Título del trabajo
Title of paper

Perceptually optimal image coding based on Human Visual System characteristics and uniform color space

Autor/ es
Author/s

Luca Leschiutta

Afiliación/ es del autor/ es
Affiliation/s of the author/s

Dipartimento di Automatica e Informatica - Politecnico di Torino

Dirección principal
Mail adress

Corso Duca degli Abruzzi, 24 - 10129 Torino, Italy

Teléfono, fax. e-mail de la persona de contacto
Phone, Fax number and e-mail address of the contact person

Tel. +390115647048
Fax. +390115647099
e-mail luca.leschiutta@polito.it

ABSTRACT

Aim of this work is to envisage an image coding technique capable to deliver a perceptually optimal output to be fed into a traditional loss-less compression algorithm such as Portable Network Graphics.
The main advantage of this two steps compression method, over lossy compression algorithms such as JPG, is that the amount of information that is discarded can be trimmed in order to meet the required output quality, the presentation device capabilities or the viewing condition.
This method does not rely on well established techniques such as Discrete Cosine or Wavelet Transforms but exploits the spatial distribution of similar pixels at microscopic level inside the image.
The coding is accomplished in two steps: in the first one, the image is converted into a perceptually uniform color space such as the CIE Lab. In the second one, each of the three image components is described via a set of subsequent rectangles, whose pixels can be associated through several rules that take into account the euclidean distance between pixels values and low perceptual impact considerations.
The coded image is therefore described with three sets of such rectangles in which an average color is associated to each rectangle.
The rectangles decomposition rules are chosen in order that the rectangles are small, in high spatial frequency regions, while the opposite is true in low spatial frequency regions. Furthermore a small amount of sensibly out of range pixels can be accepted in each rectangle.
Although the achievement of high compression ratios is outside the scope of this work, the coding alone introduces a noticeable reduction in the image size that is worth considering and computing.
In the final part of the work, the technique is tested on several images and the results, in terms of PSNR and perceived quality, are discussed.

1. INTRODUCTION

Discrete Cosine Transform and Wavelet based coding are well established techniques, for image compression, which deliver good picture quality and compression ratios [1][2].
Our goal, however, is to investigate the performance of compression algorithms that do not rely on transforms, but exploit the spatial distribution of similar pixels at microscopic level inside the image.
This work is a preliminary attempt, which is not focused on achieving high compression ratios, but tries to provide a different description of the raster image that can be later processed with an ad hoc loss-less compression algorithm.
The, hereafter described, “rectangle coding technique” is inspired by what is done in [3] with binary images.
Particular care has been placed in the investigation of rectangle coding technique biased by a Human Visual System (HVS) perceptual model.
The above has been achieved, eventually, via non-linear transforms in CIE Lab color space before processing

2. RECTANGLE CODING TECHNIQUE

This technique is based on decomposition of the image in small rectangular zones of homogeneous color or luminance.
The algorithm can be applied to any raster image, defined using an appropriate tristimulus color space [4], independently for each component of the color space. E.g. in a standard RGB image the 3 matrices R, G, and B can be processed to obtain a different description of the image as a set of 3 rectangle sequences <r>, <g> and <b>.
The following description, for simplicity purposes, is focused on gray scale images (in which it is needed to treat only one matrix), but can be easily extended to the color case repeating the procedure three times and replacing the luminance with the three color components.
Each rectangle is described by 3 parameters: height (Dy), width (Dx) and luminance (Y). The height and width are expressed in pixels, the luminance is normalized over 8 bits.
As will be discussed later, the main parameters, that can be varied, are the rectangles composition rules.
The simplest rule consists in dividing the 256 luminance levels in fixed width intervals and in searching for rectangles whose pixels all fit in one of these intervals.
One particular case is to have 256 intervals of color depth 1 (this will lead to a loss-less transformation of the image).
A preliminary analysis showed that the maximum allowed size of the rectangles should be 8*8 pixels for two reasons:

1) Larger rectangles would easily impair the image quality.
2) Considering a simple case, in which the width of the color interval is such that the image quality degradation is just noticeable, 99.9% of the rectangles found are smaller then 8*8.

