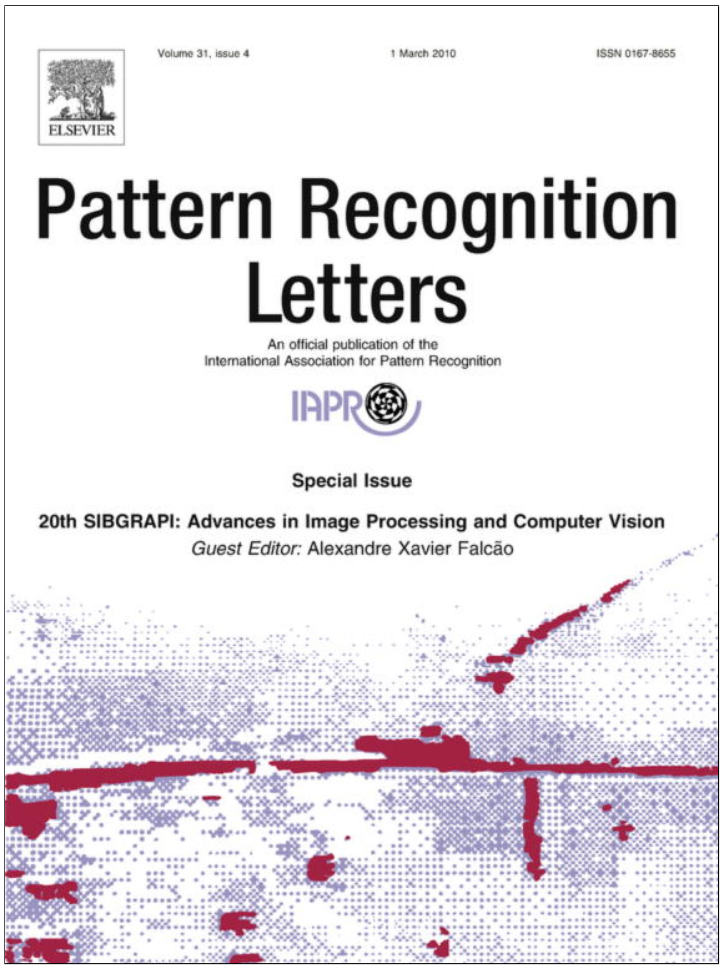

Pattern Recognition Letters 31 (2010) 269–276

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Reversible color-to-gray mapping using subband domain texturization

Ricardo L. de Queiroz*

Departamento de Engenharia Elétrica, Universidade de Brasília, CP 04591, 70910-900 Brasília, DF, Brazil

ARTICLE INFO

Article history
Available online 10 December 2008

Keywords:
Color to gray
Reversible color mapping
Color embedding
Subband transforms

ABSTRACT

The concept of reversible conversion of color images to gray ones has been recently introduced. Colors are mapped to textures and from the textures the receiver can recover the colors. This was done using the wavelet transform and replacing high-frequency subbands by the down-sampled chrominance planes. The idea is to print a color image with a black and white printer and, then, at a later time, to scan the document and recover colors. In this paper, we propose to use a more robust method, i.e. more resistant to the print-scan noise. We propose to use a largely redundant representation of the chrominance, embedding them into multiple subbands of a general subband transform. In other words, we spread the chrominance onto many subbands. We show theoretically that for minimizing the variance of the error caused by white noise, the chrominance should be replicated into many subbands and not linearly combined. We also demonstrate the method to find the best linear transform to embed the chrominance into the subbands in the more general case of colored noise. We carry a noise analysis to determine bounds to guide us on how many subbands into which to embed the chrominance. Experimental results were carried involving real printing and scanning, as well as using a simulated print-scan path.

1. Introduction

Recently, an innovative method to convert color images into gray ones was introduced (de Queiroz and Braun, 2006). Its key feature was to be reversible. The color image is converted to gray scales and one can retrieve the colors from the gray image at a later time. The application is for a user who created a color document and just has a black and white device to produce or transmit the document. For example, the user may only have easy access to monochrome office laser printers, or the user might just have to fax the color document. Thus, the color document is first converted to gray scales, then to black and white, before printing (or faxing). Later on, the user might scan (receive) the black and white document, recover the monochrome (gray) image, process it, and retrieve the colors. The process is illustrated in Fig. 1 and has been explored elsewhere (Ko et al., 2007).

Even though the proposition sounds unrealistic the method is actually fairly simple. The idea is to smoothly map colors to high frequency textures. By analysing the textures, the user retrieves the colors for each region. In the originally proposed method (de Queiroz and Braun, 2006), one has to:

1. Convert the color image to some luminance-chrominance color space such as YCbCr or CIELab (Hunt, 2000).
2. Apply the discrete wavelet transform (DWT) to the luminance channel (Vetterli and Kovacevic, 1995; Strang and Nguyen, 1996).
3. Spatially reduce (subsample) the 2 chrominance planes by a factor of 2 in each direction.
4. Replace the high-frequency wavelet subbands (Strang and Nguyen, 1996) (HL and LH) by the 2 chrominance planes.
5. Apply the inverse DWT yielding a texturized gray image (because of the embedded chrominance planes acting as subband coefficients).
6. Scale, halftone (Roetling and Loce, 1994) and print (or transmit) the texturized gray image.

The above method is illustrated in Fig. 2.

In order to recover the color from the gray texturized image, as illustrated in Fig. 3, one has to:

1. Scan (receive) the image and convert it to gray scale.
2. Apply the DWT to the gray image.
3. Retrieve the high frequency subbands and assign them as chrominance planes.
4. Set to zero the subbands which were used to embed chrominance.
5. Apply an inverse DWT to the resulting subbands, yielding the luminance plane.
6. Spatially increase (up-sample) the 2 chrominance planes by a factor of 2 in each direction.
7. Convert the luminance-chrominance planes back into RGB.

* Tel.: +55 61 99680964; fax: +55 61 32746651.
E-mail address: queiroz@ieee.org

0167-8655/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2008.11.010

Fig. 1. Illustration of the color to gray conversion application. The color image is printed in a black and white printer. Later on, the user may scan the printed document and recover the original colors, if so desired.

