Color Image Compression: Early Vision and the Multiresolution Representations

Ahmed Nabil Belbachir, Peter Michael Goebel

Vienna University of Technology, Institute of Computer Aided Automation, Pattern Recognition and Image Processing Group,
Favoritenstr. 9/183-2, 1040 Vienna, Austria.
Tel: 00431-58801 18351, Fax: 00431-58801 18392, Email:
<nabil,goe>@prip.tuwien.ac.at

Key words: Image and Video Compression, Colour Vision, Image Analysis

1. **What is the original contribution of this work?** This work investigates the emphasis of anisotropic features in the chrominance channels of color images, represented in YUV space, for compression purpose. In [3], this property is stated to be important for gray-level images by means of the contourlet transform and its pyramidal directional filter bank. This paper extends that work to color images.
2. **Why should this contribution be considered important?** In this work, it is shown that directionality is not important for chrominance channels. In [3], contourlets were successful for high-resolution gray-level images compression with preservation of details. However, this investigation shows that this method has no advantage for chrominance.
3. **What is the most closely related work by others and how does this work differ?** The most closely work is the one of reference [11], while the proposed paper additionally considers anisotropy features in color images.
4. **How can other researchers make use of the results of this work?** Other researchers may exploit our investigation for not considering anisotropy in chrominance. However, as a perspective, it is recommended to determine different contrast sensitivity functions [13] for the chrominance channels as well as for the luminance. To tune the quantization process, one can estimate the receptor noise [20] as threshold to detect the just noticeable difference based on the Human Visual System sensitivity.
5. **Has this work been presented/submitted elsewhere?** No.
6. **Which form of presentation is preferred: Oral or Poster?** Oral.

Color Image Compression: Early Vision and the Multiresolution Representations

Ahmed Nabil Belbachir, Peter Michael Goebel

Vienna University of Technology,
Institute of Computer Aided Automation,
Pattern Recognition and Image Processing Group,
Favoritenstr. 9/183-2, A-1040, Vienna
<nabil,goe>@prip.tuwien.ac.at

Abstract. The efficiency of an image compression technique relies on the capability of finding sparse M-terms for best approximation with reduced visually significant quality loss. By "visually significant" it is meant the information to which human observer can perceive. The **H**uman **V**isual **S**ystem (HVS) is generally sensitive to the contrast, color, spatial frequency...etc. This paper is concerned with the compression of color images where the psycho-visual representation is an important strategy to define the best M-term approximation technique. Digital color images are usually stored using the RGB space, television broadcast uses YUV (YIQ) space while the psycho-visual representation relies on 3 components: one for the luminance and two for the chrominance. In this paper, an analysis of the wavelet and contourlet representation of the color image both in RGB and YUV spaces is performed. A approximation technique is performed in order to investigate the performance of image compression technique using one of those transforms.

1 Introduction

The question of why some abstract shapes are more attractive to human observers than others, may be answered, in that, we probably find pleasing those forms, most closely tuned to the properties of our human visual system (HVS) [12]. An example of differential tuning is the oblique effect[1] in orientation perception. Thus aesthetic pleasure is linked in some general way to neural activity. In [7] a theory of spatial vision is given, by integrating psychophysics, neurophysiology and linear systems. They carried out many electro-physiological experiments, showing the cells of the primary visual cortex be tuned to bands in spatial frequency and orientation, which was confirmed by other physiological experiments [6]. However, the orientation properties of color perception is still an open issue. In applications where lossy compression is needed, one has to consider subjective assessments, rather than – or besides – strict PSNR values for an optimal quality of the reconstruction.

[1] horizontal and vertical lines have privileged access

As shown in a recent paper [3], considering anisotropy in luminance images in the compression stage, has led to better compression performance and better preservation of details, thus getting higher image quality. The question is: The question is: can the considerations in [3] be exploited for color images?

This paper analyzes a compression technique on color images. In the following subsections, the modeling of color in images as well as the performance of compression in color space are considered.

