IMPROVED IMAGE PARTITIONING FOR COMPRESSION AND REPRESENTATION USING THE LAB COLOR SPACE IN THE LAR IMAGE CODEC

Clément Strauss, François Pasteau, Marie Babel, Olivier Déforges, Laurent Bédat

IETR/Image group Lab
CNRS UMR 6164/INSA Rennes
20, avenue des Buttes de Coësmes
35043 RENNES Cedex, France
cstrauss@insa-rennes.fr {fpasteau, mbabel, odeforge, lbedat}@insa-rennes.fr

ABSTRACT

The LAR codec is an advanced image compression method relying on a quadtree partitioning of the image. The partitioning strongly impacts the LAR codec efficiency and enables both compression and representation efficiency. In order to increase the perceptual representation abilities without penalizing the compression efficiency we introduce and evaluate two partitioning criteria working in the Lab color space. These criteria are confronted to the original criterion and their compression and robustness performances are analyzed.

1. INTRODUCTION

JPEG committee has recently sent out a call for proposal for a new image coding standard (JPEG-AIC), and the LAR method has been proposed as a potential candidate [4]. The design of new generation image coding addresses different issues. Compression efficiency is one of the main focus but it is no longer sufficient. Other features such as content based coding and psycho-visual considerations are also expected.

Psycho-visual aspects have been considered in the past for video coding [1] and image coding such as the JPEG coder [10] where vision models have been used as improvement. Works on wavelet quantization noise visibility [13] have recently led to perceptual distortion control in JPEG2k [7]. The main drawback of this last method is that perceptual considerations are externally used to set the coder to obtain a given visual quality while this coder does not fundamentally consider the human vision properties.

Our coder philosophy is not to outperform JPEG2k in compression. Our goal is to propose an open source, royalty free, alternative image coder with highly integrated services. While keeping the compressions performances in the same range as JPEG2k with a lower and adjustable complexity, our coder also provide evolved functionalities [4] such as scalable lossy to lossless compression, region level representation and coding but also multiple services like cryptography, data hiding and error resilience.

The aim of this work is to introduce perceptual considerations in our color image coder while designing a low complexity solution. LAR coding method relies on a particular quadtree partitioning of the image [5]. This partitioning describes the image content and strongly impacts the user's quality perception of the compressed images. Our focus here is to improve the existing image partitioning method by introducing a psycho-visual color space. In order to increase the perceptual quality the partitioning process has to give a more accurate internal image representation while keeping compression performances in an acceptable range.

This paper is organized as follows. Section 2 introduces the LAR coding method. Section 3 describes the partitioning process and its influence onto our coder. Section 4 describes the partitioning process with the Lab color space. Section 5 provides comparative results between the use of the Lab color space instead of the original one.

2. LAR CODEC: PRINCIPLES

2.1 LAR overview

The LAR (Locally Adaptive Resolution) codec relies on a two layer coder system as illustrated in figure 1. The first layer, called FLAT coder, encodes the global image information leading to a low bit-rate version of the image. The second layer, called spectral layer, deals with texture, aiming at visual quality enhancement at medium/high bit-rates. Therefore, the method provides natural SNR scalability.

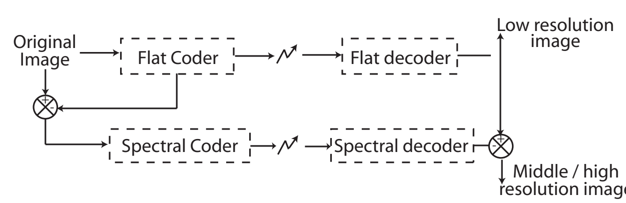

Figure 1: Overall two-layer LAR coding scheme - flat + spectral coders.

The basic concept of the LAR method is that local resolution, in other words pixel size, should be adapted to suit local activity. In the LAR codec the image is decomposed into blocks of different sizes. Those sizes depend on the local activity estimated through a local morphological gradient. This image decomposition is performed conditionally to a specific quadtree data structure, encoded in the FLAT coding stage. Thanks to this type of decomposition, block sizes implicitly give the nature of the given block: smallest blocks are located upon edges whereas large blocks map homogeneous areas. Then, the main feature of the FLAT coder consists of preserving contours while smoothing homogeneous parts of the image. This quadtree partition is the key system of the LAR codec and must be carefully computed.

