Lena," fourth row: "GoldHill."

metrics capable of properly measuring the perceptual distance between two colors. As is well known in literature [23], the RGB space, although widely used for different applications, is not suitable for accurate perceptual computations. More appropriate color spaces are the $L^*u^*v^*$ and $L^*a^*b^*$, standardized by the Commission Internationale de L'Enclairage (CIE) in 1976 as perceptually uniform. They are both equally good in providing a quantitative estimation of the perceptual distance between two colors. For our performance evaluations, we resort to use the $L^*a^*b^*$ color space where the L^* coordinate corresponds to the luminance, a^* corresponds to the red-green channel, and b^* to the blue-yellow channel. The conversion from the RGB space to the $L^*a^*b^*$ space is given in the appendix. The main property of the $L^*a^*b^*$ space is that color points at the same Euclidean distance are perceptually indistinguishable. On the other side, the perceptive distance between two colors is well

approximated by

$$\Delta E = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2}, \qquad (14)$$

where ΔE is the color error and $\Delta L^*, \Delta a^*$, and Δb^* the difference between the components L^*, a^*, and b^* components, respectively, of two colors under consideration.

Thus, the perceptive uniformity of $L^*a^*b^*$ is here exploited to evaluate the perceptual similarity between two color images, of dimension $N_1 \times N_2$, by computing the normalized color distance (NCD) [14] according to the following formula:

$$\text{NCD} = \frac{\sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} \Delta E[n_1, n_2]}{\sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} E[n_1, n_2]}, \qquad (15)$$
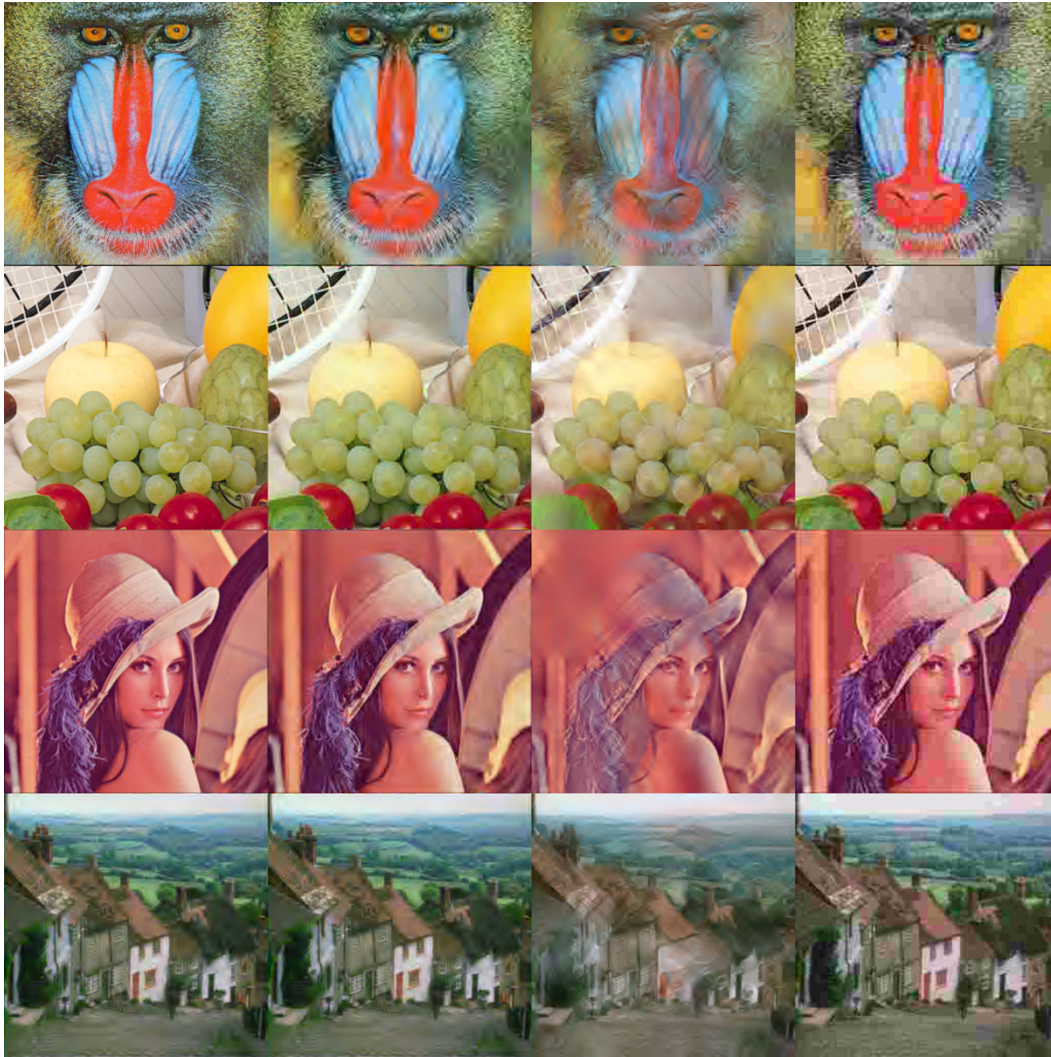
FIGURE 7: For each row from left to right: original color image (24 bits/pixel), compressed image using the proposed approach (0.45 bits/pixel), compressed image using SPIHT (0.45 bits/pixel), and compressed image using the JPEG method (0.45 bits/pixel). First row: "Baboon," second row: "Biked," third row: "Lena," fourth row: "GoldHill."

