A SMALL FOOTPRINT, STREAMING COMPLIANT, VERSATILE WAVELET COMPRESSION SCHEME FOR CAMERAPHONE IMAGERS.

L. Alacoque[1], L. Chotard[2], M. Tchagaspanian[1] and J. Chossat[2]

[1] CEA LETI-MINATEC/SCME, 17 rue des martyrs, F-38054 Grenoble, Cedex 9, France.
[2] STMicroelectronics/Imaging Division, 12 rue Horowitz, B.P. 217, F-38019 Grenoble Cedex, France.

{laurent.alacoque,michael.tchagaspanian}@cea.fr,{ludovic.chotard,jerome.chossat}@st.com

contact author: Laurent Alacoque, tel: +33 438 782 972, fax: +33 438 789 073

INTRODUCTION

The sustained effort towards higher resolution imagers has been a key point for their success. One of the drawbacks of this trend is that imaging systems are encountering difficulty to accommodate the considerable amount of data of high resolution frames. Image sensors now require several high speed 1Gbps output links to encompass the transmission bandwidth of high resolution images at video rate. Additionally, high-level image processing algorithms tend to require the storage of an increasing number of frames whereas storage ability is limited to keep system costs low.
Image sensors are therefore in need for an effective compression scheme that can accommodate imagers severely complexity constrained environment.
Although extremely efficient, compression algorithms such as JPEG rely on complex processing and require concurrent access to large parts of images, which constrains their implementation to the ultimate stage of the image processor. State of the art low complexity image compression techniques such as JPEG-LS [1] greatly helped to reduce the complexity gap. Some noticeable improvements to JPEG-LS were suggested by Bazhyna in [2] to make it more nicely accommodate linear, gamma uncorrected, Bayer CFA images. However, although efficient and compact, JPEG-LS was not designed to guarantee a fixed bitrate and can therefore not be considered for real time transmission compression. Moreover, its bitstream offers limited means of random accessibility.

This paper presents a compression scheme that aims at providing original answers to the data compression challenge of severely constrained imaging systems. It was designed to be efficient yet tiny enough to be embedded on a cameraphone image sensor, to be compliant with rolling-shutter readout scheme, to rely on a very limited (typically one line) storage capacity and to provide versatile compression modes including lossless and lossy compressions, with the abilities to target a given compressed size, to strictly enforce a bitrate and to guarantee random accessibility and decoding of the compressed bitstream.

SYSTEM OVERVIEW

Figure 1 shows a typical imaging system overview along with the targeted implementation locations for the proposed compression block and figure 2 illustrates the compression and decompression architectures. Data compression is achieved by three steps. Data is first fed to an invertible transform which is responsible for reducing signal randomness by reducing spatial redundancy; the second (optional) step is coefficient re-quantization which greatly enhances compression to the price of a hopefully imperceptible loss in image quality. The final step is effectively responsible of data compression by assigning each pixel value a binary code whose length is dependent of its probability of occurrence, hence approaching Shannon's source entropy.
Decoding process being very similar to coding process, the remainder of this paper will focus on the encoding phase.

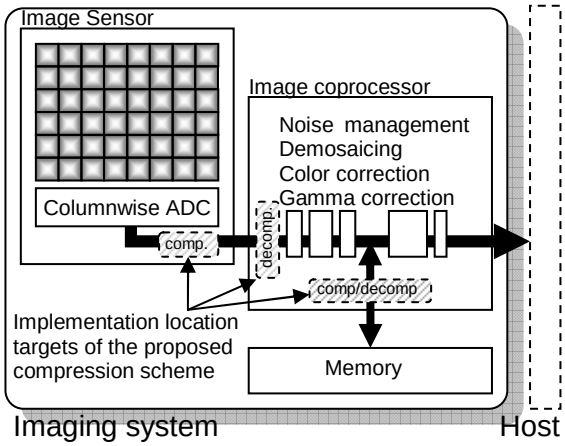

Figure 1: Imaging system overview with target implementation locations

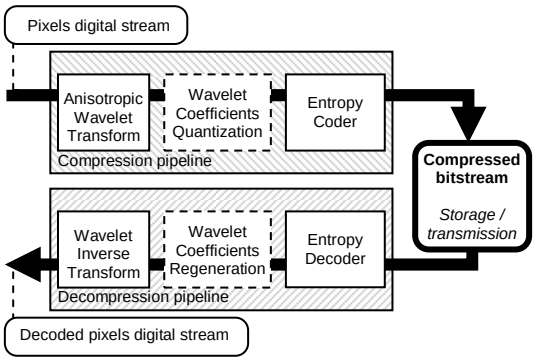

Figure 2: Compression and decompression pipelines overview

Color plane deinterleaving

The colour channels difference of Colour Filters Array raw images is responsible of a high frequency component in the raw signal. This behaviour greatly enhances signal entropy and limits signal compression. Standard methods for Bayer colour channels de-interleaving are to separately encode blue, red and green channels. Red-greens and blue-greens can be encoded separately or as a whole following *merge* or *rotate* variants. A higher correlation and hence a