The algorithm starts scanning the pixels matrix from the top left corner and verifying which, of the possible rectangles originating in that point and complying with the given set of rules, is the larger in size.
Once a rectangle is found, the corresponding pixels are marked as visited and the algorithm keeps on scanning what's left of the image, always proceeding left-right and top-down.
Some rectangles are considered “spurious” (i.e. smaller than 2 pixels) and are discarded (see figure 1).
As a result the image is described by a set of quadruples having the form <Dp_i Y_i Dx_i Dy_i>, where Dx_i and Dy_i are width and height of the i-th rectangle expressed in pixels, Y_i is the luminance of the rectangle and Dp_i is the distance from the rectangle (i-1)-th expressed in pixels (see figure 1).

Figure 1: The image is described as a set of rectangles which are uniquely identified by the quadruple <Dp_i Y_i Dx_i Dy_i>. S are spurious rectangles that are too small to be considered.

To establish the size in bits of an image coded in this way, it is necessary to analyze what are the maximum and average values of each of the above parameters.
These values depend upon the rules chosen to select the rectangles: the looser the rules the bigger the rectangles.
The following considerations are valid in the cases we are more interested in, i.e. image degradation perceptually just noticeable, and have been developed on the study of a 512 x 512 gray scale test image.
As said before, Dx and Dy are limited at 8 and according to some simple statistical considerations they can be expressed with 2.5 bits each.
Y is strictly determined by the rules chosen to find the rectangles (e. g. if the rule, on a 256 gray levels image, is to select rectangles whose pixels fall in a 5 level brightness interval, Y will be at most 52) and, in our test case, it can be conservatively expressed with 6 bits.
Dp is only limited by the width of the image, however, on a typical test image, it has an exponential distribution and can be expressed with roughly 2 bits.
As a result each rectangle can be described with 13 bits, given that not all the pixels of the image are defined.
We need, however, to better evaluate the discarding of spurious rectangles, from the decrease in image size and from the perceptual point of view.
As a matter of fact, if the rectangles were all subsequent, we would not need Dp in the coded image description and we would save 2 bits for each rectangle. The number of spurious rectangles is usually so high, however that the introduction of Dp is deemed profitable.
From a perceptual point of view, this is not cost free and the missing pixels shall be reconstructed with an appropriate algorithm in the decoding phase.
Two types of small rectangles can be distinguished: the ones surrounded by homogeneous areas and the ones included in steep brightness gradients. It is easily demonstrable that only the first type of rectangles should be discarded in order to have a good quality decoded image, wile the others will be included in the coded image.
The number of rectangles, and thus the size of the coded image, is strictly dependent on the set of rules chosen, however, to have a preliminary estimate, our test image, processed with 50 fixed width luminance intervals, produces roughly 60 000 rectangles and therefore the coded image will have a $6\ 10^4 * 13 / 8 \approx 100$ Kbytes size. However, a number of techniques can be developed to reduce this size, but these are outside the scope of this paper.

3. RECTANGLES DECOMPOSITION RULES

Aim of our study is to compare the quality of the coded image in function of the various rules chosen to select the rectangles in which the image is decomposed.
Three main methods have been followed to achieve this:
1) Fixed width intervals
2) Perceptually variant intervals.
3) Transformation in a perceptually uniform color space
In the first case the spectrum of each color is divided into a given number of intervals with a fixed size.
In the second case the width of each interval is varied according to a logarithmic law that simulates the response of the Human Visual System (HVS) to the various luminance levels.
In the third case, an evolution of the second one, instead of conceiving ad-hoc rules to segment the three different RGB color components, the image is first transformed into the perceptually uniform CIE Lab color space and then processed using the rectangle coding technique with fixed width intervals.
The main drawback of the third methodology is the computational complexity of CIE Lab non-linear transforms, however, given the constant increase of CPU performances, this is not considered as a major blocking point.
The transform–antitransform operation introduces on its own approximately 45 dB of PSNR that are completely unnoticeable and not worth accounting for.
In addition, another HVS characteristic we can use to discard some, less perceptually relevant, information, is the masking of high spatial frequencies. For this reason the acceptance of rectangles in which a small amount of pixels does not fit into the given interval it is deemed profitable; the result will be to have fewer, but bigger, rectangles in which some high spatial frequency information has been lost.

