COLOR IMAGE SCALABLE CODING WITH MATCHING PURSUIT

Rosa M. Figueras i Ventura, Pierre Vandergheynst, Pascal Frossard

Swiss Federal Institute of Technology (EPFL)
Signal Processing Institute
{rosa.figueras, pierre.vandergheynst, pascal.frossard}@epfl.ch

Andrea Cavallaro

Queen Mary, University of London
Department of Electronic Engineering
andrea.cavallaro@elec.qmul.ac.uk

ABSTRACT

This paper presents a new scalable and highly flexible color image coder based on a Matching Pursuit expansion. The Matching Pursuit algorithm provides an intrinsically progressive stream and the proposed coder allows us to reconstruct color information from the first bit received. In order to efficiently capture edges in natural images, the dictionary of atoms is built by translation, rotation and anisotropic refinement of a wavelet-like mother function. This dictionary is moreover invariant under shifts and isotropic scaling, thus leading to very simple spatial resizing operations. This flexibility and adaptivity of the MP coder makes it appropriate for asymmetric applications with heterogeneous end user terminals.

1. INTRODUCTION

Most visual coding schemes generally first compress luminance image components, and then extend coding to color components. They use color spaces with luminance and chrominance channels, where the latter may easily be downsampled, due to the fact that they carry much less information than the luminance component. Moreover, in the classical coding paradigm, decorrelation between channels is generally seen as an advantage. However, when a similar strategy is applied to scalable color image coding, it often causes colors to only appear after a certain time, or with severe distortion.

In this paper, we present a scalable color image coder, based on Matching Pursuit, that codes color information altogether, from the first information bit. The proposed image representation algorithm takes advantage of the correlation between the color channels, instead of trying to reduce it, as opposed to common coding schemes. The Matching Pursuit representation uses anisotropic refinement of wavelet-like mother functions, since such a dictionary has been shown to perform very well at low bit rates for gray scale images and is moreover intrinsically scalable [1]. The coder then chooses the same basis function for the three channels, and just recomputes the projection coefficient for each channel. Such an algorithm can be thought of as a vector greedy algorithm [2] trying to simultaneously approximate all three channels. Using a common waveform dramatically reduces the coding rate, since in Matching Pursuit (MP) representation over a redundant dictionary of atoms, indexes (i.e. position, scale and rotation, see [3]) generally carry most of the information, and thus represent the major part of the compressed stream. Such a strategy, in addition to providing excellent compression performance, allows for truly multiresolution representation of color images.

This work has been partly supported by the Swiss National Science Foundation.

This paper is organized as follows: Section 2 describes the Matching Pursuit image coder, and discusses the choice of the color space. Section 3 gives the details of the quantization and entropy coding processes, adapted to the color image representation generated by the MP algorithm. Experimental results are given in Section 4. Finally Section 5 concludes the paper.

2. MP COLOR IMAGE CODING

2.1. Matching Pursuit for Color images

Matching Pursuit (MP) image representation with a dictionary based on anisotropic refinement atoms and Gaussians has proved to give good compression results [3], in addition to intrinsic spatial and rate scalability properties [1]. MP is a greedy algorithm that iteratively chooses the atom of a redundant dictionary that provides the best correlation with the input signal (see [4, 5] for details on the algorithm). Let $\mathcal{D} = \{g_\gamma\}_{\gamma\in\Gamma}$ be a dictionary of $P > M_1 \times M_2$ unit norm vectors. This dictionary includes $M_1 \times M_2$ linearly independent vectors that define a basis of the space $\mathbb{C}^{M_1 \times M_2}$ of signals of size $M_1 \times M_2$. Also, let $R^n f$ be the residual of an n term representation of signal f, with $R^0 f = f$. The signal decomposition with MP then can be written as follows:

$$f = \sum_{n=0}^{N-1} \langle R^n f, g_{\gamma_n}\rangle g_{\gamma_n} + R^N f, \tag{1}$$

where g_{γ_n} is the dictionary vector that maximizes the energy taken from $R^n f$ at every iteration:

$$|\langle R^n f, g_{\gamma_n}\rangle| = \arg\max_{\mathcal{D}} |\langle R^n f, g_\gamma\rangle| . \tag{2}$$

Instead of performing independent iterations in each color channel, a vector search algorithm is used in the proposed color image encoder. This is equivalent to using a dictionary of P vector atoms of the form $\{\vec{g_\gamma} = [g_\gamma, g_\gamma, g_\gamma]\}_{\gamma\in\Gamma}$. In practice though, each channel is evaluated with one single component of the vector atom, whose global energy is given by adding together its respective contribution in each channel. MP then naturally chooses the vector atom, or equivalently the vector component g_γ, with the highest energy. Hence the component of the dictionary chosen at each