This quadtree based image partitioning is exploited by the LAR features and services and can in particular be exploited to get a free hierarchical region representation: from the low bit-rate image compressed by the FLAT LAR, both coder and decoder can perform a segmentation process by iteratively merging blocks into regions. A direct application is then Region Of Interest (ROI) enhancement, by first selecting regions, at the coder or the decoder, and enabling second layer coding only for the relevant blocks. It can also be used as regions based chrominance coder for the lower bit-rates. An exhaustive description of the partitioning algorithm can be found in [5]. In order to code the FLAT and spectral layer the Interleaved S+P with entropy coder is used [2]. This specific Interleaved S+P coder enables other services such as scalabity in resolution and quality, data hiding and cryptography and is commonly used in the LAR.

3. IMAGE PARTITIONING

The LAR codec is based on a variable-size block representation relying on a homogeneity criterion. The proposed approach involves quadtree partitioning with all square blocks having a size equal to a power of two. This partition is denoted $P^{[N_{max} \dots N_{min}]}$ where N_{max} and N_{min} represent respectively the maximum and minimum values of authorized block sizes.

Many methods rely also on variable-size block representation. In particular, MPEG4-AVC/H.264 intra mode authorizes a $P^{[16,4]}$ partition (it splits images into 4×4 or 16×16 blocks), where size selection operates to produce the best bit rate/distortion from a PSNR point of view. Methods based on tree structures operate from the highest level (or maximal size) by cutting down nodes into sons when a given homogeneity criterion is not met. Several homogeneity tests can be found in literature [12], [8]. In most cases, they rely on computing a $L1$ or $L2$ norm distance between the block value and the value of its four sons.

Our coder result and all of its services deeply depend on the quadtree partitioning process that gives the coder an image representation. Since most features of our coder rely on it, the splitting process must be carefully tuned in order to have an image representation coherent to the human vision.

3.1 The partitioning process

The image partitioning is based on a quadtree topology originally deduced from a morphological gradient computation (difference between maximum and minimum luminance values on a given support).

The partitioning algorithm first extracts the minimum and the maximum pixel values of a given block and uses them to compute the homogeneity. The homogeneity is then compared to a given threshold T_H.

The partitioning process starts by splitting the image into uniform N_{max} square blocks and a quadtree is built on each of these initial blocks. The algorithm splits the block into four sons until either the homogeneity criterion or the block size N_{min} is reached. This algorithm is summarized by the figure 2.

In most color images three color planes are used to represent the colors, like RGB or YUV [14] color spaces. An identical quadtree partitioning is used for all color channels. The global algorithm is unchanged and the blocks are splitted if either one color channel does not meet the homogeneity criterion.

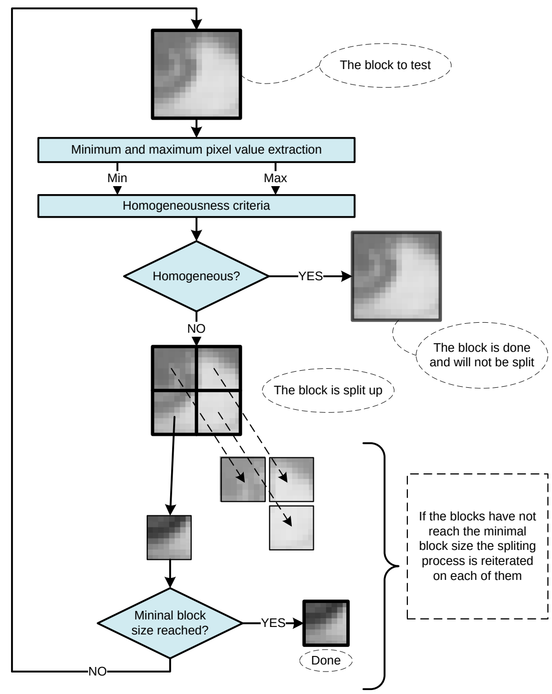

Figure 2: Partitioning Algorithm

3.2 Partitioning influence onto the coded image

Our coder performances are strongly related to this partition from lossy to lossless compression. The partitioning result has to be as accurate as possible, in terms of compression as well as representation. It is even more critical during lossy compression. Indeed the partitioning has to describe the image details accurately enough to restore visually important details without over-partitioning onto non significant areas. Over-partitioning causes an increase of the final rate as redundant data have to be encoded leading to decreased compression performances. Under-segmenting remains also critical in our scheme as necessary details are not represented in the Flat LAR layer. Therefore the texture coder has to compensate this loss of data which is very costly.