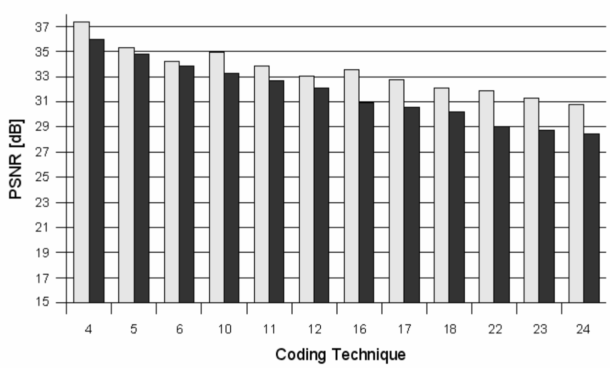

Figure 2: PSNR of coded image obtained changing the coding parameters as per table 1, in the case of fixed width intervals (light gray) and of perceptually variant intervals (dark gray).

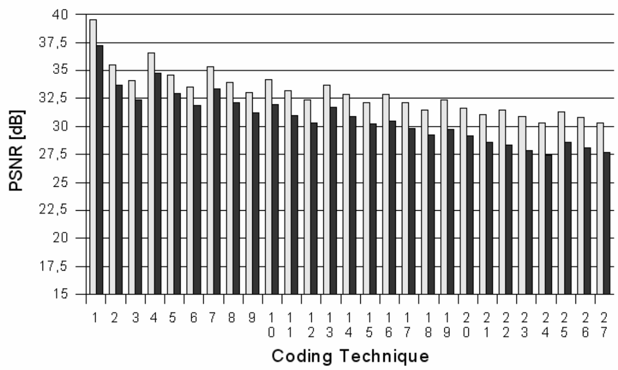

Figure 3 PSNR of coded image obtained changing the coding parameters as per table 1, in the case of fixed width intervals (light gray) and of CIE Lab variant intervals (dark gray).

4. IMAGE QUALITY ASSESSMENT

The rules described above have been used to code and decode several test images, varying the following parameters:
1) Total number of intervals
2) Amount of high spatial frequency masking
We have then assessed the relative quality of a coded image compared to the original one using both objective and subjective measurements. Even if in our study we are mainly interested in perceptive measurement, it is valuable to use an objective measure, such as PSNR, as a benchmark to compare similar images (see figure 2 and 3).
Another useful parameter to valuate the coding output is the Compression Ratio (CR) that is the ratio between the size of the original image and the one of coded image (see figures 4 and 5).

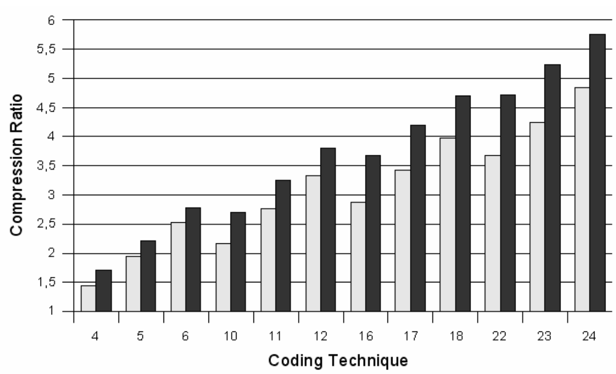

Figure 4 Ratio between original image size and coded image size obtained changing the coding parameters as per table 1, in the case of fixed width intervals (light gray) and of perceptually variant intervals (dark gray).

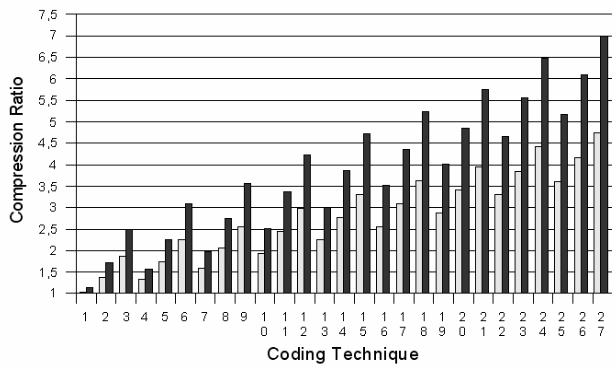

Figure 5 Ratio between original image size and coded image size obtained varying the coding parameters as per table 1, in the case of fixed width intervals (light gray) and of CIE Lab variant intervals (dark gray).