In essence, colors are converted to strong high-frequency patterns (textures). For the user they look like halftones or similar textures that the human visual system tends to ignore and blend with the object. Examples are shown in Fig. 4. The method works so well because the embedding is natural, since the colors match the objects. If we coded the chrominance and embedded the binary information into the image using any typical image watermarking technique (Cox et al., 2002), it would probably not work. The encoded information does not correlate with the image contents. In order to avoid artifacts one would have to make it subtle or invisible. Then, it is likely that the information would be removed by the halftoning, printing and scanning processes. If we make the embedded information strong enough to survive halftoning it would likely cause artifacts. At the moment, despite the work in this direction (Baharav and Shaked, 1999; Chiang et al., 2006; de Queiroz et al., 2005; Fu and Au, 2000; Knox and Wang, 1997; Liu et al., 2004; Sharma et al., 2003; Venkata et al., 2005; Villan et al., 2007; Voloshynovskiy et al., 2006; Wang and Knox, 2000) we do not have reliable non-intrusive watermarking methods for printed images, without controlling printer characteristics such as the halftone algorithm or the laser beam intensity. The proposed method produces strong but pleasant patterns, making it an excellent candidate for the application.

A preliminary version of this work was presented elsewhere (de Queiroz, 2007). Here, we include substantial theorectical results, including two theorems, comparison among subband transforms, and real print-scan tests.

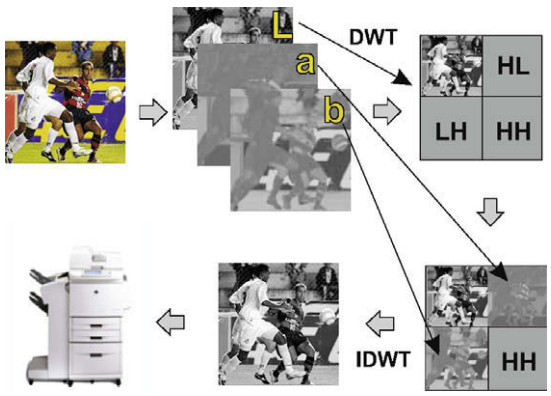

Fig. 2. The color to gray conversion method. The luminance is decomposed into subbands through a wavelet transform. The high pass bands are replaced by the chrominance planes and the image is inverse transformed generating a texturized gray image which conveys the color information.

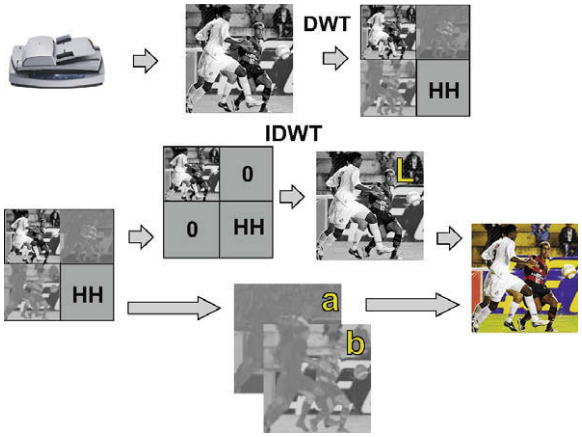

Fig. 3. The color recovery method. After scanning the image, the texturized image undergoes a DWT, the chrominance planes are extracted, the corresponding subbands are zeroed, and an inverse DWT is carried on the result, yielding the luminance channel. The luminance and chrominance channels are used to recover the color image.

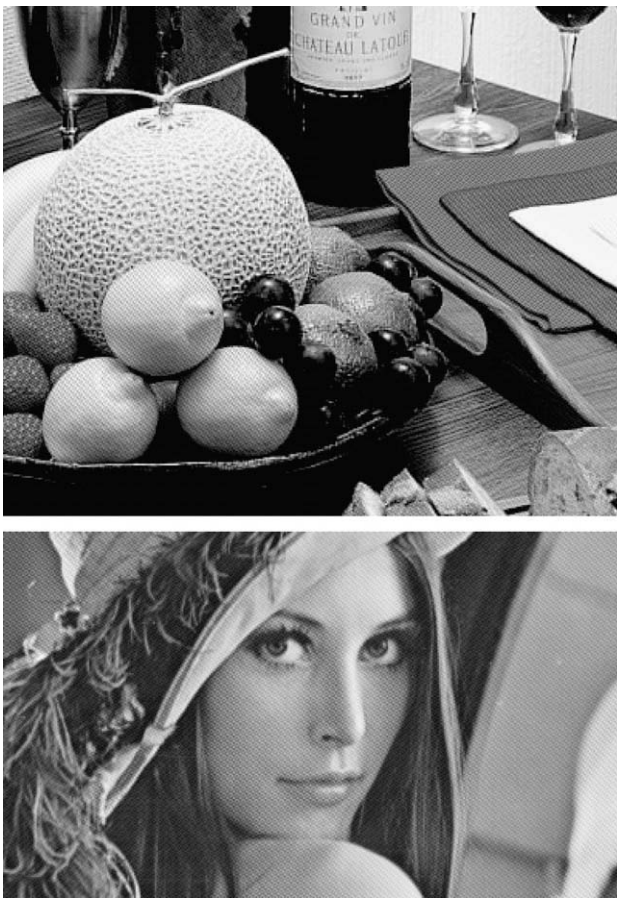

Fig. 4. Zoom of two example texturized images using the proposed method, with a 4×4 DCT and $N = 6$. The high frequency patterns tend to interfere with sampling grids resulting in annoying artifacts when displaying those texturized images at lower resolution. Changing the viewing resolution may change the artifacts. PDF conversion also causes artifacts. Please refer to the original bitmaps at http://image.unb.br/queiroz/texture to verify how they really look like.

2. Embedding chrominance into subbands

The method presented in the previous section is simplistic. If the gray image was to be scaled up, halftoned and faxed, it would

work very well, allowing the recovery of vivid colors. If the image is to be printed, the noise and errors are so intense that the simple method presented would not work. Any minimal shift of a pixel or rotations of one degree may completely impair the color recovery (de Queiroz and Braun, 2006). As a result, in the original work (de Queiroz and Braun, 2006), some non-linear redundancy was introduced in the wavelet domain to prevent sign inversion and to reduce noise sensitivity, with some success.