A good balance has to be found here in order to combine a good representation of the image content while performing well with lossless and lossy compression. Furthermore, keeping the same partitioning algorithm for lossy or lossless compression is of great importance for the method uniqueness.

The YUV criterion, originally used in our coder, does not consider Human Visual System (HVS) features. Thus the measured criterion does not fully relate to the perceived color difference (as human would do) and this can lead to a locally non consistent partitioning from the HVS point of view.

In order to correct these flaws, the partition criterion has to take advantage of the human vision. A perceptual color space such as the Lab one is then naturally preferred.

4. IMAGE PARTITIONING USING THE LAB COLOR SPACE

To take into account a psycho-visual criterion in the quadtree partitioning process, we propose to use the La*b* color space specified by the CIE (International Commission on Illumination)[6] as an analysis color space. This space is composed of a luminance channel (L) and two chrominance channels (a and b). In this space the Euclidean distance between two (L,a,b) points is supposed to be the visually perceived difference. Other improved metrics have been later proposed such as △E94 [3] or △E2000 [9].

To validate this color space for our coder we proposed to compare two partitioning criteria in the Lab space to the morphological gradient in the YUV one. More specifically, we evaluate our partition process with criteria based on the △E94 and the Euclidean distance as homogeneity criterion. Consequently the partitioning process aims at describing the edges as the human eye would actually see them. The choice of the △E94 rather than △E2000 is mostly justified by its relative computational simplicity.

4.1 Tested homogeneity criteria

All criteria take as inputs the minimum and the maximum pixel values on the three color planes on a given support. Thus, for YUV color space, $(Y_{max}, Y_{min}, U_{max}, U_{min}, V_{max}, V_{min})$ values are tested and for Lab color space, $(L_{max}, L_{min}, a_{max}, a_{min}, b_{max}, b_{min})$ values are tested.

4.1.1 *Morphological gradient: YUV color space*

$$\begin{array}{l} criterion = max(\triangle Y, \triangle U, \triangle V) \\ with: \\ \triangle Y = Y_{max} - Y_{min} \\ \triangle U = U_{max} - U_{min} \\ \triangle V = V_{max} - V_{min} \end{array} \quad (1)$$

4.1.2 *Euclidean distance: Lab color space*

$$\begin{array}{l} criterion = \sqrt{\triangle L^2 + \triangle a^2 + \triangle b^2} \\ with: \\ \triangle L = L_{max} - L_{min} \\ \triangle a = a_{max} - a_{min} \\ \triangle b = b_{max} - b_{min} \end{array} \quad (2)$$

4.1.3 *△E94: Lab color space*

$$\begin{array}{l} criterion = \sqrt{\triangle L^2 + \frac{\triangle C}{1+K_1 C_1}^2 + \frac{\triangle H}{1+K_1 C_2}^2} \\ with: \\ \triangle C = C_1 - C_2 \\ C_1 = \sqrt{(a_{max})^2 + (b_{max})^2} \\ C_2 = \sqrt{(a_{min})^2 + (b_{min})^2} \\ \triangle H = \sqrt{\triangle a^2 + \triangle b^2 - \triangle C^2} \\ K_1 = 0.045, K_2 = 0.015 \end{array} \quad (3)$$

5. RESULTS

5.1 Testing procedure

Input images are in RGB color format and each of these color planes are compressed in the RGB space by the Interleaved S+P compression algorithm with entropy coder (see figure 3). The image set is composed of five natural image: barba, lenat, parrots, peppers, and P06 from Microsoft JPEG database. The Interleaved S+P coder is set in a lossless profile and is driven by the quadtree partitioning of the image computed in either YUV space or Lab space. Compression level has not been changed in our analysis since the ideal partitioning threshold T_H has proved to be the same for all compression settings and because this analysis focuses on image representation.

