AN INTRODUCTION TO JPEG 2000 AND WATERMARKING

Richard Clark

Background

Although the name JPEG now means something to most people who have some awareness of Internet technology, it has come to refer to small part of one International Standard (IS 10918-1) and a file wrapper that encompasses that. In reality, the Joint Photographic Experts Group refers to an evolving community of experts that meet three times a year all over the world, and who are charged with delivering the standards that allow digital images to be compressed to a manageable size. Applications range from its use in Internet technology to digital cameras, pre-press, medical imaging, security applications, cultural heritage conservation and many, many more.

The group has varied in size over the years from as few as 8 or so (when the boring work of attending to conformance was being dealt with) to well over 200 at present. Currently, at least 100 of these attend each meeting and participate actively in its work electronically at other times. All are nominated through their national standards body. Whilst national requirements vary from a fairly minimal attendance at local meetings to a requirement for active contribution via a one week pre-meeting discussion, those who become JPEG experts never fail to be surprised at the depth of and opportunity for discussion.

The JPEG work started in the late 70's with a requirement emanating from the various national videotex systems. The pre-cursor to Internet technology, videotex was always felt to need 'real' images to be a success. Its applications were those of the Internet today. With access speeds ranging from 1200 to 9600 bit/s for the lucky few, for successful implementation the compression of images was not just a nicety but an essential feature. As the oldest remaining JPEG member, the author was lucky enough to be in at the start, when the more dominant skills were those of the character coding world, and the need was for simplicity and elegance of the compression algorithms, with little attention to more sophisticated usage.

As more members became involved, and the computer revolution took hold, it became obvious that the narrow field of adding small images to terminal technology was insufficient, and an array of new features were gradually added to the first JPEG standard. Many of these are still being 'discovered' as we move into a new generation of coding systems - progressive image build-up is quite a recent addition, and the intention of point transforms, arithmetic coding and hierarchical coding are a grounding for today's work on JPEG 2000.

It would not be fair to leave the early JPEG work without mention of three groupings that probably delivered more to the work of the JPEG community than any other. These are

- the small group of European researchers in the COST 211 and other projects that provided a solid base for the use of the discrete cosine transform (DCT) in image compression,
- the many experts from IBM's Thomas J Watson Research Labs (and especially Joan Mitchell and Bill Pennebaker as editors for the original ISO standard text, and the definitive book on JPEG)
- and lastly, and by no means least, Tom Lane and the Independent JPEG Group (IJG) who, without ever attending a JPEG meeting, have probably done more than anyone to make the JPEG standards

Managing Director - Elysium Ltd (http://www.elysium.ltd.uk)

Chair BSI/IST37 (MPEG/JPEG/MHEG/JBIG)

Webmaster and Editor, JPEG (mailto:richard@elysium.ltd.uk)

The paper also reflects contributions made by fellow contributors to the CEC funded SPEAR, Eurostill and Migrator2000 projects, whose collaboration is gratefully acknowledged.

(and even the Internet) a reality. Their license fee free code (long before Open Source was a well defined concept) is of the highest quality and is still the de facto standard for implementers to follow.

These are personal acknowledgements - many others could have been provided, including the hard work of all the JPEG Convenors, but it is time to move on.

JPEG 2000 - the new pretender?

Whilst computing power, memory allocation and real time decoding of a 64 kbit/s stream were major factors in deciding the technology for the original JPEG standard, it was clear that mistakes were being made, and evolution was not taking place. In part it is probable that the widening of the application area for JPEG technology had led to confusion, both amongst implementers and technologists. The original JPEG standard had in excess of 40 options, many of which were mutually exclusive, and some of which required costly licensing to avoid IPR constraints. Its potential for use across a much wider range of applications meant that the requirements were constantly changing, and the standard developed as a shopping list of compression components rather than as an integrated whole.

JPEG tried to counter these problems and criticism firstly by extending the standard, providing for so-called pyramidal coding, improving compression options via variable quantisation and other techniques, and adding SPIFF, a file format (regrettably omitted in the original standard). It then tried to further improve options and meet the needs of some users by offering a new lossless standard, JPEG-LS. Although these are just as valid as improvements to the current state of the art, at least in practical terms, the marketplace has reacted mostly with stoical indifference. The sub-baseline implementations of the original JPEG standard are out there, and there is a huge legacy of existing imagery and code to support them. The IJG has shown little interest in augmenting the existing de facto JPEG code for compression and decompression as its users tend to adopt a minimal subset of its functionality in any event. There was little option if improvement was to take place other than to take a radical step forward.

Ricoh (with CREW) and others had been unsuccessful in getting their code accepted as suitable for JPEG-LS. There was considerable discussion within JPEG as to how some of the obviously useful features which were proposed might be introduced effectively, and it was agreed that the time had come to try and produce a next generation standard, capable of dealing with digital imagery well into the next millennium. And so the concept of JPEG 2000 was born.

Work on wavelet compression had in the meantime gathered pace. Although much of the foundation for this is over 50 years old, it really began to develop in the 80's - by which time JPEG standardisation of the DCT was already well under way. Mallat's work on pyramid algorithms and the decomposition of images led to Daubechies' derivation of a set of orthonormal basis functions which are fundamental to the practical use of wavelet technology, and indeed are used in JPEG 2000. As with JPEG DCT though, the theoretical needed practical demonstration and code, and the FBI in the US provided this in the early 90's. Their co-operation on a fingerprinting standard with NIST led to considerable interest in the use of wavelet technology as a practical alternative to DCT - as expected this became a major feature in the early proposals for JPEG 2000.