We propose to use a more aggressive redundancy scheme to cope with noise. Indeed, we propose to:

- Use general M-band subband transforms rather than the DWT, allowing the use of the discrete cosine transform (DCT) (Rao and Yip, 1990) or general lapped transforms (de Queiroz and Tran, 2001; Malvar, 1992).
- Embed the chrominance planes into a linear combination of multiple subbands per chrominance channel, as illustrated in Fig. 5.
- Decide upon the redundancy based on a signal to noise model analysis of the embedding mechanisms.

The questions that need to be answered before we apply such a method are:

(a) What subband transform should we use?
(b) How many channels M should it have?
(c) How many subbands N_s should we use per chrominance?
(d) How do we decide upon the linear combination of the subband, i.e. how do we distribute the chrominance information onto the N_s channels?

An M-band subband transform (a) can be constructed through the general association of 2-band filter banks (i.e. discrete wavelet packets) or directly. Both approaches are roughly equivalent in this context, with the latter having faster algorithms. We can use the DCT, the lapped orthogonal transform (LOT), or the modulated (extended) lapped transform (MLT or ELT), (Malvar, 1992). The choice on the number of bands (question b), we believe to be a function of N_s. We do not want to use too many subbands to carry chrominance information, because it may impair the quality of the luminance information. For example, one can use a 4×4-band ($M = 16$) transform to embed up to 10 bands ($N_s = 5$). Higher numbers may be accommodated into an 8×8 transform. Note that the larger M is, the better the luminance quality, but the worst the chrominance representation since the chrominance planes must be sub-sampled accordingly to fit the subbands. Hence, we want M to be as small as possible as long as M is substantially larger than N_s. A detailed analysis on deciding upon N_s will be carried later on and questions (c,d) will be answered in the following sections.

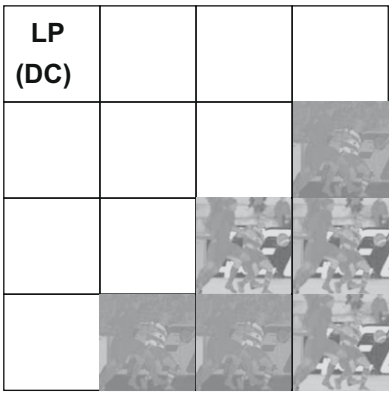

Fig. 5. Illustration of the proposed method, showing an example of the subbands of a 4×4 transform, indicating the low-pass or DC band. In this example, 3 copies of each chrominance channel were alternately embedded into the highest frequency subbands.

3. Multiple description efficiency

Let the mapping from M to N source samples be linear. In other words, the information M-tuple $\mathbf{c} = [c_0, c_1, \ldots, c_{M-1}]^T$ is mapped into an N-tuple vector $\mathbf{y} = [y_0, \ldots, y_{N-1}]^T$ as

$$\mathbf{y} = \mathbf{A}\mathbf{c}, \tag{1}$$

where $\mathbf{A}$ is an $N \times M$ matrix. We assume that $\mathbf{A}$ has full rank and $N > M$. The $\{y_i\}$ are then conveyed to the decoder, being corrupted by noise in the process. The decoder has to try to recover the information, i.e. the $\{c_i\}$, from the noisy version of $\{y_i\}$.

Theorem 1. *Let the noise be zero-mean white and additive, i.e.* $\hat{\mathbf{y}} = \mathbf{y} + \mathbf{n}$ *where* $\mathbf{n}$ *is a noise vector. The signal is reconstructed linearly as*

$$\hat{\mathbf{c}} = \mathbf{B}\hat{\mathbf{y}} \tag{2}$$

and the error is $\mathbf{e} = \hat{\mathbf{c}} - \mathbf{c}$. *The average reconstruction error power is* $\sigma_e^2 = E\{\frac{1}{N}\mathbf{e}^T\mathbf{e}\}$. *Optimal recovery, in the sense of minimizing* σ_e^2, *is achieved when the columns of* $\mathbf{A}$ *form an orthogonal basis.*

Proof 1. – Since

$$\hat{\mathbf{c}} = \mathbf{B}(\mathbf{y} + \mathbf{n}) = \mathbf{B}\mathbf{A}\mathbf{c} + \mathbf{B}\mathbf{n},$$

then, if we define $\boldsymbol{\Phi} = \mathbf{B}\mathbf{A} - \mathbf{I}$,

$$\mathbf{e} = (\mathbf{B}\mathbf{A} - \mathbf{I})\mathbf{c} + \mathbf{B}\mathbf{n} = \boldsymbol{\Phi}\mathbf{c} + \mathbf{B}\mathbf{n},$$

so that the cost function $J = N\sigma_e^2 = E\{\mathbf{e}^T\mathbf{e}\}$ is

$$J = E\{\mathbf{c}^T\boldsymbol{\Phi}^T\boldsymbol{\Phi}\mathbf{c} + \mathbf{c}^T\boldsymbol{\Phi}^T\mathbf{B}\mathbf{n} + \mathbf{n}^T\mathbf{B}^T\boldsymbol{\Phi}\mathbf{c} + \mathbf{n}^T\mathbf{B}^T\mathbf{B}\mathbf{n}\} \tag{3}$$

Since the noise is uncorrelated with $\mathbf{c}$ and has zero-mean, $E\{\mathbf{n}\} = \mathbf{0}$ and the second and third terms, above, are zero. The first term is quadratic and positive semi-definite. Hence, it will be minimal if null. In other words, if

$$\mathbf{B} = \mathbf{A}^+ = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T, \tag{4}$$

then $\boldsymbol{\Phi} = 0$ and the first term is minimized. The choice of $\mathbf{B}$ impacts the first and last terms. The first term attains its absolute minimum and the last is not restricted by the choice of $\mathbf{B} = \mathbf{A}^+$. Since $\mathbf{A} = \mathbf{B}^- = \mathbf{B}^T(\mathbf{B}\mathbf{B}^T)^{-1}$, then all degrees of freedom are maintained, i.e. every full rank matrix $\mathbf{A}$ corresponds to a full rank matrix $\mathbf{B}$. We are then left with minimizing the last term. We assumed the noise to be white and stationary, i.e. $\mathbf{R}_{nn} = \sigma_n^2\mathbf{I}$, so that we can rewrite the cost as

$$\begin{aligned} J =& E\{\mathbf{n}^T\mathbf{B}^T\mathbf{B}\mathbf{n}\} = Tr\{\mathbf{B}E\{\mathbf{n}\mathbf{n}^T\}\mathbf{B}^T\} \\ =& Tr\{\mathbf{B}\mathbf{R}_{nn}\mathbf{B}^T\} = \sigma_n^2 Tr\{\mathbf{B}\mathbf{B}^T\} \\ =& \sigma_n^2 Tr\{(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\} \\ =& \sigma_n^2 Tr\{(\mathbf{A}^T\mathbf{A})^{-1}\}. \end{aligned} \tag{5}$$

