White paper on image compression using the Fast Wavelet Transform

Daniel Lemire[†] , Ph.D. and Alain Béliveau[‡] Ph.D.

This white paper presents the image compression technology developed by *Compression d'Images en réseaux et Automatismes* (CIRA).

I. Introduction

Wavelet theory is still somewhat recent. We usually consider the French engineer Morlet as the inventor (beginning of the eighties). Later on, a French mathematician, Meyer, was the first to give a mathematical foundation to this early theory (1986). However, history will remember the American mathematician Daubechies as the first to come up with wavelets of industrial strength (1992).

Daubechies says that she imagined her wavelets while she was in Montreal invited by professors Deslauriers and Dubuc at the University of Montreal, both experts in iterative interpolation which is one of the pillars of the modern wavelet theory. Since it was during winter in Montreal and it was very cold, she was stuck in her hotel room forced to work...

One of the first really exciting application of wavelets is the result of a study done by the FBI concerning a new compression algorithm for fingerprints. Indeed, in the United States alone, 40 000 fingerprints must be treated, classified and compared every day. Beside the FBI, all of the police force in the United States must manage multiple fingerprint databases and the costs are rather high. It was first believed that the choice of the new compression algorithm would fall on the JPEG standard (*Joint Picture Expert Group*) which is based on a DCT (*Discrete Cosine Transform*) on blocks of 8 by 8 pixels (see figures 1 and 2). In the case of the FBI, images are monochrome (256 tints), are always of the same size and finally, all look similar. In figure 1, we can see very clearly the 8 by 8 blocks of the compressed image on the left.

[†] Biomedical Engineering Institute, University of Montreal, Montreal, Canada, Tel. (514) 524-3517, Fax (514) 524-0141, email Daniel.Lemire@Videotron.Ca

[‡] Applied Mathematics Department, *École Polytechnique de Montréal*, Montreal, Canada, Tel. (514) 340-4711 #4292, Fax (514) 340-4463, email Beliveau@MathAppl.PolyMtl.Ca

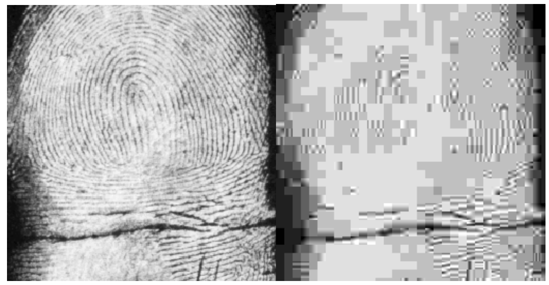

Figure 1. JPEG compressed fingerprint

While JPEG does a very good job for lower compression ratio (5:1 to 10:1), results are often unacceptable at ratios of 15:1 to 20:1. It isn't often the case with wavelets. Wavelets are now used every day by the FBI and a large fraction of the police force in the United States to process thousands of fingerprints every day.

picture	DCT on 8x8 blocks	quantization	entropy coding

Figure 2. JPEG Algorithm

One of the advantages of wavelets is that it is a global approach and not a block by block approach : the blocking artifacts are rare contrary to JPEG.

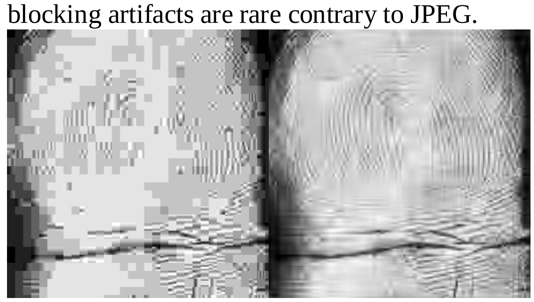

Figure 3. Qualitative comparison between JPEG (on the left) CIRA technology (on the right).

II. PROPOSED SCHEME

At first, the researchers working for the FBI had imagined using optimization algorithms such as *Matching Pursuit* (Stéphane Mallat) or *Best Basis* (David Donoho). However, this type of algorithm is slower than the *Fast Wavelet Transform* and the improvement can be small. The chosen model is an astute compromise. The FBI observed that the optimization algorithms always made the same choices. From this fact, it was easy to get a very fast algorithm without the optimization overhead. They could achieve this because fingerprints are fairly similar in general.