where $\Delta E[n_1, n_2]$ is given by (14) particularized to the colors of the two pixels in position $[n_1, n_2]$ of the two images under analysis and

$$E[n_1, n_2] = \sqrt{(L^*[n_1, n_2])^2 + (a^*[n_1, n_2])^2 + (b^*[n_1, n_2])^2} \qquad (16)$$

is the Euclidean norm of the pixel in position $[n_1, n_2]$ belonging to the uncompressed image.

This quantitative performance evaluation, at different bit rates (0.15 bpp, 0.30 bpp, and 0.45 bpp), is performed on the images obtained by applying our compressive data hiding approach, JPEG and SPIHT with respect to the original image and the results are shown in Table 4.

The quantitative evaluation of the performance of our method in comparison with JPEG and SPIHT is in agreement with a subjective evaluation performed on the images

displayed in Figures 5, 6, and 7, from which it is evident that the proposed method better performs than JPEG and outperforms SPIHT.

The pyramidal wavelet-like decomposition described in Section 3.4.1 also provides a tool to device a progressive coding strategy. In fact, in some applications such as supervised human retrieval from still images databases, it is not always necessary to retrieve, in one shot, a fully detailed color image because the user can decide at a first glance, from a rough gray-scale reproduction of the image, whether it is of interest. Therefore, the progressive approach could lead to dramatic bit saving in critical applications such as progressive display over a bandwidth limited wireless channel. Using our approach a first sketch of the image is obtained by transmitting only the $Y_{LL}[n_1, n_2]$ component; further details can be added by means of the $Y_{HH}[n_1, n_2]$ subband. Moreover, the transmission of $Y_{HL}^{(e)}[n_1, n_2]$ and $Y_{LH}^{(e)}[n_1, n_2]$ allows, not only to add

TABLE 4: NCD evaluation for the compression rates 0.15 bpps, 0.30 bpps, 0.45 bpps.

	Bit rate	Data hiding	SPIHT	JPEG
Baboon	0.15	0.1492	0.1985	0.19
	0.30	0.1368	0.1932	0.1793
	0.45	0.1225	0.1896	0.1545
Biked	0.15	0.0859	0.1213	0.1202
	0.30	0.0713	0.1112	0.0938
	0.45	0.0661	0.1052	0.0780
Lena	0.15	0.0932	0.1650	0.1753
	0.30	0.0807	0.1595	0.1504
	0.45	0.0695	0.1563	0.0881
GoldHill	0.15	0.1076	0.1725	0.1630
	0.30	0.0893	0.1030	0.1385
	0.45	0.0798	0.0963	0.0930

more details to the reconstructed image but even to recover the color information.

To summarize, in this paper, a progressive data hiding-based compression scheme, properly designed in order to trade off between the goals of data hiding and perceptual coding, is proposed. After having performed an unconventional two level wavelet decomposition of the luminance component of a color image, the perceptually irrelevant subbands are properly selected, zeroed, and replaced with a parsimonious representation of the chrominance components. This leads to a gray scale image in which the color information is piggybacked without impairing the overall perceptual quality of the embedded image. This gives the opportunity of viewing the image progressively from a monochrome version, at different details level, to a color one according to the user's needs. Moreover, our method allows to achieve better quality with respect to well consolidated coding schemes such as JPEG and SPIHT at low bit rates.