better compression is assumed when green pixels are encoded as a whole. This was proved to be true but of very limited (0.1 bit per pixel in average) efficiency. The reader is referred to [2] for an in depth illustration.
In [3], Zhang showed that a specific wavelet decomposition, the Mallat wavelet packet transform, can directly and nicely handle Bayer data and exhibits better compression performances than state of the art JPEG-LS algorithm. Although efficient, the fifth order wavelet decomposition proposed in this paper has the major drawback to require the concurrent access to several lines of the image, hence impacting the required memory and streaming ability that are two severe constraints of cameraphone imaging systems.
In the proposed approach, de-interleaving is done at the pixel stream memory stage. Each colour plane is separately fed to the compression pipeline as colour-uniform blocs.

Low data dependency signal entropy reduction

Signal randomness attenuation is usually achieved by either transform or prediction based approaches.
In JPEG-LS, a predictor is used to guess the most probable value of the current pixel based on a limited knowledge of its neighbours [1]. This prediction can be reproduced on the decoder side and allows the encoding of the pixel value by transmitting only the (much lower entropy) difference between prediction and actual pixel value.
Alternatively, transform based approaches like baseline JPEG Discrete Cosine Transform, or JPEG 2000 Wavelet Transforms take advantage of the stationary behaviour of natural images to segregate image information into few high energy meaningful coefficients and many low energy highly predictable ones. Transform based approach further have the interest of presenting image features on a very useable form for post processing. For example defects, textured zones, noise and edges are easily identified by high magnitude coefficients and can readily be processed in the transform domain. This point is emphasized by state of the art transform based demosaicing and noise filtering algorithms.
The proposed entropy attenuation scheme relies on a very low complexity wavelet transform. The transform handles 2x16 pixels blocs and produces 32 coefficients wavelet packets. The transform kernel is based on a slightly modified fourth order Haar wavelet transform with the first transform level being applied vertically and horizontally and the three following levels horizontally only.
This original rectangle shape decomposition has a very low (1 line) data dependency and makes it compatible with pixel stream inputs at the moderate price of a one line memory. Its average complexity is below 3 add operations by pixel. This figure is believed to be lower than the predictor of JPEG-LS and should be compared to 64 multiply accumulate by pixel in the case of baseline JPEG.

Figure 3 shows the effect of anisotropic wavelet decomposition on a raw image colour plane and its effect on image histogram.