However, CIRA must use a more flexible algorithm which is not necessarily optimal for a given image set, but which performs well in general. In short, we must provide an alternative to JPEG which must be clearly better in some cases, in particular when compressing aggressively large images.

One of the particular concern for CIRA is the time to uncompress an image. Also, since the algorithm must replace the JPEG format, it must use 24 bits (also called *true color*).

A. Channel-based approach

Instead of considering an image as being a block of pixels, it is useful to use a n-channel representation (see figure 4). The most famous n-channel representation is certainly RBG (*Red-Green-Blue*). Beside the mathematical concepts, there are three reasons to use a n-channel model :

1. the functions called are the same to process a gray image (8 bits and 1 channel) and a true color image (24 bits and 3 channels);
2. it is then easy to optimize the algorithm for human perception (see below);
3. it is then possible to program the display of an image in a progressive manner, one channel at a time.

The second point must be explained further. The human eye doesn't see very well chromatic shifts (Cb/Cr ou I/Q), but is very sensible to luminance Y. It is why researchers recommend YCbCr or YIQ models (and not RGB) based on this cognitive fact. We get better compression ratios by using weaker channels for chromatic components (Cb/Cr ou I/Q). In television, they use this same approach in digital cameras : the Cb and Cr (or I and Q) components are undersampled four times.
A even greater flexibility can be achieved by considering each wavelet frequency band as a channel within each “colour” channel.

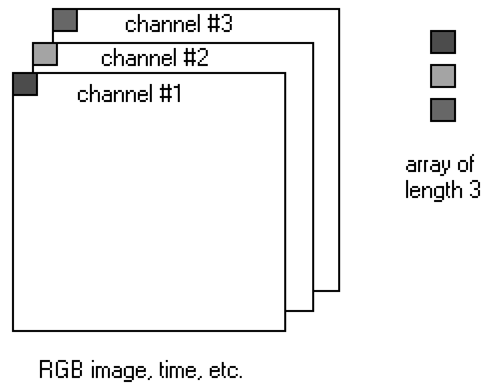

Figure 4. Channel-based representation of an image

B. Wavelets

We use the *Fast Wavelet Transform* on each channel. These last few years, two types of wavelets were found to be best for image compression : Cohen-Daubechies-Feauveau and symmetric Daubechies (symmlet). All these wavelets are (almost) symmetric. For this project, the wavelets are those chosen by the FBI : the symmetric Daubechies wavelets. They are chosen because the information content is never moved in the coefficients blocks during coding (see figure 5).

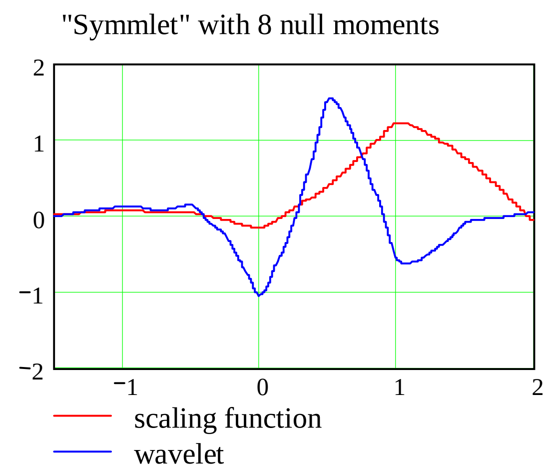

Figure 5. Wavelets chosen for the compression algorithm

C. Zerotree

The wavelet coefficients are not coded using bytes (values between 0 and 255), contrary to the original channel-based coding of the image, but in this case, by floating numbers. Therefore, there has been no compression so far, quite the contrary since floating numbers are more expensive to code than bytes. However, without getting into the details, it is possible, in this new representation, to set a large number of coefficients to zero (by using a threshold) without changing the image much. We can then code cheaply the zeroes using a *zerotree* which takes advantage of the fact that if in a certain area of the image, there are no sharp edges, all corresponding wavelet coefficients will be zero. This set of zero coefficients is what we call a zerotree.

D. Quantization

Following these steps, and before saving the compressed data, all floating numbers must be converted to integers (*quantization*). Since the eye sees according to a logarithmic scale, the error margin is then very small when approximating small coefficients, but much larger for larger coefficients. This quantization can be multi-channel or channel per channel.