As far as the subjective image quality measurement is concerned, a test using the double-stimulus continuous quality-scale (DSCQS) method described in [5] has been performed.
Each assessor, after an opportune training period, has been invited to evaluate the quality of several images, rated on a 5 levels scale, against the original image (Bad = 1, Poor = 2; Fair = 3; Good = 4 and Excellent = 5).
The assessor had the capability to switch between the test and the original image at any time.

5. RESULTS

The reported objective results have been obtained over a set of test images, while the subjective test has been performed on the first image of [6]. In order to simplify the process only the most promising method (CIE Lab) has been subjectively tested against the fixed width interval one.
Figure 2 and 3 show that, with fixed width intervals, the PSNR is slightly higher (about 2 dB) than the variable cases. This result is expected given that the scope of our work, if considered by a different prospective, is to introduce higher objectively measurably errors that are not perceived so by the Human Visual System.
The analysis of the CR graphs (figure 4 and 5) tells us that some data compression is introduced by the algorithm even in the loss-less cases (Coding Technique 1), however this ratio is so small, compared to other compression method, that the rectangle decomposition algorithm alone doesn't seem promising from this point of view; e.g. Portable Network Graphic applied on the same set of reference pictures introduce a 1.85 compression ratio which is higher than the 1.13 of the CIE Lab method.
At any rate, the capability to decompose an image in perceptually homogeneous zones is meant to be exploited, in a lossy contest, together with some other traditional loss-less compression techniques.
The most interesting conclusion that can be drown from the a. m. graphs is that, the more we move towards lossy coding techniques, the higher is the compression ratio of

perceptually variable methods (most noticeably the CIE Lab one), with respect to the one of fixed width intervals methods.

To have a better understanding of the performance of the proposed methods it is necessary to proceed along two paths:

1) Subjective test.
2) Have a “fair” comparison in which each CIE Lab image is compared with a fixed width image having the same compression ratio or PSNR.

Figure 6 shows the results of the perceptive test in which are compared the vote given to images, obtained with different techniques of table 1, selected in order to have the same compression ratio and coded with the two methods (fixed width intervals and CIE Lab transform).

Figure 7 shows the results of the test in which are compared the votes obtained by images having the same PSNR.

From a general point of view the CIE Lab methods doesn't perform significantly better than the traditional one with high quality images.

Moving towards the images of poorer quality the opposite is true, with some outstanding cases where the compression ratio is about 2 – 2.3 (figure 6) and where the PSNR is around 28.5 – 29.5 dB (figure 7)

We need to understand if these results are repeatable. Considering that this method does not seem fit for high quality application, such as personal computer LCD or CRT monitors, we can concentrate on the 28.62 dB case of figure 7 that corresponds to coding technique 9 for CIE Lab and 23 for RGB. These characteristics are typical of the range of quality of devices, that are having a huge growth in these years (PDA, smartphones, picture viewing on home TV set, ...) and that are characterized by low resolution screens and low mass memory amount.

From an objective point of view, the results are confirmed by the analysis of aggregated data reported in figure 3 and 5 for the a.m. techniques

From a visual inspection point of view, instead, we detect a small problem in the high light color backgrounds of some images and this is probably due to the CIE Lab transformation. However we think that this could be easily corrected, at no extra cost for the coded image size, just adding, in the decoding phase, some filtering to smooth the artefacts visible on these backgrounds.

6. CONCLUSIONS

In this paper we tested a general purpose algorithm for images compression.

The main characteristics of this algorithm are:

- it does not rely on DCT or other transforms;
- it divides the image into small homogeneous rectangular zones;
- it is scalable from loss-less to lossy;
- it is possible to adapt it according to the HVS perceptive characteristics;
- it is possible to identify coding techniques that are perceptually optimal given a certain presenting device.

From a compression ratio point of view, this algorithm alone is not satisfactory for high quality images, because there are too many small rectangles and, particularly for loss-less coding, there is almost no compression at all (CR = 1.13).