Matching Pursuit iteration satisfies:

$$\gamma_n = \arg\max_{\gamma_n} \sqrt{\langle R^n f^1, g_{\gamma_n}\rangle^2 + \langle R^n f^2, g_{\gamma_n}\rangle^2 + \langle R^n f^3, g_{\gamma_n}\rangle^2},$$

where $R^n f^i$, $i = 1, 2, 3$ represents the signal residual in each of the color channels. Note that this is slightly different than the algorithm introduced in [2], where the sup norm of all projections is maximized :

$$\max_i \sup_{\mathcal{D}} |\langle R^n f^i, g_{\gamma_n}\rangle| .$$

All signal components are then jointly approximated through an expansion of the from :

$$f^i = \sum_{n=0}^{+\infty} \langle R^n f^i, g_{\gamma_n}\rangle g_{\gamma_n}, \ \forall i = 1, 2, 3.$$

Note that channel energy is conserved, and that the following Parseval-like equality is verified :

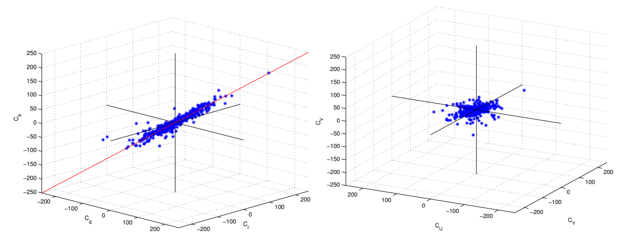
$$\|f^i\|^2 = \sum_{n=0}^{+\infty} |\langle R^n f^i, g_{\gamma_n}\rangle|^2, \ \forall i = 1, 2, 3.$$

The search for the atom with the highest global energy necessitates the computation of the three scalar products $c_n^i = \langle R^n f^i, g_{\gamma_n}\rangle$, $i = 1, 2, 3$, for each atom g_{γ_n}, and for each iteration of the Matching Pursuit expansion. The number of scalar products can be reduced by first identifying the color channel with the highest residual energy, and then performing the atom search in this channel only. Once the best atom has been identified, its contribution in the other two channels is also computed and encoded. The reduced complexity algorithm obviously performs in a suboptimal way compared to the maximization of the global energy, but in most of the cases the quality of the approximation does only suffer a minimal penalty (Fig. 2(b) is an example of MP performed in the most energetic channel).

2.2. Color space choice

The choice of the color space has to be adapted to the proposed coding strategy, and minimize the distortion introduced to the image due to the coding scheme itself and the quantization. Simultaneously, the color space has also an impact on the compression performance. Interestingly, the MP encoder seems to favor correlated color spaces, since atom indexes induce higher coding costs than color coefficients.

In order to reach a maximal correlation between the color channels, and thus the possibility to reach higher compression ratios, the RGB color space seems a natural choice. This choice can be justified by the following simple experiment. The coefficients $[c_n^1, c_n^2, c_n^3]$ of the MP decomposition can be represented in a cube, where the three axes respectively corresponds to the red, green and blue components (see Fig. 1(a)). It can be seen that the MP coefficients are interestingly distributed along the diagonal of the color cube, or equivalently that the contribution of MP atoms is very similar in the three color channels. This very nice property is a real advantage in overcomplete expansions where the coding cost is mainly due to the atom indexes. In the contrary, the distribution of MP coefficients resulting from the image decomposition in the YUV color space, does not seem to present an obvious structure (see Fig. 1(b)). In the coding algorithm proposed in this paper, these coefficients become much more difficult to code efficiently. In addition, the YUV color space has been shown to give quite annoying color distortion for some particular images (see Fig. 2(a)).



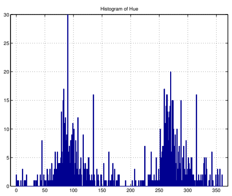
(a) RGB (b) YUV

Fig. 1. Distribution of the MP coefficients when MP is performed in the RGB color space (left), with the diagonal of the cube shown in red, and in the YUV color space (right).

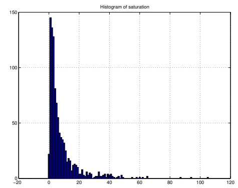
(a) YUV most energetic (b) RGB most energetic

Fig. 2. Japanese woman coded with MP using the most energetic channel search strategy in YUV color space 2(a) and in RGB color space 2(b).