APPENDIX

$L^*a^*b^*$ COLOR SPACE

The transform from the RGB space to the $L^*a^*b^*$ space is as follows [19]:

$$X = 0.607R + 0.174G + 0.200B,$$
$$Y = 0.299R + 0.587G + 0.114B,$$
$$Z = 0.000R + 0.066G + 1.116B.$$
$$L^* = 25\left(\frac{100Y}{Y_0}\right)^{1/3} - 16 \quad 1 \leq 100Y \leq 100, \quad \text{(A.1)}$$
$$a^* = 500\left[\left(\frac{X}{X_0}\right)^{1/3} - \left(\frac{Y}{Y_0}\right)^{1/3}\right],$$
$$b^* = 200\left[\left(\frac{Y}{Y_0}\right)^{1/3} - \left(\frac{Z}{Z_0}\right)^{1/3}\right],$$

with reference to the white tristimulus values X_0, Y_0, Z_0.

REFERENCES

[1] J. Lacy, S. R. Quackenbush, A. R. Reibman, D. Shur, and J. H. Snyder, "On combining watermarking with perceptual coding," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, vol. 6, pp. 3725–3728, Seattle, Wash, USA, May 1998.

[2] H.-J. Wang and C.-C. J. Kuo, "An integrated progressive image coding and watermark system," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, vol. 6, pp. 3721–3724, Seattle, Wash, USA, May 1998.

[3] J. Meng and S.-F. Chang, "Embedding visible video watermarks in the compressed domain," in *Proc. IEEE Int. Conf. on Image Processing*, vol. 1, pp. 474–477, Chicago, Ill, USA, October 1998.

[4] S. Bhattacharjee and M. Kutter, "Compression tolerant image authentication," in *Proc. IEEE International Conference on Image Processing*, vol. 1, Chicago, Ill, USA, October 1998.

[5] T.-Y. Chung, M.-S. Hong, Y.-N. Oh, D.-H. Shin, and S.-H. Park, "Digital watermarking for copyright protection of mpeg2 compressed video," *IEEE Transactions on Consumer Electronics*, vol. 44, no. 3, pp. 895–901, 1998.

[6] J. Lacy, S. R. Quackenbush, A. R. Reibman, and J. H. Snyder, "Intellectual property protection systems and digital watermarking," *Optics Express*, vol. 3, no. 12, pp. 478–484, 1998.

[7] D. Kundur and D. Hatzinakos, "Mismatching perceptual models for effective watermarking in the presence of compression," in *Proc. SPIE, Multimedia Systems and Applications II*, vol. 3845, pp. 29–42, September 1999.

[8] D. Kundur, "Energy allocation principles for high capacity data hiding," in *Proc. IEEE Int. Conf. on Image Processing*, vol. 1, pp. 423–426, Vancouver, Canada, September 2000.

[9] C. Fei, D. Kundur, and R. H. Kwong, "The choice of watermark domain in the presence of compression," in *Proc. IEEE Int. Conf. on Information Technology: Coding and Computing*, pp. 79–84, April 2001.

[10] B. Zhu and A. H. Tewfik, "Media compression via data hiding," in *Thirty-First Asilomar Conf. on Signals, Systems, and Computers*, vol. 1, pp. 647–650, 1997.

[11] N. Jayant, J. Johnston, and R. Safranek, "Signal compression based models of human perception," *Proceedings of the IEEE*, vol. 81, pp. 1383–1422, October 1993.

[12] A. Said and W. A. Pearlman, "A new fast and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 6, pp. 243–250, 1996.

[13] S. Mallat, "Multifrequency channel decompositions of images and wavelet models," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 37, pp. 2091–2110, 1989.

[14] K. N. Plataniotis and A. N. Venetsanopoulos, *Color Image Processing and Applications*, Springer-Verlag, Berlin, 2000.

[15] M. Ramkumar and A. N. Akansu, "Theoretical capacity measures for data hiding in compressed images," in *Proc. SPIE, Multimedia Systems and Applications*, vol. 3528, pp. 482–492, Boston, Mass, USA, November 1998.

[16] S. Mallat, "Wavelet for a vision," *Proceedings of the IEEE*, vol. 84, no. 4, pp. 604–614, 1996.

[17] I. J. Cox, J. Kilian, T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Trans. Image Processing*, vol. 6, no. 12, pp. 1673–1687, 1997.

[18] F. Hartung and M. Kutter, "Multimedia watermarking techniques," *Proceedings of IEEE*, vol. 87, no. 7, pp. 1079–1107, 1999.

[19] A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1989.