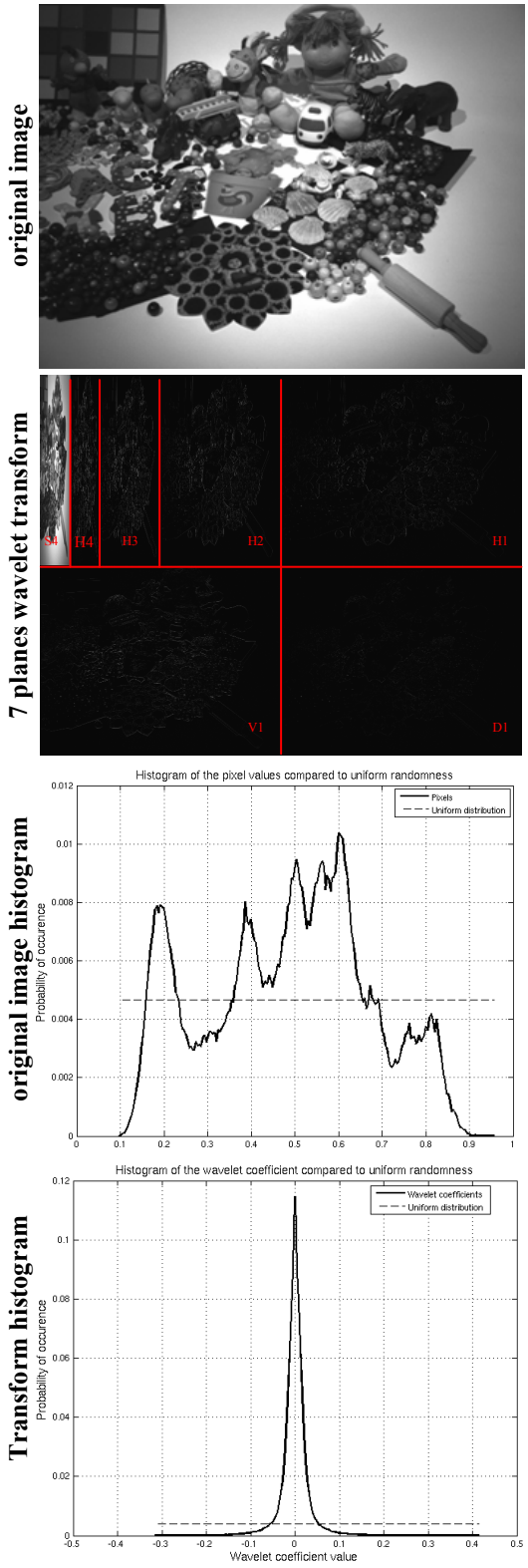

Figure 3: Original image (top) and its wavelet transform (second row) along with their histograms (bottom). The dashed lines represent uniform randomness. As it can be seen, raw image histogram resembles uniform randomness whereas wavelet coefficients are highly predictable.

Variable length entropy coder (VLEC)

Wavelet transform histogram clearly shows that wavelet coefficient obey a Laplacian distribution which denotes an exponentially decreasing probability of occurrence when

events move away from zero. Golomb codes [4] were proven to be optimal for exponentially decaying (geometric) probability distributions. Given a parameter *m* each Golomb code of a symbol *n* is made out of two codewords, a position coded MSB part [n/m] and a fixed length LSB part [n mod m]. Rice [5] proposed a very low complexity implementation of Golomb codes by limiting the parameter *m* to the powers of two. JPEG-LS uses Golomb-Rice codes at the VLEC stage and make use of contexts and embedded alphabet to determinate *m* at each pixel without explicitly sending it.

The proposed variable length entropy coder was designed to be straightforward enough to allow on the fly encoding. It relies on Golomb-Rice encoding of a full wavelet packet. This method allows the computation of the optimal *m* parameter that minimizes the packet encoded size. This parameter is explicitly transmitted as a header of each compressed packet with a limited overhead of a few bits per packet (typically 3 bits per 32 pixels).

Figure 4 shows performance results for the variable length entropy coding of the coefficient of the wavelet H1 plane. Red line shows each symbol intrinsic entropy (optimal codeword length) and blue line represents VLEC assigned codewords lengths.

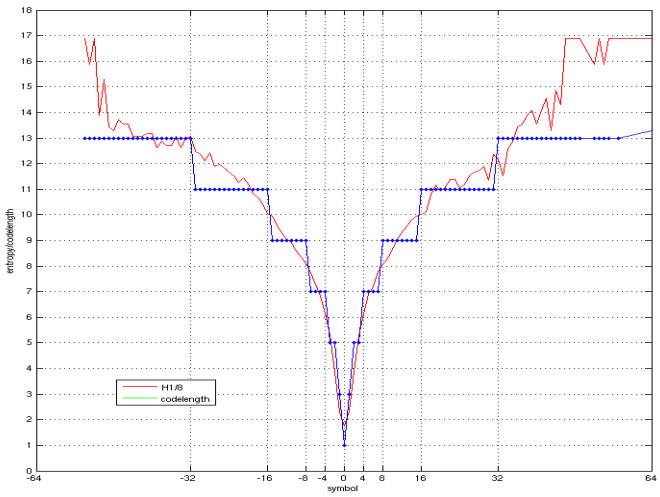

Figure 4: Symbol entropy and codeword assigned length (bits) vs. symbol value (log).

Progressive transmission, random accessibility and constant bitrate

The VLEC was designed to allow progressive mode encoding. In this mode of operation, each new transmitted bit is required to be of lower importance than the previously sent ones. This allows stopping bit transmission at any time and still guaranteeing the best possible reconstruction.

The proposed VLEC ensures progressive mode by sending bits of the 32 wavelet coefficients following bitplanes and coefficients prioritization to ensure that each bit is transmitted by order of quality relevance.

This feature is used by the constant bitrate mode where packet transmission is stopped after a predefined number of bits, thus enforcing a constant bitrate.

In constant bitrate mode, each 2x16 pixel area is encoded on a known number of bits which makes it easy to compute the packet start address in the compressed bitstream. This, in turn, provides a way to randomly access pixels embedded in the compressed stream.