1.1 Psycho-visual Representation of Color Images

The quantization of color perception in the HVS is a challenging process. There exist different theoretical studies [21, 19] of the chromatic psycho-visual components that the HVS could perceive. The theories show that the HVS separates light information, received by the retina, into three distinct, non independent components. Thus, HVS color perception can be modeled in two steps [21]. In the first step, the received light is transformed into three bio-electrical signals (L,M,S) as received by the three types of cones on the retina [4]. In the second step, the resulting signals are combined into three psycho-visual components: one achromatic (A) and two color components (C1 and C2).

Using this model, the achromatic component A is calculated by $A = L + M$ [1]. The first color component $C1$ corresponds to the chromatic axis Red-Green $C1 = L - M$. The second color component C2 corresponds to the chromatic axis Blue-Yellow $C2 = S - 0.5(L + M)$. Note also that if $C1$ corresponds to the a axis from $L{*}ab$ space defined by the CIE-Lab [16], then $C2$ diverges from the b component of this space. An investigation has been made in [1] on the possible relationships between the three components A, C1 and C2. It was shown a strong interaction between A and C1, and C1 and C2, while C2 has a limited influence on the perception of A. Thus, efficient quantization process should take into account these relationships instead of handling the three components independently.

1.2 Color compression standards

Lossy compression mainly consists of de-correlation and quantization stages that reduce the image size by permanently eliminating certain information. The decorrelation stage of the image compression algorithm is usually done by a transformation from one space to another space to facilitate compaction of information. One approach is the use of the **D**iscrete **C**osine **T**ransform (DCT), which is used in the JPEG (baseline) industry standard [17].

The recent standard "JPEG 2000" [18] exploits the sparse approximation of the **D**iscrete **W**avelet **T**ransform (DWT) for image compression purpose, which is free from the blocking effect artifacts of DCT. DWT allows the decomposition of an image into a set of contributions at different frequency bands and resolution levels. The corresponding coefficients of the different decomposition levels are correlated and show characteristic trends for additional compression potential.

Many coding techniques of the DWT coefficients have been investigated, like the **E**mbedded **Z**ero-**T**ree **W**avelet (EZW), the **S**et **P**artitioning **I**n **H**ierarchical **T**rees (SPIHT) and the **E**mbedded **B**lock **C**oding with **O**ptimized **T**runcation (EBCOT) [17, 18].

1.3 Color Images modeled in RGB and YUV

There exist many possible color space definitions [15] and this paper addresses those commonly required by image compression techniques. For digital still images, the **R**ed-**G**reen-**B**lue color space, known as RGB, is commonly used. RGB is an additive color model [16], which appropriately fits the physics of usual capturing and display devices. A RGB representation of an image is generated by spectral primary filtering an arbitrary color scene. Filters generate three channels by the spectral subbands red, green and blue, which usually overlap. Therefore the RGB model representation is a very redundant one.

The combining of these three channels of light produces a wide range of visible colors. All color spaces are three-dimensional orthogonal coordinate systems, meaning that there are three axes (in this case the red, green, and blue color intensities) that are perpendicular one to another. The luminance is given by $Y = 0.299R + 0.587G + 0.114B$ and is represented by the area of the triangle, spanned from the three RGB vectors. All three components must be tuned proportionally for changing the luminance, therefore, RGB is not optimal for compression purpose. Thus, a changeover to a color space with uncorrelated components is more convenient and provides better compression performance.

Television broadcast makes use of color spaces based on luminance and chrominance, which correspond to brightness and color respectively. These color spaces are denoted as YUV and YIQ [15]. The YUV space is used for the PAL broadcast television system standard in Europe and the YIQ color space is used for the NTSC broadcast standard in North America.

Chrominance carries only the differences R-Y and B-Y, which is the principle advantage of using YUV or YIQ for broadcast. Thus the amount of information is significantly reduced to define such a color television image. From the past, the compatibility with monochrome receivers should be noticed.