[20] B. Chen and G. W. Wornell, "Quantization index modulation: a class of provably good methods for digital watermarking and

information embedding," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1423–1443, 2001.

[21] T. Cover and J. Thomas, *Elements of Information Theory*, John Wiley & Sons, Toronto, Canada, 1991.

[22] J. Li, P.-Y. Cheng, and C.-C. J. Kuo, "A wavelet transform approach to video compression," in *Proc. SPIE Wavelet Applications II*, Orlando, Fla, USA, April 1995.

[23] A. K. Jain, "Color distance and geodesics in color 3 space," *Journal of the Optical Society of America*, vol. 62, no. 11, pp. 1287–1290, 1972.

Patrizio Campisi received the "Laurea" degree in Electrical Engineering, "summa cum laude," at the University of Roma "La Sapienza," Roma, Italy, and received his Ph.D. degree in Electrical Engineering from the University of Roma "Roma Tre," Roma, Italy, in 1995 and 1999, respectively. In 1997 and 2000, he was a visiting research associate at the Communication Laboratory of the University of Toronto, Canada. He presently holds an associate research position at the University of Roma "Roma Tre," where he is also lecturer for the graduate course in Signal Theory. His research interests are in the area of digital signal and image processing with applications to wireless communications and multimedia.

Deepa Kundur is an Assistant Professor in the Edward S. Rogers Sr. Department of Electrical and Computer Engineering at the University of Toronto. She holds the title of Bell Canada Junior Chair-holder in Multimedia and is also an Associate of the Nortel Institute for Telecommunications. She received her B.A.Sc., M.A.Sc., and Ph.D. degrees from the Electrical and Computer Engineering Department at the University of Toronto. Prof. Kundur's research interests span the areas of multimedia security, data hiding and covert communications, and nonlinear and adaptive communication algorithms. Deepa is a member of the IEEE (Communications and Signal Processing Societies) and the Professional Engineers of Ontario (PEO).

Dimitrios Hatzinakos received the Diploma degree from the University of Thessaloniki, Greece, in 1983, the M.A.Sc. degree from the University of Ottawa, Canada, in 1986 and the Ph.D. degree from Northeastern University, Boston, M.A., in 1990, all in Electrical Engineering. In September 1990 he joined the Department of Electrical and Computer Engineering, University of Toronto, where now he holds the rank of Professor with tenure. Also, he serves as Chair of the Communications Group of the Department since 1 July, 1999. His research interests are in the areas of digital communications and signal processing with applications to wireless communications, image processing and multimedia. He has organized and taught many short courses on modern signal processing framework and applications devoted to continuing engineering education and given numerous seminars in the area of blind signal deconvolution. He is author/co-author of more than 100 papers in technical journals and conference proceedings and he has contributed to 5 books in his areas of interest. His experience includes consulting through Electrical Engineering Consocates Ltd., and contracts with United Signals and Systems Inc., Burns and Fry Ltd., Pipetronix Ltd., Defense Research Establishment Ottawa (DREO), Vaytek Inc., Nortel Networks, and Vivosonic Inc. He is an Associate Editor for the IEEE Transactions on Signal Processing since July 1998. Also, he was the Guest Editor for the special issue of Signal Processing, Elsevier, on Signal Processing Technologies for Short Burst Wireless Communications which appeared in October 2000. He was a member of the IEEE Statistical Signal and Array Processing Technical Committee (SSAP) from 1992 till 1995 and Technical Program co-Chair of the 5th Workshop on Higher-Order Statistics in July 1997. He is a senior member of the IEEE and member of EURASIP, the Professional Engineers of Ontario (PEO), and the Technical Chamber of Greece.

Alessandro Neri was born in Viterbo (Italy) in 1954. In 1977 he received the Doctoral Degree in Electronic Engineering from the University of Rome "La Sapienza." In 1978 he joined the Research and Development Department of Contraves Italiana S.p.a. where he gained a specific expertise in the field of radar signal processing and in applied detection and estimation theory, becoming the chief of the advanced system group. In 1987, he joined the INFOCOM Department of the University of Rome "La Sapienza" as Associate Professor in Signal and Information Theory. In November 1992, he joined the Electronic Department of the University of Rome "Roma Tre" as Associate Professor in Electrical Communications. Since September 2001 he is full professor in Telecommunications at the University of Rome "Roma Tre." Since 1987, his research activity has mainly been focused on information theory, signal theory, and signal and image processing and their applications to both telecommunications systems and remote sensing.