Next Generation Image Compression and Manipulation Using CREW

Martin Boliek, Michael J. Gormish,
Edward L. Schwartz, Alexander Keith

29 July 1997

CRC-TR-9728

Keywords:

CREW Image Compression

RICOH California Research Center
2882 Sand Hill Road, Suite 115
Menlo Park, CA 94025
(1 415) 496-5700, (1 415) 854-8740 fax
Internet: cip@crc.ricoh.com, WWW: http://www.crc.ricoh.com/CREW

Next Generation Image Compression and Manipulation Using CREW

Martin Boliek, Michael J. Gormish, Edward L. Schwartz, Alexander Keith
RICOH California Research Center, 2882 Sand Hill Road, Suite 115, Menlo Park, CA 94025
crew@crc.ricoh.com, http://www.crc.ricoh.com/CREW/

Abstract

As the applications of digital imagery expand in resolution and pixel fidelity there is a greater need for more efficient compression and extraction of images and sub-images. No longer is it sufficient to compress and decompress an image for a specific target device. The ability handle many types of image data, extract images at different resolutions and quality, zoom and pan, and to extract regions-of-interest are the new measures of image compression systems.

CREW is a high quality image compression system that is progressive from high compression to lossless, pyramidal, supports regions-of-interest, and multiple image types. This paper describes the CREW system and format, shows how the correct data can be quickly extracted from a CREW file to support a variety of target devices, describes the mechanisms needed for panning, zooming, and fixed-size compression, and explains the superior performance on bi-level and graphic images.

1 Introduction

Wavelet transforms have been studied for use in image compression systems for about ten years [1]. The results are generally considered superior to Discrete Cosine Transform systems, such as the JPEG Still Image Compression standard [2][3]. Because wavelets have compact support and are overlapped, blocking the image is unnecessary. Thus, no block artifacts occur.

Wavelet image compression systems vary widely in complexity and character. However, except for a few notable systems, quantization is performed while encoding, just like JPEG, allowing only one possible decoding of the image.

Zerotree-based systems [4][5][6] are embedded (progressive by pixel fidelity), therefore, the quantization is a function of how much of the compressed file is decoded. Although progressive with respect to pixel fidelity, zerotree-based systems do not allow efficient extraction of the image at different resolutions. The wavelet transform's natural "pyramid" of image resolutions are not exposed, or available, in the zerotree-based methods.

The new FlashPix file format [7] has a pyramidal structure that allows the extraction of different resolution images. This structure allows zooming and panning through the image. However, it is achieved through the use of redundant copies of the image. Further, the only compression currently allowed is the lower quality lossy JPEG with only one quality available per redundant image.

Compression with Reversible Embedded Wavelets (CREW) [8][9][10] is a wavelet compression system that allows the best quality compression and enables great flexibility in quantization of the image while *decoding*[1]. CREW technology enables

- progressive quality from high compression to lossless,
- pyramidal resolution from thumbnail to full size,
- fixed-rate and fixed-size encoding and decoding,
- region-of-interest (tiled) decoding,
- idempotency,
- color, gray-level, graphic, text, bi-level images,
- "parsable" codestream without decoding,
- and different coding styles for multiple components.

CREW's features enable many modern imaging systems to utilize image data in compressed form. For example, an image browser can pan and zoom through a compressed image using the progressive, pyramidal, and region-of-interest features.

In another example, using the fixed-size and idempotency or the lossless features, an image editor can call up the appropriate potions of an image, perform editing or filtering, and recompress those portions into the codestream. These operations can be performed at the appropriate resolution or postponed until that resolution is needed.

Finally, a WWW example is perhaps most illustrative of the need for this type of compression system. Using the pyramidal, progressive, and parsing features, a number of thumbnail images could be displayed on a web page. The