If $\mathbf{A} = [v_0, v_1, \ldots, v_{M-1}]$, then $\mathbf{A}^T\mathbf{A}$ has elements $\{\ell_{ij}\}$ such that $\ell_{ij} = v_i^T v_j$. The variance and energy of the input c_i can be normalized without loss of generality. Note that $\mathbf{y} = c_1 v_1 + \ldots + c_{M-1}v_{M-1}$ and we can normalize the input signals c_i and compensate any normalization through scaling the $\{v_i\}$. Hence, we can assume, without sacrificing generality, that the $\{v_i\}$ have unity norm, i.e. $\ell_{ii} = 1$ for $0 \leqslant i < M$.

As a symmetric matrix, the singular value decomposition of $\mathbf{A}^T\mathbf{A}$ is $\mathbf{A}^T\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T$, where $\mathbf{U}$ is some unitary matrix and $\boldsymbol{\Lambda}$ is a full-rank diagonal matrix with real positive diagonal entries λ_k (Horn and Johnson, 1985). Then,

$$J/\sigma_n^2 = Tr\{(\mathbf{A}^T\mathbf{A})^{-1}\} = \sum_{k=0}^{M-1}\frac{1}{\lambda_k}. \quad (6)$$

Note that

$$Tr\{\mathbf{A}^T\mathbf{A}\} = \sum_{k=0}^{M-1}\lambda_k = \sum_{k=0}^{M-1}\ell_{kk} = M. \quad (7)$$

As a result, with the constraint in (7), then the cost in (6) is minimized iff all $\lambda_i = 1$. This implies that $\mathbf{A}^T\mathbf{A} = \mathbf{I}$, i.e. that the set of $\{v_i\}$ are orthogonal, hence form an orthonormal basis. **Q.E.D.**

An important consequence of the above theorem is that not only the columns of $\mathbf{A}$ should be orthogonal, but any set of orthogonal vectors leads to an optimal choice of $\mathbf{A}$. Being that the case, one can chose the simplest possible orthogonal set. For example an incomplete permutation set, i.e. the v_i are formed by ones and zeros only. Another implication of the theorem is that $\mathbf{B} = \mathbf{A}^+ = \mathbf{A}^T$.

Hence, the optimal transform can be made deterministic and signal independent: create the matrix $[\mathbf{I}_M\mathbf{I}_M\cdots]^T$ by replicating the $M \times M$ identity matrix as many times as necessary to generate more than N rows. Then, set $\mathbf{A}$ as a truncated version of the replicated matrix to have exactly N rows. Finally, set $\mathbf{B} = \mathbf{A}^T$. In summary, the multiple description can be optimally achieved through replication rather than linear combination.

If we set the optimized values $\lambda_k = 1$ in (6), then the error variance is given by

$$\sigma_e^2 = \frac{M\sigma_n^2}{N}. \quad (8)$$

In other words, the error is distributed evenly across the redundant information. The more redundancy, the lower the error, as expected.

In the coloured (correlated) noise case, the solution is less incisive, nevertheless equally interesting. First, one needs to readdress the problem. In (5) the term to be minimized is $Tr\{\mathbf{B}\mathbf{R}_{nn}\mathbf{B}^T\}$. However, this is only part of the problem because there are constraints to obey in order to make the solution useful. We know $\mathbf{B}$ needs to be inverted, hence to have full-rank. Without further restraints, in order to minimize the cost fucntion, one needs only to find a full-rank matrix $\mathbf{B}$ and multiply it by a sufficiently small number $\epsilon \to 0$. However, since $\mathbf{A} = \mathbf{B}^-$, $\mathbf{y}$ would contain huge numbers, proportional to $1/\epsilon$, and the method would not be very useful. One potential solution is to reduce the spread of the non-zero eigenvectors of $\mathbf{B}$ and $\mathbf{A}$. Alternatively, one can simply limit the condition number (Horn and Johnson, 1985) (η) of $\mathbf{B}\mathbf{B}^T$ and normalize the range of its eigenvalues from $1/\sqrt{\eta}$ to $\sqrt{\eta}$, for $\eta \geqslant 1$. If we limit $\eta = 1$, then the rows of $\mathbf{B}$ are orthonormal to each other, and so are those of $\mathbf{A}$.

Theorem 2. *Let the noise be zero-mean additive with correlation matrix $\mathbf{R}_{nn}$. Then, in the conditions of Theorem 1, the orthonormal matrix $\mathbf{A}$ that yields optimal recovery, in the sense of minimizing σ_e^2, is the pseudo-inverse of a matrix with the eigenvectors of $\mathbf{R}_{nn}$ corresponding to its M smallest eigenvalues.*

Proof 2. $\mathbf{R}_{nn}$ is a Hermitian full-rank matrix and from previous results in trace inequalities for orthogonal regression problems (Coope and Renaud, 2000), we know that the trace of $\mathbf{B}\mathbf{R}_{nn}\mathbf{B}^T$ will be minimized when $\mathbf{B}$ is made of the eigenvectors corresponding to M smallest eigenvalues of $\mathbf{R}_{nn}$. From the derivation of Theorem 1 we also know that $\mathbf{A}$ is the pseudo inverse of $\mathbf{B}$. **Q.E.D.**

For a more general condition number η, then

$$J = Tr\{\mathbf{B}\mathbf{R}_{nn}\mathbf{B}^T\} = Tr\{\mathbf{R}_{nn}\mathbf{B}\mathbf{B}^T\}$$