E. Entropy coding

Once the floating numbers have be converted to integers, we need to do entropy coding. Indeed, just like the Morse code takes advantage of the fact that, in English, the letter e comes up more often, and so codes it very efficiently (using a single dot), we can compress considerably the image by using probabilistic techniques like *Huffman coding* or LZ which are still the fastest approaches for raw data. To help entropy coding further, we use a variable bit rate per scale, giving more importance to large scale coefficients and less to small scale ones.

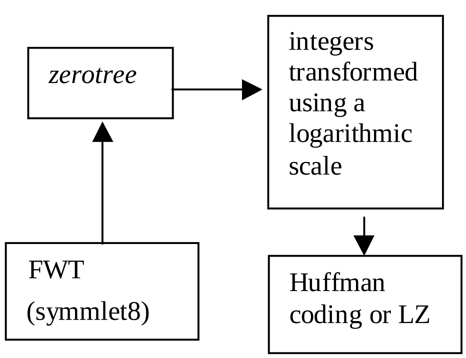

Figure 6. Proposed compression scheme

III. RESULTS

We have tested all of this algorithm with a Java prototype and then with a C/C++ version (Microsoft Visual C++).

The C/C++ version is made of two DLLs which go respectively from RGB to CIRA COD and from CIRA COD to RGB. From these DLLs, it is possible to write software compressing just about all image formats. As an example, we wrote executables compressing JPEG files and extensive testing was performed.

As far as speed is concerned, the Java prototype would compress and uncompress an image in about 4 s whereas the C/C++ version will do the same job in about 1 s. A rewrite of the Java version will bring it closer to the C/C++ version.

The qualitative results are presented in figures 7, 8 and 9. We see that in addition to a much better compression rate (about 6 times better) the CIRA COD format allows a much better qualitative result.

Figure 7. Original image (261 KB)

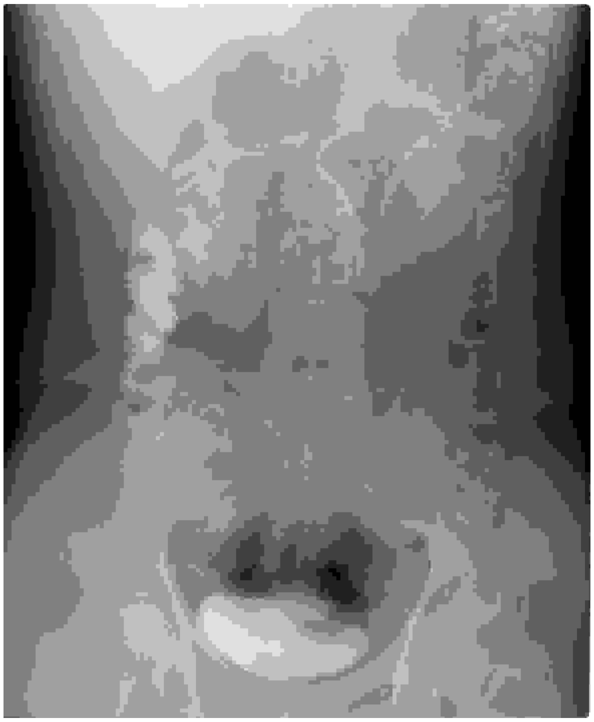
Figure 8. JPEG compressed image (28 KB)


Figure 9. CIRA compressed image (5 KB)

I. Options

This software can be extended, in particular the following options are available :

1. special filters on the wavelet coefficients, in particular, for mammography applications;
2. image completion upon request to add missing information;
3. progressive display of the image;
4. display of only a small part of the image;
5. Fast image segmentation using the COD format (muscle, bones, etc.);
6. XML interface to control preceding options (2 to 5) using a browser;
7. updates to emerging standards (JPEG2000, etc.);
8. update of the algorithm according to progresses in entropy coding technology;
9. port to other OS (PowerPC, RISC, Java, DSP chips, etc.).

Société de Compression d'Images en Réseaux et Automatismes (CIRA)
30, allée du Vert Galant - 78240 CHAMBOURCY
Tél 01 39 30 42 00
Fax 01 39 30 42 09