Based on psycho-visual properties of the human eye, luminance is more important for subjective good image quality than chrominance. Therefore chrominance components can be downsampled to achieve better compression performance. So the formats YUV:4:2:2 and YUV:4:1:1 were generated, especially for video compression applications.

1.4 Multiresolution Representations of Images

For compression purposes, sparse representation[2] of images is an important issue, where the compact approximation is mainly achieved by means of quantization,

[2] A general WVT series expansion by Φ_n for a given image I, such that $I = \sum_0^\infty C_n \Phi_n$, where C_n are the transform coefficients. The best M-term approximation, using this expansion, is defined as $I_M = \sum_{n \in G_M} C_n \Phi_n$, where G_M is the set of indexes of the

which also yields to a degradation of the images. Separable WVT [9] offers an acceptable tradeoff between visual quality and information sparsity. Thus, WVT has been adopted for gray-level and color image compression standards (i.e. JPEG2000) [13, 17, 18].

The quality of the approximation images strongly depends on the quantization noise. For gray-level images, this noise is uniform as one quantization procedure is used. However, three quantization procedures are used for the three image components that may yield a quantization noise localized on certain structures. The minimization of such noise depends on the capability of exploiting the interaction between the three components.

For the assessment of the reconstruction quality another important aspect is the relationship between the contrast sensitivity and the spatial frequency. This is described by the **C**ontrast **S**ensivity **F**unction (CSF) [13]. For compression purposes, CSF is exploited to regulate the quantization step-size to minimize the visibility of artifacts. This approach can be used for luminance as well as for color images. For color perception three different CSFs describe the sensitivity for the respective color bands. Implemented in JPEG2000 compression standard on the WVT coefficients, CSF schemes provide visually more efficient image approximation than conventional hard thresholding.

2 The Contourlet Transform

Efficient image representations require that coefficients of functions, which represent the regions of interest, are sparse. Wavelets can pick up discontinuities of one dimensional piecewise smooth functions very efficiently and represent them as point discontinuities. 2D WVT obtained by a tensor product of one-dimensional wavelets are good to isolate discontinuities at edge points, but cannot recognize smoothness along contours. Numerous methods were developed to overcome this limitation by adaptive [14], Radon-based [5], or filter bank-based techniques [8].

2.1 Contourlets for Luminance Images

Inspired by Candes theory [5], Do and Vetterli [8] proposed the **P**yramidal **D**irectional **F**ilter-**B**ank (PDFB), which overcomes the curvelet in sparsity using a directional filter bank, applied on the whole scale, also known as CTT. Besides their parsimony, CTT offers the directionality and anisotropy to image representation that are not provided by separable WVT.

In [3], the potential of the CTT for image compression has been demonstrated on gray-level images. The advantage of CTT over WVT, is the sparse approximation of images with smooth contours. We did experiments with over 100 high resolution images, and it was proved, that the smoothness of the contours within an image is coupled with the spatial resolution of a desired scale.

M-largest C_n. The quality of the approximated function I_M relates to how sparse the WVT expansion by Φ_n is, or how well the expansion compacts the energy of I into few coefficients.

It is found, that below a spatial resolution around 2^8 pixels, the application of the CTT carries no advantage compared with WVT in terms of compaction of energy. By combining CTT and WVT as unique image transform, an interesting gain in image quality has been demonstrated. These results motivate the interest for extending this approach to color images.

2.2 Contourlets for Color Images

The proposed approach for color images approximation is very close to the one used in [3] for gray-level images. Consequently, the main idea is to combine CTT-WVT for the approximation of the three image components Y, and also U and V. The adopted procedure for image compression is as follows:

1. The used digital images are stored using RGB coordinates. In the first step, the RGB component of the color image are transformed into YUV space as defined in [16].
2. In case of image size of $\leq$ 512 x 512 pixels, four decomposition levels are used, where the two fine scales are CTT, (with l=8 directions) and the remaining two are obtained using WVT. However, in case of images with $\leq$ 2048 x 2048 pixels, five decomposition steps are using with three CTT levels (l=16 directions).
3. This decomposition is used to each component independently Y, U and V.
4. Simple thresholding of the coefficients from Y space is performed as in [3].
5. The same thresholding strategy, as in 4., is performed for the resulting coefficients from U image decomposition.
6. In the V decomposition, the coefficients located in the same position as the truncated Y coefficients as excluded from the thresholding. Other V coefficients follow the same thresholding strategy.