Let $\mathbf{U}\Sigma\mathbf{U}^T$ be the singular value decomposition of $\mathbf{R}_{nn}$, i.e. the eigenvectors of $\mathbf{R}_{nn}$ are in $\mathbf{U}$ while its eigenvalues are the diagonal elements of Σ. Same thing for the decomposition of $\mathbf{B}\mathbf{B}^T = \mathbf{V}\mathbf{S}\mathbf{V}^T$. Then, if $\mathbf{W} = \mathbf{U}^T\mathbf{V}$, then

$$J = Tr\{\mathbf{U}\Sigma\mathbf{U}^T\mathbf{V}\mathbf{S}\mathbf{V}^T\} = Tr\{\Sigma\mathbf{W}\mathbf{S}\mathbf{W}\}.$$

The solution can be found through unconstrained optimization, by factorizing $\mathbf{W}$ into $M(M-1)/2$ Givens rotations (Golub and Van Loan, 1989) and limiting the search of the M diagonal elements of $\mathbf{S}$ within the $[1/\sqrt{\eta}, \sqrt{\eta}]$ range. In other words, a search over $M(M+1)/2$ parameters. Once $\mathbf{W}$ and $\mathbf{S}$ are set, $\mathbf{V}$ and $\mathbf{B}\mathbf{B}^T$ can be found. From $\mathbf{B}\mathbf{B}^T$, the factorization of $\mathbf{B}$ is not unique and can be found as

$$\mathbf{B} = \mathbf{V}\mathbf{S}_r\boldsymbol{\Psi},$$

where $\mathbf{S}_r$ is a diagonal matrix with the positive square roots of the diagonal entries of $\mathbf{S}$, and $\boldsymbol{\Psi}$ can be any orthogonal matrix. Finally, $\mathbf{A} = \mathbf{B}^T(\mathbf{B}\mathbf{B}^T)^{-1}$. □

4. Noise analysis and the amount of redundancy

Let us move closer to our original problem of embedding chrominance into luminance subbands. The input signal is supposed to be a triplet $\mathbf{x}(n_1,n_2) = [l(n_1,n_2), a(n_1,n_2), b(n_1,n_2)]^T$, where $\{l(i,j)\}$ denotes the luminance image pixels and a, b are the same for the chrominance image planes. We will drop the pixel index (n_1,n_2) wherever convenient. The recovered image is $\hat{\mathbf{x}}$. The distance between colors is taken as the Euclidean distance between points in the Lab color space for every pixel (Hunt, 2000), i.e.

$$\begin{aligned}\sigma_d^2 &= E\{||\hat{\mathbf{x}} - \mathbf{x}||^2\}\\ &= E\{|\hat{l} - l|^2\} + E\{|\hat{a} - a|^2\} + E\{|\hat{b} - b|^2\}\\ &= \sigma_{el}^2 + \sigma_{ea}^2 + \sigma_{eb}^2. \end{aligned} \quad (9)$$

Let us use an $m \times m$ transform, i.e. $M = m^2$. Note that the chrominance planes have a resolution that is m times larger than any of the subbands in each direction. Hence, the chrominance planes are to be decimated and will suffer losses. If we ignore that and assume the noise that will affect the gray-scale (texturized) image is white with zero-mean and variance σ_n^2, and the subband transform is orthogonal, then σ_{ea} and σ_{eb} would be given by (8). However, to account for sub-sampling we should add a term to that.

Let us model the signal as a separable autoregressive process of first order, AR(1), with autocorrelation $r_c(n_1,n_2) = \sigma_c^2\rho^{|n_1|}\rho^{|n_2|}$, and corresponding power spectral density $S_c(e^{j\omega_1})S_c(e^{j\omega_2})$. If the anti-aliasing filter is perfect, the sub-sampling by a factor of m in each direction would remove all frequency contents above π/m. Hence, sub-sampling causes an error whose variance is

$$\sigma_{ss}^2 = \sigma_c^2 - \left(\int_0^{\pi/m} S_c(e^{j\omega})d\omega\right)^2. \quad (10)$$

For an AR(1) signal with unit variance, we can show that:

$$\int_\alpha^\beta S_c(e^{j\omega})d\omega = \frac{2}{\pi}arctan\left[\frac{1+\rho}{1-\rho}tan\left(\frac{\omega}{2}\right)\right]\Bigg|_{\omega=\alpha}^{\omega=\beta}. \quad (11)$$

Let

$$\varphi(\rho,m) = \frac{2}{\pi}arctan\left[\frac{1+\rho}{1-\rho}tan\left(\frac{\pi}{2m}\right)\right]. \quad (12)$$

Thus,

$$\sigma_{ss}^2 = \sigma_c^2 - \sigma_c^2\varphi^2(\rho,m). \quad (13)$$

Note that if the signal was uncorrelated, i.e. $\rho = 0$, then the noise variance would reduce to

$$\sigma_{ss}^2 = \sigma_c^2 \frac{M-1}{M}. \tag{14}$$

However, the chrominance image is actually supposed to have a correlation coefficient very close to 1. Then,

$$\sigma_{ea}^2 = \sigma_{eb}^2 = \frac{2\sigma_n^2}{N} + \sigma_c^2(1-\varphi^2(\rho,m)) \tag{15}$$

If there was no embedding, $\sigma_{el}^2 = \sigma_n^2$, but that is not true in general. When we use one subband to embed a chrominance plane or its replica, we actually erase the luminance in that subband. So, as the distortion decreases because we improve chrominance reproduction, at the same time distortion also increases because luminance reproduction suffers.