1. For more information CREW see http://www.crc.ricoh.com/CREW/.

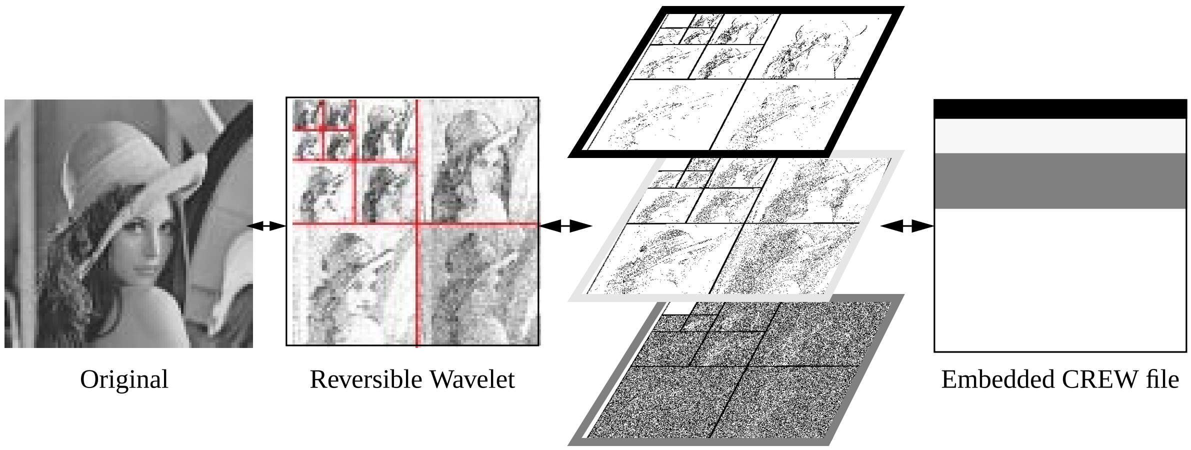

Figure 1 - CREW transform-style compression path

client can select the appropriate image and the server could parser the data needed for a screen resolution image. When the client wants to print, the remaining data needed for a print resolution image is parsed by the server and sent to the client.

After a brief introduction of the CREW technology, this paper describes how to extract (quantize at decode) these different image sizes, regions, and fidelities.

2 CREW System

To illustrate the basic workings of CREW, the compression of a grey-level (single component) image using only the wavelet transform compression style is described first. Figure 1 shows this image data path. The color, bi-level, and multiple tiles capabilities of CREW are mentioned briefly.

2.1 Grey-level transform-style compression path

The image is transformed by the “Two-Ten” reversible wavelet transform based on a decomposition of the Le Gall-Tabatabai polynomial offered in the literature [11][8][10]. This filter pair implementation has the special properties of reversibility (exact reconstruction with finite precision arithmetic) and no precision growth in the low pass coefficients.

The resulting coefficients are “aligned” into “importance levels.” That is, the coefficients are shifted with respect to each other according to some user defined criteria, such as best rate-distortion performance with respect to MSE.

This alignment is a precursor to quantization. A given alignment makes certain quantizations easier than others. The alignment does not necessarily preclude other quantizations, however. Figure 2 and Figure 3 show two possible alignments.

The importance levels are encoded as if they were bit-planes in a binary image, much like JBIG [12]. The highest importance level comes first and the lower ones follow in order. To encode the importance levels a context model followed by a binary entropy coder (Finite State Machine [13]) are used. The context model and binary coder can be reset at the beginning of any importance level, providing an “entry point” for random access.

In the simplest decode operation, the importance levels are decoded in order, from most important to least important. Of course, the decoder can stop at any point, such as when a certain bit-rate, quality, or resolution is obtained. Then the quantized coefficients are reconstructed and transformed into the image.

A description of variations on this decoding is offered below.

2.2 The CREW file system

CREW can handle multiple component images. Components can be interleaved with entry points or can be non-interleaved. If the components are classic RGB or CYM color planes, CREW offers a reversible color space conversion that is tuned to the Human Visual System for quantization.

For random access of regions of the image, the image may be tiled using a regular rectangular grid of user specified size. Thus, each component is divided into “tile-components.”

CREW allows for a binary-style in addition to the wavelet transform-style of coding. For tile-components that have graphic or text regions (such as overlays, alpha planes, and transparency planes), binary-style coding can