Optional lossy compression

Experimentations on a batch of 10 bits per pixel raw Bayer images from a 3 Mpixels cameraphone exhibit compression rates ranging from 6 to 7 bits per pixel in lossless mode; this is comparable although slightly less efficient than state of the art small footprint Bayer level compression presented in [2].

To achieve higher compression rates, lossy compression is needed; it is achieved by reducing wavelet coefficient precision, hence inducing some error in the reconstruction. However, selecting which bits to remove out of 32 wavelet coefficient spread on seven planes of different importance is not straightforward.

A statistical approach allowed us to model coefficient requantization impact on both Mean Square Error and compression gain. Figures 5 and 6 show the comparison of both models (dashed lines) and actual observations (colour lines) for each wavelet plane.

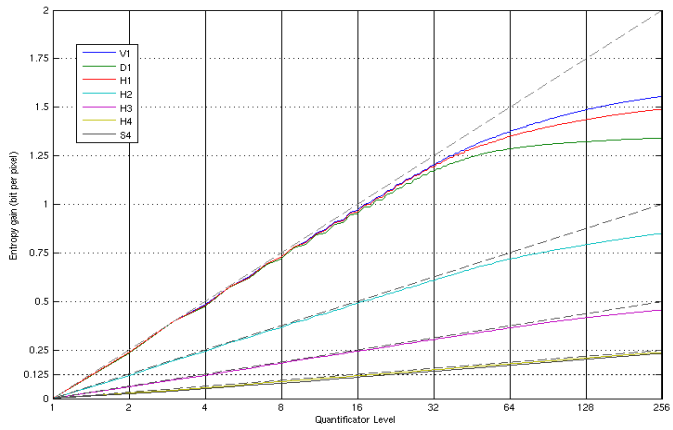

Figure 5 Compression gain (bpp) vs. Coefficient requantization level for each wavelet plane. Predicted (dashed lines) and observed (colour lines)

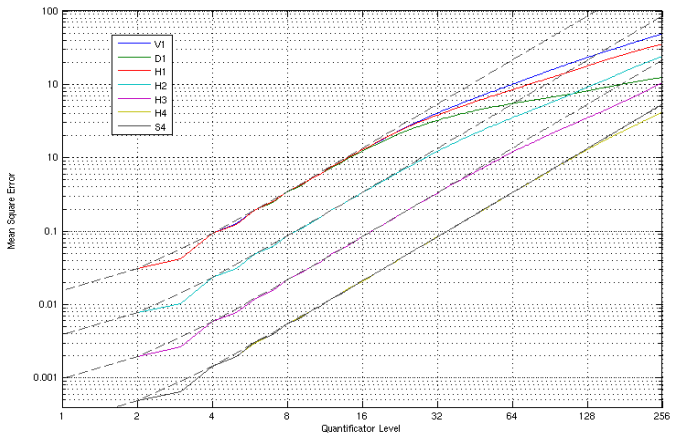

Figure 6 Induced Mean Square Error vs. Coefficient requantization level on each wavelet plane. Predicted (dashed lines) and observed (colour lines).

By coupling the two approaches, we were able to link the target compression to statistically optimal quantifier values in a way that guarantees minimal mean square error introduction hence approaching the optimal compression vs. quality trade-off.

To keep complexity low, this requantization process was limited to powers of two to allow coefficients requantization to be performed by simple level shifters. The coefficient requantization stage can operate in two modes. In target compression mode, quantization parameters are computed to approach a given target compression. In maximum error mode, quantization parameters are computed to obtain the maximum compression given a maximum allowed reconstruction error.

RESULTS

The presented compression scheme is fully described by a bit-true implementation ready model and associated micro-architecture. Processing complexity is estimated to be around 20,000 NAND-equivalent gates (excluding memory) with a surface of 0.04 mm² in 65nm technology which makes it suitable for focal plane implementation even on cameraphone image sensors.

Figure 8 shows an example of an uncompressed and a 2.9 bits per pixel lossily compressed image and figure 7 shows PSNR versus compression performances of the proposed compression block. The vertical bold line at 10 bits per pixel represents uncompressed image, the solid line represents achieved compression and the dashed line the wavelet coefficient entropy.

The difference between entropy and actual compression shows that current implementation is only limited by the VLEC for high compressions thus keeping room for improvement for high compression at constant quality.