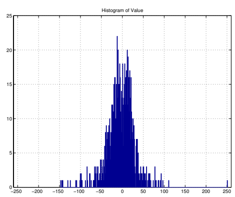
Most of the image coding techniques use color spaces such as YUV, LAB of CrCbCg. These color spaces have less redundancy among channels than RGB. For example, YUV has all the luminance information in the Y channel, and the U and V channels have less information. LAB and CrCbCg color spaces have the drawback that they have some user defined parameters, not standardized for all the displays. They do not present the same amount of redundancy among channels as RGB does. The techniques which use the YUV color space generally take profit from the fact that the human eye is less sensitive to color than luminance localization in images, and hence downsample the U and V channels in order to reduce the amount of data. Such a downsampling is not helpful anymore in the context of the proposed MP coder, since the same function is used for the three color channels, in order to limit the coding cost of atom indexes. Since, in addition, the use of the same function in Y, U and V channels may induce color distortion, RGB becomes clearly the preferred color space for the MP coder.



(a) Hue



(b) Saturation



(c) Value

Fig. 3. Histograms of the MP coefficients when represented in HSV coordinates.

3. PARAMETER CODING

3.1. Coefficient quantization

A fundamental goal of data compression is to obtain the best possible fidelity for a given data rate or, equivalently, to minimize the rate required for a given fidelity. Due to the structure of the coefficient distribution, centered around the diagonal of the RGB cube, an efficient color quantization strategy is not anymore to code the raw value of the R, G and B components, but instead to code the following parameters: the projection of the coefficients on the diagonal, the distance of the coefficient from the diagonal and the direction where it is located. This is equivalent to coding the MP coefficients in an HSV color space, where V (Value) becomes the projection of RGB coefficients on the diagonal of the cube, S (Saturation) is the distance of the coefficient to the diagonal and H (Hue) is the direction perpendicular to the diagonal where the RGB coefficient is placed. The HSV values of the MP coefficients present the following distribution: the Value distribution is Laplacian centered in zero (see Fig. 3(c)), Saturation presents an exponential distribution (see Fig. 3(b)), and a Laplacian-like distribution with two peaks can be observed for Hue values 3(a). Once the HSV coefficients have been calculated from the available RGB coefficients, the quantization of the parameters is performed as follows:

- V is exponentially quantized with the quantizer explained in [6]. The choice of exponential quantization is driven by the exponential decay of V coefficients as the iteration number increase. The value that will be given as input to the arithmetic coder will be $N_j(l) - Quant(V)$, where $N_j(l)$ is the number of quantization levels that are used for coefficient l.
- H and S are uniformly quantized, since the iteration number does not seem to have any particular influence on their magnitude.

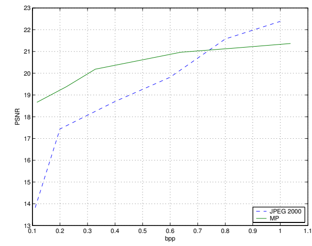


Fig. 4. PSNR comparison between JPEG 2000 and MP. The PSNR has been computed in the CIELAB color space.

3.2. Entropy coding of the parameters

The entropy coding of the quantized parameters is performed through and adaptive arithmetic coder based on [7], with the probability update method used in [8]. The detailed implementation of this arithmetic coder is given in [3], where it was used for a MP scalable image coder for gray scale images. The only difference is the number of fields to be entropy coded (color images need two extra fields to represent color information). Note that this coder is still under study, and currently does not yet behave optimally. A more elaborated study of the entropy of the parameters, and improved histogram initialization and update methods are necessary to obtain better compression performance. However, the bit-stream obtained is still representative enough of the behavior of the MP coder.

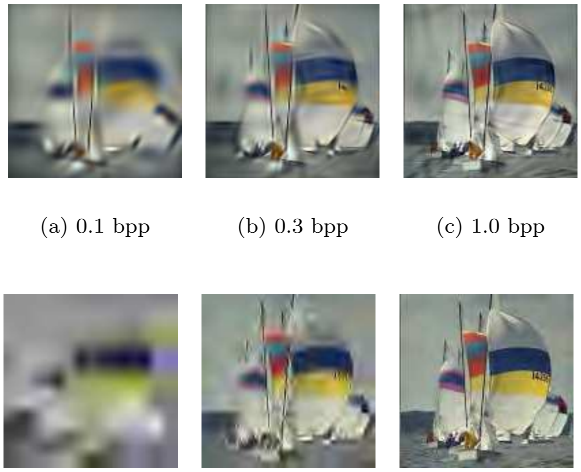
4. RESULTS

In the previous section it has been shown that the choice of the RGB color space is the most adapted to the MP algorithm which uses the same mother function for the three color channels. All results presented in this section thus use RGB color space for MP expansion.