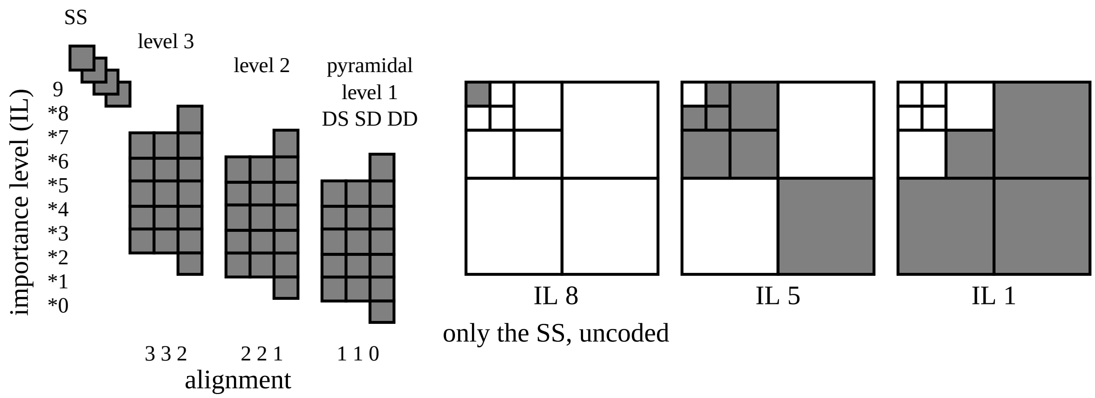

Figure 2 - Normal alignment and example importance levels

achieve higher quality than transform-style coding. Essentially, the transform is bypassed for the designated tile-components and the component bit-planes are encoded like importance levels.

3 Decoding with CREW

CREW is unique in that several types of reconstruction or quantization decisions can be made by the user or client at decode time. Below are some examples.

3.1 Alignment for progressive pixel fidelity

Consider a tile-component of an image with 4 bits per pixel. The mathematics of the filters lead to a maximum magnitude of Detail-Detail (DD) coefficients of 8 bits signed, and the Detail-Smooth (DS) and Smooth-Detail (SD) coefficients of 6 bits signed, while the single level of Smooth-Smooth (SS) coefficients are 4 bits unsigned. (The sign bit is handled separately and is not shown.)

The left side of Figure 2 shows the maximum magnitude of the coefficients in "normal" alignment. This special alignment has the property of being progressive by fidelity in a linear coded bit stream. That is, every successive decoded bit contributes monotonically to a better, full resolution, image with respect to SNR or MSE. Thus, the quantization can be controlled by simply truncating the coded bit stream. This truncation can be at a target bit-rate, transmission time, or quality level.

The right side of Figure 2 has three example importance levels, 8, 5, 1. The gray areas represent coefficients that have magnitude bits at that level. The white areas are deterministically known to contain no data and are not encoded. The uncoded SS coefficients are packed into the codestream first.

3.2 Alignment for pyramidal resolution

Figure 3 has the same tile-component coefficient data arranged in pyramidal alignment. This alignment has the property that simple truncation leads different resolutions.

3.3 Quantization regardless of alignment

Note that, in the pyramidal alignment, if there are entry points at importance level 20, 14, and 6 (denoted by the asterisk in Figure 3), it is possible to decode each pyramidal level independently. It is, therefore, possible to achieve "normal" quantization at any resolution. For example, if IL21, IL20 through IL16, IL13 through IL10, and IL6 through IL3 are decoded, this is equivalent to decoding IL9 though IL3 in the normal alignment.

It is also possible to decode pyramidally when aligned in normal mode. Referring to Figure 2, if there is an entry point at every importance level, then each importance level can be decoded until the target resolution is reached. Then

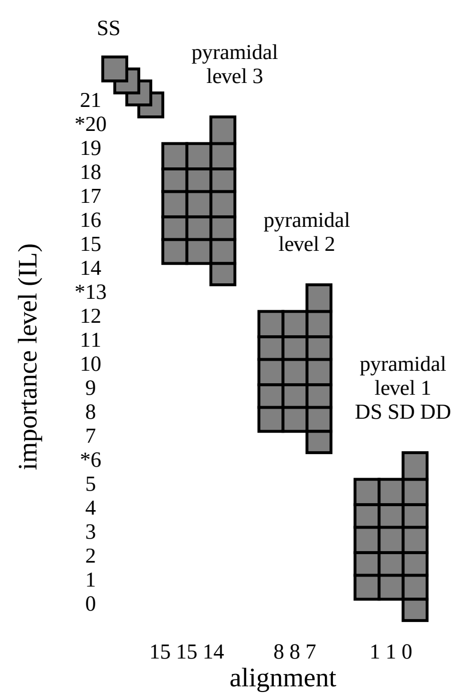

Figure 3 - Pyramidal alignment

the decoder skips to the next importance level entry point and starts again.

3.4 "Custom" alignments

The CREW syntax allows any custom alignment to be used. These alignments can be used for special quantization based on Human Visual System characteristics of the image and the target device or other norms like L_1 and L_∞. They could also be used for special functions such as edge detection.