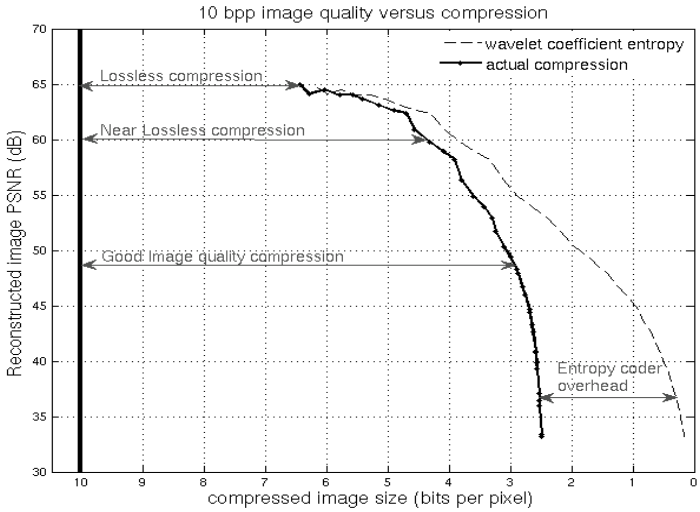

Figure 7: Reconstruction PSNR vs. image compression for a Bayer 3 MPix 10 bits cameraphone image (solid line).

REFERENCES

[1] M.J. WEINBERGER, G. SEROUSSI AND G. SAPIRO, "The LOCO-I lossless image compression algorithm: principles and standardization into JPEG-LS", IEEE Trans. on Image Processing, vol. 9, Aug. 2000 pp. 1309-1324.

[2] A. BAZHYNA AND K. EGIAZARIAN, "Lossless and Near Lossless Compression of Real Color Filter Array Data", IEEE Transactions on Consumer Electronics, vol 54, N°4, nov. 2008, pp. 1492-1500.

[3] N. ZHANG AND X. WU, "Lossless Compression of Color Mosaic Images", IEEE Transactions on Image Processing, vol. 15, N°6, jun. 2006, pp. 1379-1388.

[4] S.W. GOLOMB, "Run-length encodings", IEEE Transactions on Information Theory, vol. 12, jul. 1966, pp. 399-401

[5] R.F. RICE,"Some practical universal noiseless coding techniques – part iii", Technical Report JPL-91-3, Jet Propulsion Laboratory, Pasadena, CA, mar. 1979.

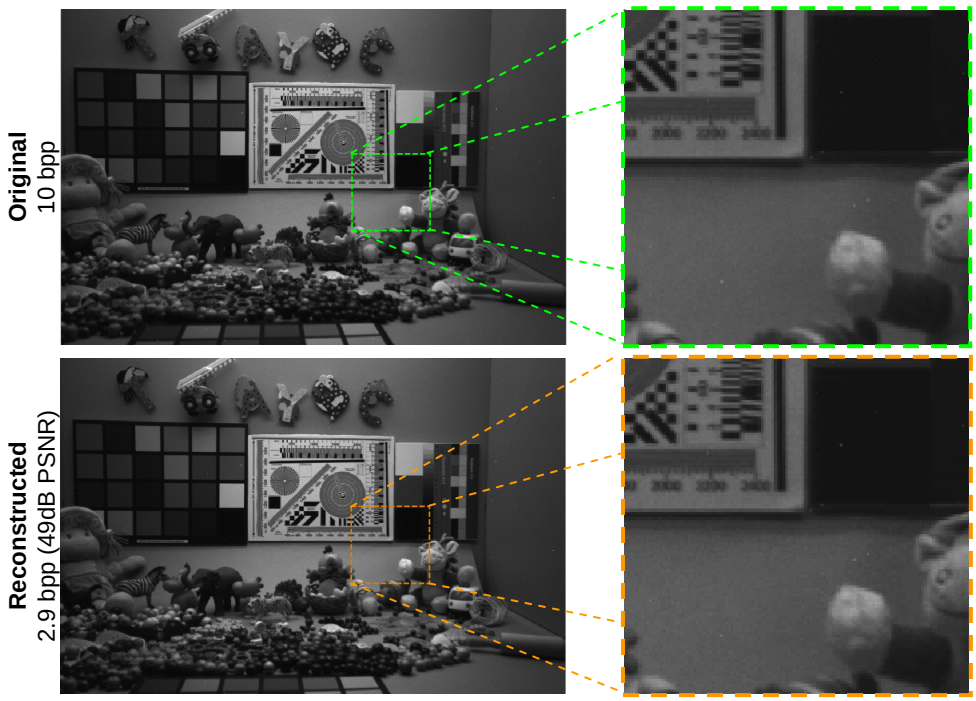

Figure 8: Detail of a 3MPixel original raw image at 10bpp, for clarity only showing green colour plane (top), lossily compressed image at 2.9 bpp / 49dB PSNR (bottom)