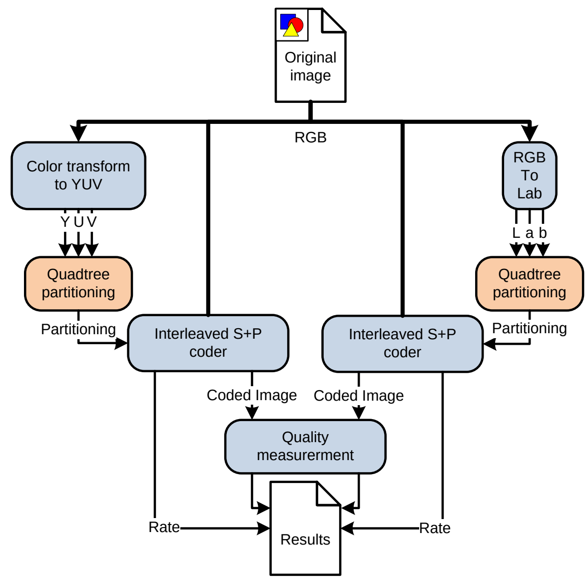

Figure 3: coder configuration

The images shown are issued from the FLAT coder. The visual quality of those images directly depend on the quadtree partitioning since FLAT images are splitted into blocks following the partitioning. The representation quality of the partitioning is then easy to visually evaluate.

In order to fairly compare the partitioning computed with the YUV color space and the Lab color space the ideal YUV partitioning threshold T_H is computed. The ideal threshold T_H is the one giving the lower bit-rate in lossless compression. The Lab partitioning is set with a threshold T_H giving the closest number of bloc to the ideal YUV partitioning. A reasonable assumption is that an equivalent number of blocks also gives an equivalent bit-rate and vice versa. Therefore the representations are compared at the best compression performances and at similar number of blocks.

5.2 Quality metrics

Image quality is usually evaluated with a PSNR based metric even if in the same time it is well none that PSNR is not well correlated to human observation. It is even worst with color images where three color planes have to be analyzed and some metrics do not give the same score in RGB or YUV color spaces. Several objective metric asses this issue by including human vision system properties in the score calculation improving the correlation with human observers.

In the LAR case the PSNR is the worst metric to consider [11]. It gives incoherent results, especially when comparing to other codecs. As an example the FLAT image barba presented in figure 4 is clearly improved by the Lab partitioning when the WPSNR_PIX shown in table 1 is worst for a comparable bit-rate.

Rate/distortion	YUV	Lab Euclidean	Lab Δ94
WPSNR_PIX	24.909dB	24.887dB	24.896dB
FLAT bit-rate:	1.243bpp	1.241bpp	1.238bpp

Table 1: *WPSNR_PIXRGB* and bit-rate, YUV and Lab representations (image: barba)

One explanation is coming from the fact that the LAR codec slightly move the contours witch is interpreted as an error by some metrics even if visually there is nothing noticeable.

C4 metric has been proved to be the most efficient metric on computing the LAR image quality and correlates the most with human observation [11]. This metric evaluates the quality from 1 to 5 as in subjective assessment rating where 1 is the poorest quality. This metric is used in the following results.

5.3 Comparison of YUV and Lab representation

Rate variation between the lossless compression with YUV partitioning and Lab partitioning is very small. Usually the Lab representation causes an average rate increase of 0.01bpp as the table 2 shows. Therefore the Lab partitioning can also be used for compression purpose without significant compression ratio loss.

bpp	YUV	Lab Euclidean	Lab Δ94
Barbara	13.567	13.577	13.566
Lena	13.394	13.393	13.391
P06	12.173	12.166	12.170
Parrots	9.791	9.833	9.797
Peppers	14.849	14.858	14.854
Average	12.755	12.765	12.755

Table 2: Lossless coding costs in bpp with different partitioning criteria and optimal threshold T_H

Rates being the same, representations quality remain to be evaluated and compared. The representation can be both evaluated visually and with the C4 metric. The C4 metric is used here by comparing the original image with the FLAT image.

In order to evaluate the representation difference between the several partitioning, FLAT images have to be looked at and visually compared. On some images the differences are barely noticeable but on others the Lab representation with Euclidean criterion clearly outperforms the YUV and Lab-Δ94 representation by introducing less noticeable artifacts. The differences usually occur in the red tone color and is especially visible in fading areas. YUV partitioning is less discriminant in the red tones and produces larger blocks while the Lab produces smaller and more visually accurate blocks. Figure 4 shows a good example of the superiority of the Lab-Euclidean representation.

The visual subjective analysis is correlated with the C4 metric analysis. Lab partitioning with the Euclidean criterion produces more visually accurate representation as shown in table 3.