Without loss of generality, let us number the M subbands one-dimensionally, arranging them such that the embedded luminance subbands would be the last N. Let the resulting filter for the kth luminance subband be $f_k(n_1,n_2)$, be it the result of the cascade of filters like in wavelet packets, or not. Then, the variance of each luminance subband is

$$\sigma_{s_k}^2 = \frac{1}{\pi^2}\int\int_0^\pi |F_k(e^{j\omega_1},e^{j\omega_2})|^2 S_l(e^{j\omega_1},e^{j\omega_2})d\omega_1 d\omega_2, \tag{16}$$

and because of the assumed orthogonality,

$$\sigma_l^2 = \frac{1}{M}\sum_{k=0}^{M-1}\sigma_{s_k}^2. \tag{17}$$

As we embed the chrominance into the last N subbands, such that these subbands disappear, the distortion on the luminance becomes:

$$\sigma_{el}^2 = \sigma_n^2 + \frac{1}{M}\sum_{k=M-N}^{M-1}\sigma_{s_k}^2. \tag{18}$$

Now that we have computed the distortion for all planes, the signal-to-noise ratio (SNR) of the embedding process can be written as:

$$SNR = \frac{\sigma_l^2+\sigma_a^2+\sigma_b^2}{\sigma_{el}^2+\sigma_{ea}^2+\sigma_{eb}^2} = \frac{\sigma_s^2}{\sigma_d^2}. \tag{19}$$

Since σ_s^2 is a property of the image itself, and is independent of M and N, we are concerned with

$$\begin{aligned}\sigma_d^2 &= \sigma_n^2 + \frac{1}{M}\sum_{k=M-N}^{M-1}\sigma_{s_k}^2 + 2\frac{2\sigma_n^2}{N} \\ &\quad + 2\sigma_c^2(1-\varphi^2(\rho,m)) \\ &= \left(1+\frac{4}{N}\right)\sigma_n^2 + \frac{1}{M}\sum_{k=M-N}^{M-1}\sigma_{s_k}^2 \\ &\quad + 2(1-\varphi^2(\rho,m))\sigma_c^2.\end{aligned} \tag{20}$$

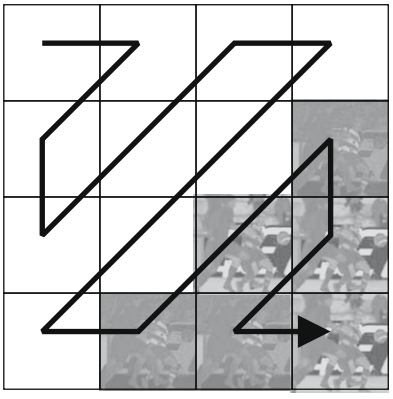

Fig. 6. Zigzag ordering of subbands.

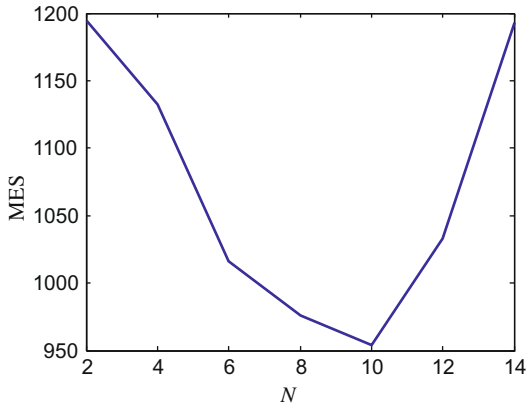

Fig. 7. MSE plot for a 4×4 DCT embedding framework. Printing simulation used $K=4$ and $\sigma_p=0$.

Note that there are three terms contributing to noise. One depends on the system noise σ_n^2, the other on the luminance subband variances and the last on the chrominance energy σ_c^2. The latter does not depend on N. The first term decreases and the second grows with N. So the question we ask ourselves, and is the purpose of this Section, is whether it is worthy to embed more chrominance replicas into more luminance subbands? More specifically, if we have already N subband embedded, is it worthy to embed two more, one for each chrominance channel? If we grow the number of subbands from N to $N+2$ the chrominance error (apart from sub-sampling) will decrease from $\sigma_n^2\frac{4}{N}$ to $\sigma_n^2\frac{4}{N+2}$. In the other hand, the luminance error variance will increase by the sum of the variances of subbands s_{M-N-1} and s_{M-N-2} scaled by $1/M$. Thus, it is easy to see that it is only advantageous to embed chrominance into two more subbands in terms of SNR if

$$\sigma_n^2 > \frac{N(N+2)}{8M}(\sigma_{s_{M-N-1}}^2 + \sigma_{s_{M-N-2}}^2). \tag{21}$$

The above condition tells us that if the textured image suffers small noise levels, it is not advantageous to add chrominance redundancy in expense of luminance quality. In the extreme case of noise free transmission, one can perfectly recover both chrominance channels with only two subbands. No redundancy is necessary. As the noise increases, the chrominance information disappears as noise corrupts the textures, and it becomes necessary to add redundancy (more subbands) to protect the channel. This equation also tells us that, as N grows, it becomes harder and harder to get any benefit from embedding even more subbands. Of course, it all depends on the variance of the luminance subbands that will be discarded. The more intense their energy, the more intense the noise has to be to make it worthy embed more subbands.

Table 1
Comparison of MSE results for embedding chrominance using simulated printing and many transforms ($M=16, N=10$).

Transform	Image "wine" σ_p^2			Image "Lena" σ_p^2		
	10	25	50	10	25	50
DCT	962	998	1122	811	850	998
MLT	1033	1075	1180	929	969	1118
ELT	1028	1067	1196	936	975	1114
LOT	1018	1058	1187	923	962	1101
LBT	1148	1182	1265	958	990	1089
DWT	1258	1306	1516	1017	1078	1302

MSE computed in RGB space.

5. Experiments

The condition in (21) is image dependent and hard to compute, since it is not trivial to determine σ_n^2. We have assumed the print-scan distortion to be additive white Gaussian noise. It is a reasonable model if geometric distortions, caused by paper warping and misplacement, and halftoning, can be at least partially corrected after scans. The results will be heavily dependent on the correction

Fig. 8. (Top) Original color images "mandrill" and "Barbara", (middle) their respective texturized grey versions, (bottom) and the recovered color images without printing by scanning, or any other noise. In essence the bottom images are recovered directly from those in the middle row.