3 Experimental Results and Analysis

This analysis has been performed on more than 50 digital color images from a multimedia database. The WVT and combined CTT-WVT approach have been analyzed for compression purpose. A set of images have been selected for this paper that are depicted in Figure 1.

Most of those images has been used in [11]. The image A and C are the Lena and F16 color images respectively, with 512x512 pixels, which are usually used in standard image compression literate. In B, there is Zaza image with 512x512 pixels that has the property of containing smooth contours without a complex structure. The image in D is the Art image with 2048x2048 pixels that contains a diversity of colors.

Compression results using WVT and CTT-WVT on the four selected images, out of 50, are reported in this Section. The evaluation between these approaches is performed using the PSNR criterion. The images have been reconstructed from the remaining significant coefficients and the reconstruction error (PSNR)

Fig. 1. Example of color images used for the analysis: A. Lena(512x512), B. Zaza(512x512), C. F16(512x512) and D. Art image(2Mx2M).

has been derived. The normalized log energy entropy measures indicate the potential of energy compaction for the desired transformation. The factor between the original image size and the non-zeros coefficients size is used to derive the compression ratio. Note that state-of-the-art coders (e.g. EBCOT contextual coding [18]) also exploit high order statistics. Thus, the used measures (entropy) are an approximation that can only give a hint on the relative performance. The entropy measure is given over the whole YUV components as well as over the luminance component Y.

It can be noticed that the combination CTT-WVT provides higher reduction than that of WVT. However, the entropy measure of the resulting YUV space coefficients seems to be advantageous for WVT such that the remaining coefficients can be efficiently coded. On the other hand, the entropy results of the Y space coefficients show that CTT-WVT provides a sparser representation for luminance than that with WVT while chromatic components seems to be compactly represented by WVT. Therefore, one possible altenative to investigate would be the combination of CTT and WVT for luminance and chrominance

Images	PSNR	Entropy YUV	Entropy Y	Compression Ratio
Lena (WVT)	30.1	0.054	0.054	21.8
Lena (CTT-WVT)	29.9	0.065	0.053	20.7
F16 (WVT)	26.8	0.035	0.033	31.4
F16 (CTT-WVT)	26.8	0.043	0.032	32
Zaza (WVT)	29.7	0.031	0.029	38.1
Zaza (CTT-WVT)	29.6	0.034	0.028	39.3
Art (WVT)	28.1	0.057	0.048	34.8
Art (CTT-WVT)	28.2	0.061	0.045	36.7

Table 1. Compression results.

components respectively. In such a scheme, CTT can be used to code the Y channel while WVT can be applied for U and V components

4 Conclusions

This paper investigates the potential of the **C**on**T**ourlet **T**ransform (CTT) for compression of color images. In [3], it was demonstrated the potential of CTT, in providing a more compact representation of the energy, compared to that of the **W**a**V**elet **T**ransform (WVT), used in the compression standard JPEG2000, for gray-level images. CTT shows less information loss and artifacts in the reconstructed images after simple thresholding of the transformed coefficients. Indeed, WVT exhibits a large number of coefficients for representing smooth contours, which are usually present in high resolution images.

For this reason, this paper investigates the extension of the CTT-WVT approach to color images. Several color spaces has been discussed including the space with the psycho-visual perception of colors. The CTT potential has been investigated in YUV space, the one used in JPEG and JPEG2000 standards. After the quantization step, CTT provides a reduced size of coefficients, compared to WVT. Unfortunately, the log energy entropy measure indicates that the resulting WVT coefficients may be compacter coded. This is apart to the Y channel, which can be coded better by processing by CTT. However, as a perspective, it is recommended to determine different contrast sensitivity functions [13] for the chrominance channels as well as for the luminance. To tune the quantization process, one can estimate the receptor noise [20], as threshold, to detect a just noticeable difference (Weber-Fechner law), based on the Human Visual System sensitivity.