3.5 Region-of-interest, zooming and panning

Regions of interest can be decoded independently with the use of tiling. This allows partial decoding and panning through the image. Pyramidal decomposition allows zooming into regions of the image. These two features can be combined to allow panning and zooming with a minimum of coded data transmitted.

3.6 Parsing a compressed image

Parsing refers to extracting the right data from a CREW file to reconstruct a quantized image *without* decoding. Imagine using this feature for the WWW client-server protocol. A high resolution lossless or near-lossless image is stored on the server. Perhaps a thumbnail image is extracted from the server, sent in a minimal bit stream, and decoded at the client. Then, according to the instructions from the user or client, a pyramidal segment of the image is sent to fill the window area of a particular screen. Finally, the full resolution image could be sent for printing. All of this is done without decoding at the server or sending redundant data.

Nearly all of the quantization that can be performed by a decoder can also be performed by a parser.

4 Future plans and research

CREW will be submitted to the JPEG 2000, the next generation still image compression standard effort. In fact, the CREW technology was key in prompting this standardization effort [14].

The flexibility of the CREW system suggests a number of areas of research to optimize its use for specific applications. For example medical imaging, facsimile, digital cameras, scanners, printers, copiers, CD-ROM archival systems, image browsing, remote sensing, and the World Wide Web all have different image types and system requirements. All can benefit from the flexibility of CREW.

5 Conclusions

Accessibility and flexibility of an image compression system is of more importance as the size and pixel depth of images in modern digital imaging systems increases. Functions such as access to different resolutions and pixel fidelities, zooming and panning, region of interest, and fixed-size can be achieved with CREW without sacrificing compression performance or image quality.

Bibliography

[1] O. Rioul and M. Vetterli, "Wavelets and Signal Processing," IEEE SP Magazine, pp. 14-38, October 1991.

[2] W. B. Pennebaker and J. L. Mitchell, *JPEG still image data compression standard,* Van Nostrand Reinhold, New York, NY, 1992.

[3] JPEG, *Digital Compression and Coding of Continuous-tone Still Images, Part 1: Requirements and guidelines*, ISO/IEC DIS 10918-1, 1992.

[4] A. Lewis and G. Knowles, "Image compression using the 2-D wavelet transform," *IEEE Trans. Image Proc.*, vol. 1, pp. 244–250, April 1992.

[5] J. Shapiro, "An embedded hierarchical image coder using zerotrees of wavelet coefficients," *Proc. IEEE Data Compression Conference*, Snowbird, UT, pp. 214–223, March 1993.

[6] A. Said and W. Pearlman, "Reversible image compression via multiresolution representation and predictive coding," *SPIE Visual Communications and Image Processing*, vol. 2094, pp. 664-674, November 1993.

[7] FlashPix format, www.kodak.com/daiHome/flashPix/flashPix.shtml.

[8] A. Zandi, J. Allen, E. Schwartz, and M. Boliek, "CREW: Compression with reversible embedded wavelets," *IEEE Data Compression Conference*, Snowbird, UT, pp. 212–221, March 1995.

[9] E. L. Schwartz, A. Zandi, M. Boliek, "*Implementation of Compression with Reversible Embedded Wavelets*," *SPIE*, vol. 2564, San Diego, CA, 11 July 1995.

[10] M. J. Gormish, E. L. Schwartz, A. Keith, M. Boliek, and A. Zandi, "Lossless and near-lossless compression for high quality images," SPIE, San Jose, CA, February 1997.

[11] D. Le Gall and A. Tabatabai, "Sub-band coding of digital images using symmetric short kernel filters and arithmetic coding techniques," *IEEE International Conference on Acoustics, Speech and Signal Processing*, New York, NY, pp. 761–765, 1988.

[12] JBIG, *Information technology - Coded representation of picture and audio information - Progressive bi-level image compression*, ISO/IEC IS 11544, 1993.

[13] M. J. Gormish and J. Allen, "Finite state machine binary entropy coding," *Proc. Data Compression Conference*, Snowbird, UT, p. 449, March 1993.

[14] A. Zandi, M. Boliek, E.L. Schwartz, M.J. Gormish, J.D. Allen, "*CREW Lossless/Lossy Image Compression Contribution to ISO/IEC JTC 1.29.12*," proposal submitted to the ISO/IEC JTC1/SC29/WG1 (JPEG/JBIG), CRC-TR-9524, 30 June 1995.