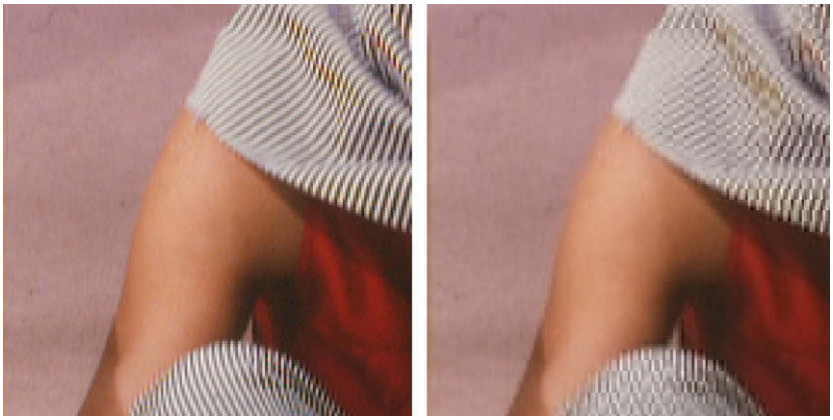

Fig. 9. Enlarged portions of original (left) image "barbara" and its recovered version (right) without printing, scanning or other sources of noise.

algorithms and on the distortions they aim to correct. Furthermore, a study on modelling the print-scan noise will be presented elsewhere.

Here, the subsampled chrominance planes are alternately embedded into the last N (out of M) high frequency subbands according to the zigzag path as shown in Fig. 6. Both real and simulated print-scan paths are used. In the simulations, we scale (interpolate) the gray image up by a factor of K in each direction, halftone the resulting image using error diffusion (Roetling and Loce, 1994), filter the bi-level image using a short Gaussian filter, add zero-mean white noise with variance σ_p^2, and downsample the noisy, filtered, halftoned image by averaging $K \times K$-pixel blocks.

As a first test we used the 4×4 DCT and varied N in the simulated print-scan path. Mean squared error (MSE) plots are shown in Fig. 7 for an average of few images. Results indicate that $N = 10$ (out of 16) seems to be an efficient compromise. Results for simulated print-scan and $M = 16, N = 10$ for many transforms are shown in Table 1, from which one can see that the DCT, despite its simplicity, outperforms other subband transforms.

With these results in mind we then tested a few images using 4×4 DCT and $N = 10$. Grey Images were printed at an HP LaserJet 4250 DTN printer and scanned using an HP Scanjet 3570c scanner. Two example images are shown in Fig. 8, along with their respective texturized versions. It also shows the recovered color images without any printing, scanning, or halftoning. Even though there is no noise source, it is quite interesting to see that the images at the bottom would be produced exclusively by directly processing the images in the middle row. The recovered images look very good. Nevertheless, a closer look at the details, as shown in Fig. 9, can easily reveal the artifacts.

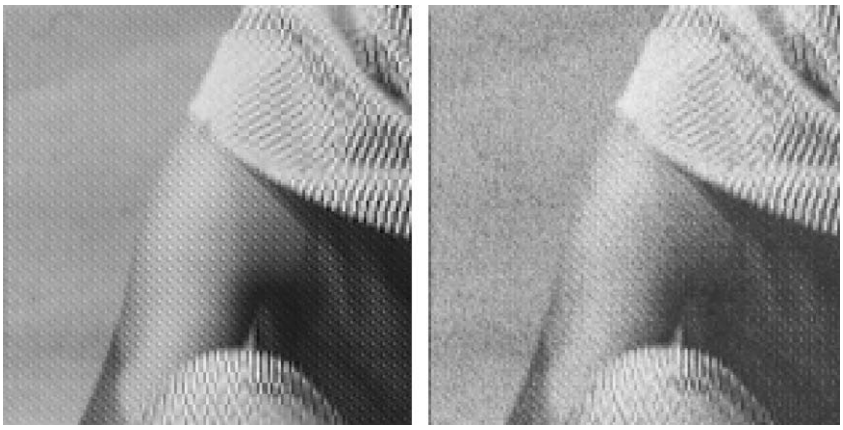

Fig. 10. Enlarged portion of texturized image, shown along with the result after printing, scanning and applying geometric correction. Note the large distortion that is imposed to the texture.

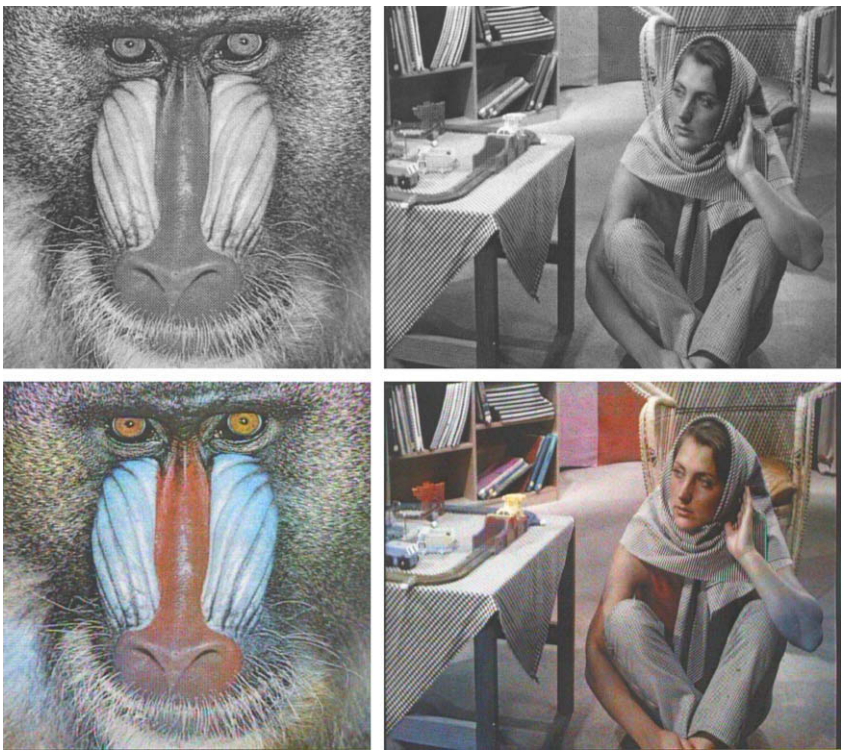

Fig. 11. Scanned and corrected images (top) and the corresponding recovered color images (bottom).