References

1. D. Barba. Codage avec Compression d'Images Couleur. Ecole du Printemps, Pau, France, 2001.

2. L. Bedat. Aspects Psychovisuels de la Perception des Couleurs. Application au Codage d'Images Couleur Fixes avec Compression de l'Information. PhD Thesis at University of Nantes, France, Oct. 1998.
3. A.N. Belbachir and P.M. Goebel. A Sparse Image Representation Using Contourlets. 10th Computer Vision Winter Workshop, Zell an der Pram, Austria, February 2005.
4. P. Le Callet and D. Barba. A Robust Quality Metric for Color Image Quality Assessment. Proceeding of the International Conference on Image Processing, Vol 1, pp. 437-440, Sep. 2003.
5. E.J. Candes and D.L. Donoho. Curvelets - A Surprisingliy Effective Non-Adaptive Representation for Objects with Edges. in Curve and Surface Fitting, A. Cohen, C. Rabut, and L.L. Schumaker, Eds. Saint Malo: Vanderbilt University, 1999.
6. J. G. Daugman. Complete discrete 2D Gabor transforms by neural networks for image analysis and compression. IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. **36**, No.7, pp. 1169-1179, 1988.
7. K. and R. De Valois. Spatial Vision. Oxford Psychology Series No. **14**; Oxford University Press, 1990
8. M.N. Do and M. Vetterli. The Contourlet Transform: An Efficient Directional Multiresolution Image Representation. IEEE Transactions on Image Processing, Dec. 2004.
9. I. Daubechies. Ten Lectures on Wavelets. CBMS-NSF Lecture Notes Nr. 61, Seventh Printing SIAM, 2002.
10. D. Donoho, M. Vetterli, R.A De Vore and I. Daubechies. Data Compression and Harmonic Analysis. IEEE Transactions on Information Theory, vol. **44**, no. 6, pp. 24352476, October 1998.
11. D.M. Gomes, W.T.A. Filho and A.D.D. Neto. A Method for Selective Color Images Compression. Proceedings of the International Joint Conference on Neural Networks, Volume **2**, 20-24 July 2003.
12. B. C. Hansen, J. K. DeFord, M. J. Sinai and E. A. Essock. Anisotropic processing of natural scenes depends on scene content. Journal of Vision, 2(7), 500a, http://journalofvision.org/2/7/500/, doi:10.1167/2.7.500, 2002.
13. M.J. Nadenau, J. Reichel and M. Kunt. Wavelet-Based Color Image Compression: Exploiting the Contrast Sensivity Function. IEEE Transactions on Image Processing Vol.**12**, NO.1, January 2003.
14. E.L. Pennec and S. Mallat. Image Compression with Geometric Wavelets. IEEE International Conference on Image Processing, 2000.
15. C.A. Poynton. A Technical Introduction to Digital Video. John Wiley, NY, 1996.
16. M. Richter. Einfuehrung in die Farbmetrik. W..D Gruyter, Berlin-NY, 1981.
17. K. Sayood. Introduction to Data Compression. *Second Edition, Morgan Kaufmann, 2000.*
18. D. Taubman. High Performance Scalable Image Compression with EBCOT. IEEE Trans. on Image Processing, VOL. **9**, NO. 7, July 2000.
19. K.K. De Valois, M.A. Webster and E. Switkes. Orientation and Spatial Frequency Discrimination for Luminance and Chromatic Gratings. JOSA, Vol. **7**, N6, pp. 1034-1049, 1990.
20. M. Vorobyev and D. Osorio. Receptor Noise as a Determinant of Colour Thresholds. Pro. Royal Society, London **265**, pp. 351-358, 1998.
21. D.R. Williams, J. Krauskopf and D.W. Heeley. Cardinal Direction of Color Space. Vision Research, Vol.**22**, pp. 1123-1131, 1982.