Fig. 4 first compares the compression performance of the proposed MP encoder with state-of-the-art JPEG-2000. It can be seen that MP advantageously compares to JPEG-2000, and even performs better at low bit rates. This can be explained by the property of MP to immediately capture most of the signal features in a very few iterations. At higher rates, the advantage of redundant expansion obviously decreases in comparison to orthogonal signal decomposition as performed by JPEG-2000. Note that the PSNR values have been computed in the Lab color space [9], in order to closely evaluate the Human Visual System perception. Fig. 5 proposes visual comparisons between MP and JPEG-2000, and it can be clearly seen that MP indeed offers better visual quality at low bit rate.

In addition to interesting compression properties, the MP bitstream offers highly flexible adaptivity in terms of rate and spatial resolution. Fig. 6 shows the effects of truncating the MP expansion at different number of coefficients. It can be observed that the MP algorithm will first describe the main objects in a sketchy way (keeping the colors) and then it will refine the details. The stream generated by MP thus truly offers an intrinsic rate scalability in image representation.

Finally, MP decompositions with anisotropic refinement atoms are covariant under isotropic dilations [3, 1]. This



(a) 0.1 bpp (b) 0.3 bpp (c) 1.0 bpp

(d) 0.1 bpp (e) 0.3 bpp (f) 1.0 bpp

Fig. 5. Top row, MP of sail with coefficients quantized in HSV color space. Bottom row, JPEG2000 for the same bit-rate.

(a) 50 coeffs (b) 150 coeffs (c) 500 coeffs

Fig. 6. MP of sail for 50, 150 and 500 coefficients.

means that the compressed image can be resized with any ratio α, including irrational factors (see Fig. 7). These spatial scalability properties, added to very low complexity decoder implementation, make MP especially useful for asymmetric applications with heterogeneous end-user terminals.

5. CONCLUSIONS

In this paper we have extended to color images a scalable coding scheme initially developed for gray scale images. The MP color image coding takes advantage of the redundancy of the color channels to reduce the coding rate. The compression performance efficiently compares to state of the art techniques like JPEG 2000 at very low bit-rates, even though it can still be improved by taking advantage of the coefficient distribution. The MP color image coder presented in the scope of this paper additionally allows for rate and spatial scalability, with the particularity that colors already appear from the first decoded bits. In addition, the distortion introduced by MP for very low bit-rate images is a sort of "sketching" of the image, instead of ringing artifacts or blocks. This sketchy image is quickly understood and accepted by the human observer, in contrary to highly blocky images.

(a) $\alpha = .5$ (b) Original size (c) $\alpha = \sqrt{2}$

Fig. 7. Example of spatial scalability with MP streams.

6. ACKNOWLEDGMENTS

Thanks to Elena Salvador for fruitful discussions.

7. REFERENCES

[1] P. Frossard, P. Vandergheynst, and R. M. Figueras i Ventura, "High flexibility scalable image coding," in *Proceedings of VCIP 2003*, July 2003.

[2] A. Lutoborski and V. N. Temlyakov, "Vector greedy algorithms," *Journal of Complexity*, vol. 19, no. 4, pp. 458–473, 2003.

[3] R. M. Figueras i Ventura, P. Vandergheynst, and P. Frossard, "Highly flexible image coding using nonlinear representations," Tech. Rep. TR-ITS-03.02, ITS, EPFL, June 2003.

[4] S. Mallat and Z. Zhang, "Adaptive time-frequency decomposition with matching pursuits," in *Time-Frequency and Time-Scale Analysis, 1992, Proceedings of the IEEE-SP International Symposium*, October 1992, pp. 7–10.

[5] S.G. Mallat and Zhifeng Zhang, "Matching pursuits with time-frequency dictionaries," in *IEEE Trans. on Signal Processing*, December 1993, vol. 41, Issue 12, pp. 3397–3415.

[6] P. Frossard, P. Vandergheynst, R. M. Figueras i Ventura, and Murat Kunt, "A posteriori quantization of progressive matching pursuit streams," *To appear in IEEE Trans. on Signal Processing*, 2003.

[7] I. H. Witten, R. M. Neal, and J. G. Cleary, "Arithmetic Coding for Data Compression," *Communications of the ACM*, vol. 30, no. 6, pp. 520–540, June 1987.

[8] D. L. Duttweiler and C. Chamzas, "Probability estimation in arithmetic and adaptive-huffman entropy coders," *IEEE Trans. on Image Processing*, vol. 4 Issue: 3, pp. 237–246, Mar 1995.

[9] M .D. Fairchild, *Color Appearance Models*, Addison–Wesley Publishing Co., 1998.