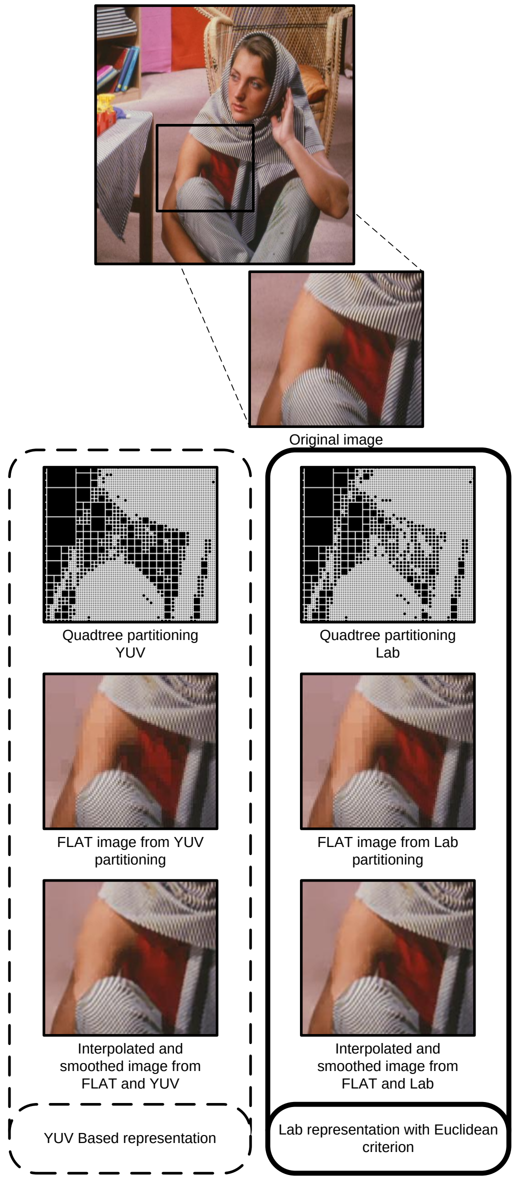

Figure 4: YUV versus Lab partitioning

C4 mesurement	YUV	Lab Euclidean	Lab Δ94
Barba	3.320	3.256	3.143
Lenat	3.703	3.916	3.734
P06	2.953	3.049	2.973
Parrots	4.393	4.363	4.299
Pimen	2.894	3.038	2.912
Average	3.453	3.524	3.412

Table 3: C4 quality assessments with different partitioning criteria and optimal threshold T_H

The LAR codec embeds a dedicated FLAT image smoothing and interpolation process currently at an experimental stage. This post processed FLAT image can be used as a low bit-rate final image and a better representation will naturally lead to better interpolated images as figure 4 shows.

Although criteria perform almost equally around their ideal threshold, experiments showed different behaviors around this point. For these experiments the thresholds are experimentally set for each criterion. This is not acceptable from an user point of view in a real case scenario. Therefore the threshold has to be found *a priori* without compression trials. In order to simplify an automatic estimation of this parameter, rate and quality have to remain stable around the ideal threshold. This robustness aspect is discussed in 5.4.

5.4 Robustness of the criteria

Our codec does not currently involve a specific technique to *a priori* determine the ideal partitioning threshold. For every image this parameter is different and depends on the nature and complexity of the image. Two setting approaches can thus be considered. The first one would be fully automatic where the threshold is set depending on an image analysis. The second approach would simply set the threshold to a standard value depending on the kind of image to process. In both approaches the threshold value set would probably be off the ideal threshold.

To observe the coding cost overhead produced by using a non optimal threshold T_H, a test procedure has been developed. First, for all images in the dataset and all partitioning criteria, the optimal threshold is experimentally determined. An average threshold along the different images is calculated for each criteria. These average thresholds are used to perform the partitioning and lossless coding of each image. Overhead coding cost is finally calculated by subtracting optimal coding costs from ones produced with average thresholds.