However, using CIE Lab transform to code an image provides some more interesting results, as it is possible to identify cases in which it performs much better with respect to comparable non perceptual algorithms

The next steps into this work consist in:

- identify a traditional loss-less algorithm that performs well on the CIE Lab coded image;
- apply some filtering of the high light backgrounds in the decoding phase;
- extend the algorithm in the video compression research field in order to build parallelepipeds over a Group Of Pictures. The base of each parallelepiped will be one of the rectangles found with the 2D algorithm while the height will represent the number of frames in which the base is identifiable with a given amount of uncertainty.

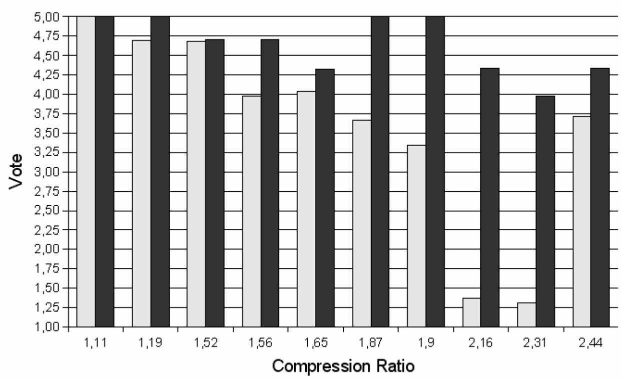

Figure 6: results of the perceptive test comparing images obtained with different coding methods and having similar Compression Ratio.
Legend: light gray – Fixed width intervals
dark gray – CIE Lab perceptually variant intervals.

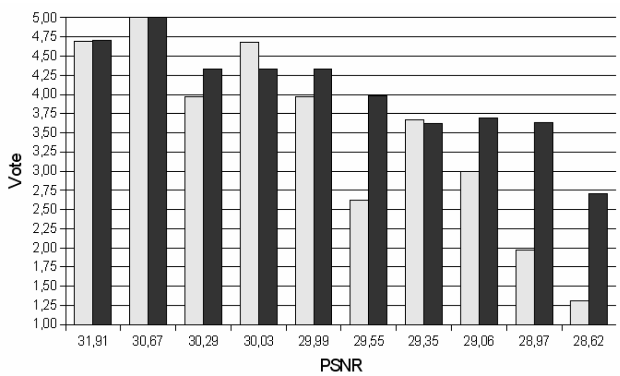

Figure 7: results of the perceptive test comparing images obtained with different coding methods and having similar PSNR.
Legend: light gray – Fixed width intervals
dark gray – CIE Lab perceptually variant intervals.

7. REFERENCES

[1] Watson A.B., "Image Compression Using the Discrete Cosine Transform", *Mathematica Journal*, 4(1), 1994, p. 81-88

[2] Samet A., Ayed M.A.B., Loulou M., Masmoudi N., "Perceptual Evaluation of JPEG2000", *European Transactions on Telecommunications*, Volume 15, Issue 2, 2004, Pages 135-143

[3] Mohamed, S.A.; Fahmy, M.M, "Binary image compression using efficient partitioning into rectangular regions", *Communications, IEEE Transactions o*n , Volume: 43 , Issue: 5 , May 1995, Pages:1888 - 1893

[4] Poynton C. A., "Frequently asked questions about color", *http://www.poynton.com/ColorFAQ.htm*l, 2000

[5] Rec. ITU-R BT.500-11, "Methodology for the subjective assessment of the quality of television pictures", 2002

[6] http://r0k.us/graphics/kodak/ Kodak Lossless True Color Image Suite, 2004

CT	CI	AE	CT	CI	AE	CT	CI	AE
1	1	0	**10**	4	0	**19**	7	0
2	1	2	**11**	4	2	**20**	7	2
3	1	4	**12**	4	4	**21**	7	4
4	2	0	**13**	5	0	**22**	8	0
5	2	2	**14**	5	2	**23**	8	2
6	2	4	**15**	5	4	**24**	8	4
7	3	0	**16**	6	0	**25**	9	0
8	3	2	**17**	6	2	**26**	9	2
9	3	4	**18**	6	4	**27**	9	4

Table 1: Coding Technique (CT) used for the simulation. CI represents the width of the color intervals in each color component. AE represents the amount of variation, from the given color interval, accepted to exploit the high spatial frequency masking.