Once we print and scan the gray textured images, much of the texture disappears, contrast is reduced and a very large amount of noise is added. It can be seen in Fig. 10 the huge amount of degradation. Despite the fact that we process and correct the scanned image using an affine transform, there are still many non-linear effects that are not countered. In Fig. 11 we show the scanned images corresponding to those in Fig. 8, along with the respective recovered color images. The colors in Fig. 11 are not as saturated as those in Fig. 8. Furthermore, there are image regions where the colors were not recovered. These effects are caused by the very intense distortion produced by the print-scan path.

6. Conclusions

In this paper, we propose a method for reversible conversion from color images to gray ones based on redundant color information embedded into subbands of an *M*-band subband transform. Among the contributions, we propose to use a general subband transform rather than wavelets, since it generates more subbands, and the insertion of color information with strong linear redundancy. We have shown that the ideal distribution of the chrominance energy among many subbands is to replicate the chrominance plane for each subband. We also have calculated the SNR performance of the method and the bounds that limit the effectiveness of chrominance embedding, as well as the amount of redundancy. The bounds tell us that if the noise level is high enough one should spare more luminance subbands to protect the chrominance channel. Tests have shown the effectiveness of the method using both real and simulated print-scan paths.

Real print-scan paths have shown to cause severe degradation, reducing color saturation. Future work is planned to carry further non-linear processing on the recovered color image in order to remove luminance artifacts, saturate colors, etc. We also plan to improve the scan correction process to see if we improve the quality of the scanned image and reduce the noise of the print-scan path.

Acknowledgement

The author would like to thank Anderson A. Neves and Carlos V. F. Silva for their invaluable help in debugging the scan correction program and carrying printing and scanning tests.

References

Baharav, Z., Shaked, D., 1999. Watermarking of digital halftones. J. Electron. Imaging 3657, 307–313.

Chiang, P.J., Mikkilineni, A., Delp, E.J., Allebach, J.P., Chiu, G.T.-C., 2006. Extrinsic signatures embedding and detection in electrophotographic halftone images through laser intensity modulation. In: Proceedings of 22nd International Conference on Digital Printing Technologies, NIP 2006, pp. 432–435.

Coope, I.D., Renaud, P.F., 2000. Trace inequalities with applications to orthogonal regression and matrix nearness problems. Research Report No. UCDMS2000/17, Canterbury University.

Cox, I.J., Miller, M.L., Bloom, J.A., 2002. Digital Watermarking. Academic Press.

de Queiroz, R.L., Tran, T.D., 2001. Lapped Transforms for Image Compression (Chapter 5). In: Rao, K.R., Yip, P. (Eds.), The Handbook on Transforms and Data Compression. CRC Press.

de Queiroz, R.L., Braun, K., Loce, R., 2005. Detecting spatially varying gray component replacement with application in watermarking printed images. J. Electron. Imaging 14 (3), 033016.1–033016.9.

de Queiroz, R.L., Braun, K., 2006. Color to gray and back: Color embedding into textured gray images. IEEE Trans. Image Process. 15 (6), 1464–1470.

de Queiroz, R.L., 2007. Improved reversible mapping from color to gray. In: Proceedings of Brazilian Symposium on Computer Graphics and Image Processing, SIBGRAPI, Belo Horizonte, MG, Brazil.

Fu, M.S., Au, O.C., 2000. Data hiding in halftone image by pixel toggling. In: Proceedings of SPIE Security and Watermarking of Multimedia Contents II, Vol. 3971, San Jose, CA, US.

Golub, G.H., Van Loan, C.F., 1989. Matrix Computations, second ed. Johns Hopkins University Press, Baltimore, MD.

Horn, R.A., Johnson, C.R., 1985. Matrix Analysis. Cambridge University Press, Cambridge, UK.

Hunt, R.W.G., 2000. The Reproduction of Color. Fountain Press, Toolworth, England.

Knox, K.T., Wang, S., 1997. Digital watermarks using stochastic screens – a halftoning watermark. In: Proceedings of the SPIE International Conference on Storage and Retrieval for Image and Video Databases V, Feb. 1997, vol. 3022. San Jose, CA, US. pp. 310–316

Ko, K.-W., et al., 2007. Color recovery from gray image based on the analysis of wavelet packet subbands. In: Proceedings of SPIE-IS&T Electronic Imaging 2007, Color Imaging, San Jose, CA, USA.

Liu, C., Wang, S., Xu, B., 2004. Authenticate your digital prints with Glossmark images. In: Proceedings of NIP20: The 20th International Congress on Digital Printing Technologies, Salt Lake City, UT, US.

Malvar, H.S., 1992. Signal Processing with Lapped Transforms. Artech House, Norwood, MA.

Rao, K.R., Yip, P., 1990. Discrete Cosine Transform. Algorithms. Adv. Appl.. Academic Press, San Diego, CA, US.

Roetling, P.G., Loce, R.P., 1994. Marcel Dekker, New York.

Sharma, G., Loce, R.P., Harrington, S.J., Zhang, Y., 2003. Illuminant multiplexed imaging: GCR and special effects. In: Proceedings of IS&T/SID Eleventh Color Imaging Conference: Color Science, Systems and Applications, Scottsdale, AZ. pp. 266–271.

Strang, G., Nguyen, T., 1996. Wavelets and Filter Banks. Wellesley-Cambridge, Wellesley, MA.

Vetterli, M., Kovacevic, J., 1995. Wavelets and Subband Coding. Prentice-Hall, Englewood Cliffs, NJ.

Venkata, N.D., Yen, J., Monga, V., Evans, B.L., 2005. Hardcopy image barcodes via block-error diffusion. IEEE Trans. Image Process. 14 (12), 1977–1989.

Villan, R., et al., 2007. Tamper-proofing of electronic and printed text documents via robust hashing and data-hiding. In: Proceedings of SPIE-IS&T Electronic Imaging 2007, Security, Steganography, and Watermarking of Multimedia Contents IX, San Jose, CA, USA.

Voloshynovskiy, S., et al., 2006. Information-theoretic analysis of electronic and printed document authentication. In: Proceedings of SPIE-IS&T Electronic Imaging 2006, Security, Steganography, and Watermarking of Multimedia Contents VIII, San Jose, CA, USA, January 15–19.

Wang, S., Knox, K.T., 2000. Embedding digital watermarks in halftone screens. In: Proceedings of SPIE Security and Watermarking of Multimedia Contents II, vol. 3971, San Jose, CA, US.