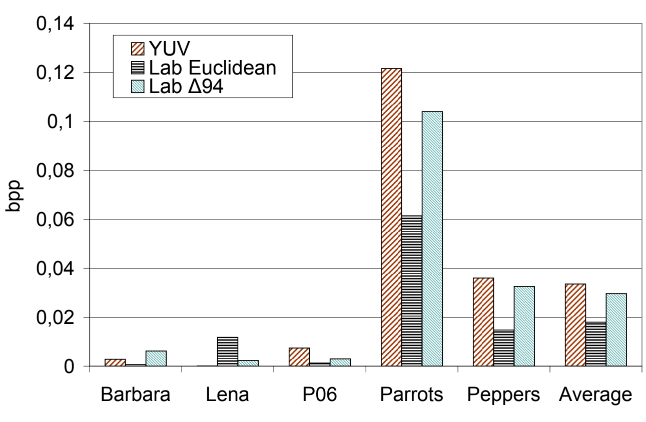

Figure 5: Lossless coding cost overhead due to non-optimal threshold T_H

As shown in figure 5 on average Lab Euclidean and Lab Δ94 criteria perform better than the YUV criterion. The coding cost overhead is reduced by half when using the Lab Euclidean criterion instead of the YUV criterion. This way, even if the threshold is off the ideal threshold the compression results and the quality remain close to the ones obtained with the ideal threshold.

6. CONCLUSION

The LAR image coding method relies on a variable block-size partitioning of the images. In this paper, we propose to introduce a perceptual color space (Lab) in order to increase the perceptual representation, and to study the impact on compression and representation. Results show that it is actually possible to combine a better representation efficiency with equivalent compression performances.

From all the Lab based criteria, Euclidean distance gives excellent results. Compared to the original YUV based criterion, it enables a better robustness against partitioning threshold value variations, a better visual representation combined with an equivalent compression efficiency.

These results open perspectives for further works especially concerning the region based color representation in the LAR coder. The accuracy of the regions representation is highly correlated to the partitioning efficiency. Moreover, the segmentation is currently performed in YUV space, and should be perceptually improved considering a perceptual color space.

REFERENCES

[1] W. A. B. Perceptual-components architecture for digital video. *Journal of the Optical Society of America. A: Optics and image science, volume 7, issue 10*, pages 1943–1954, Oct. 1990.

[2] M. Babel, O. Déforges, and J. Ronsin. Interleaved S+P pyramidal decomposition with refined prediction model. In *IEEE International Conference on Image Processing, ICIP'05*, pages 750–753, september 2005.

[3] CIE. CIE publication 116-1995 "Industrial Colour-Difference evaluation", 1995.

[4] O. Déforges and M. Babel. WG1N4692 - The LAR method for lossy and lossless image compression. Technical report, ISO/ITU JPEG commitee, Poitiers, July 2008.

[5] O. Déforges, M. Babel, L. Bédat, and J. Ronsin. Color LAR codec: a color image representation and compression scheme based on local resolution adjustment and self-extraction region representation. *IEEE Transactions on Circuits and Systems for Video Technology*, 17(8):974–987, August 2007.

[6] M. Fairchild. *Color Appearance Models*. John Wiley & Sons, Ltd, England, 2005.

[7] Z. Liu, L. Karam, and A. Watson. JPEG2000 encoding with perceptual distortion control. In *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, volume 1, pages I–637–40 vol.1, 2003.

[8] C. A. Shaffer and H. Samet. Optimal quadtree construction algorithms. *Comput. Vision Graph. Image Process.*, 37:402–419, 1987.

[9] G. Sharma, W. Wu, and E. N. Dalal. The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. *Color Research & Application*, 30:21–30, 2005.

[10] G. Sreelekha and P. Sathidevi. An improved jpeg compression scheme using human visual system model. *Systems, Signals and Image Processing, 2007 and 6th EURASIP Conference focused on Speech and Image Processing, Multimedia Communications and Services. 14th International Workshop on*, pages 98–101, June 2007.

[11] C. Strauss, F. Pasteau, F. Autrusseau, M. Babel, L. Bédat, and O. Déforges. Subjective and Objective Quality Evaluation of LAR coded Art Images. In *IEEE International Conference on Multimedia and Expo, ICME'09*, pages 1–4, Cancun Mexique, 07 2009.

[12] P. Strobach. Tree-structured scene adaptive coder. *Communications, IEEE Transactions on*, 38:477–486, 1990.

[13] A. Watson, G. Yang, J. Solomon, and J. Villasenor. Visibility of wavelet quantization noise. *Image Processing, IEEE Transactions on*, 6(8):1164–1175, 1997.

[14] W. Wharton and D. Howorth. *Principles of Television Reception*. Pitman Publishing, 1971.