JPEG 2000 - what, where and when?

So what is JPEG 2000 intended to provide? Key in this is an architecture for handling all aspects of digital still imaging rather than a toolkit. This even extends into the bite-size chunks of moving images and audio that the next generation of digital cameras will offer. Learning both from its mistakes and its success, JPEG intend (at least as far as is possible in today's litigious and IPR conscious environment) to preserve some basic success factors. These include:

- the model which meant that at least a basic set of functionality could be offered, without requiring developers to pay a software licensing or royalty fee.
- as the work of the IJG has been so instrumental in the success of the standard, JPEG intend to offer high quality C / Java source code to implementers
- visually lossless image quality should be achievable at a level which is at least as good as current JPEG DCT images

- delivery should be possible, and mechanisms provided to allow both clients and originators to define how images are received and displayed, including transmission up to a completely lossless stage of compression
- better quality imaging at high compression ratios, which when coupled with progressive spatial or accuracy mean that more meaningful parts of an image will display earlier in bandwidth restricted environments
- avoiding some of the more unpleasant JPEG DCT artefacts - the ringing near sharp edges, the clear tile boundaries, and the harsh colour quantisation are all undesirable at high compression ratios
- and lastly, but by no means least, the capability to offer the features needed in today's e-commerce environment - protection of the image by encryption, watermarking, licensing and registration and so-on.

All of this is handled, and is gradually being publicised by JPEG's own web site at *http://www.jpeg.org*. The site is shared with the JBIG group who work on bi-level (and reduced grey scale) images, and this collaboration is resulting in a gradual merging of the standards needed for colour facsimile and cheque processing, and those of the photographic image community. JBIG are producing an enhancement to their original standard (known as JBIG-2), and not only are they following the JPEG model by intending that this should be implementable free of license fees, their arithmetic entropy coder is forming a key component in the JPEG 2000 deliverable.

The work on JPEG 2000 started in 1996 - already 12 one week meetings have been held at a wide variety of venues to discuss the options, and for the last two years these have been attended by well over 100 representatives of all the leading players in this field. The work culminated in December 1999 with the release of the first Committee Draft of Part 1 of the standard. The timetable for production of the standard means that initial text for a final CD was made available at the end of March 2000 - this is downloadable from the links at the JPEG web site at *http://www.jpeg.org/JPEG2000.htm.* The hope is currently that the documents will enter the ISO standards track process in July. At this time ISO rules will prevent further electronic publication without payment for the document.

The final version of the standard will be issued in at least five parts - hopefully relatively close together as drafts exist for much of the content already. The final document will be known as IS 15444 and will be identical to the ITU-T Recommendation T.800. The parts planned at present are:

- Part 1, JPEG 2000 Image Coding System
- Part 2, Extensions
- Part 3, Motion JPEG 2000
- Part 4, Conformance
- Part 5, Reference software

So how does JPEG 2000 stack up?

JPEG 2000 follows a similar progression to any transform technique for image compression. For those who are interested, a number of papers cited from *http://www.jpeg.org/JPEG2000.htm* provide further detail, and only a brief outline is tabled below.

Initially, the components of the image are passed through an optional level shifting operation and a decorrelation transform suitable for RGB images (either a simple reversible version, or a more accurate irreversible transform for lossy compression). Following this a wavelet transform is performed - three types of decomposition are supported, but by default a Mallat decomposition is used, and then typically a Daubechies 9/7 wavelet transform is used for irreversible coding, or a simple symmetric 5/3 filter if lossless transmission is the goal. The output is then organised into blocks using the 'EBCOT' techniques proposed by David Taubmann, effectively creating a final bit-stream that consists of a number of independent layers. These can be given different priorities during transmission to achieve differing image build-up features, and truncated as appropriate during lossy compression. The output coefficients are quantised, and here another feature of JPEG 2000 is introduced, as this process can be controlled to provide a given compression rate and hence

output file size. Finally, a low complexity arithmetic coding stage is added to entropy code the delivered bit-stream. Decoding is simply a reversal of these stages.

Estimates vary of the increase in complexity introduced, but in general it is felt that JPEG 2000 decoders will be approximately an order of magnitude more complex than their DCT forebears. The range of additional features offered is however significant. These include:

- imaging transmission going from lossy through to lossless
- the ability to define Regions of Interest which can be coded at higher resolution, or losslessly to preserve image features - important for example in medical imaging
- the use of 'resync' markers to allow JPEG 2000 to be effective (with a recovery mechanism) on data channels with high error rates - applications include the new mobile telephony channels
- 'fair' quality image reproduction at rates down to 0.1 bits per pixel and below
- the ability to customise encoders and decoders to suit low memory or fixed size applications
- random codestream access and processing
- high quality and fidelity colour image processing, with wider bit depths and larger image sizes
- use of alpha channels and other features to meet future graphic arts and Internet needs
- features to provide for image security and content metadata inclusion

A word of caution here! Claims are sometimes seen of compression ratios of 200 to 1, or vast improvement over existing JPEG DCT technology. To date, the author has found little evidence that these are true except in particularly contrived cases. Often, the comparisons are made at high compression rates, as the differences at lower ones are very difficult to gauge. JPEG DCT, in its common baseline form, suffers in these comparisons, as the resolution is then too high for effective compression. Simply sub-sampling the image by two in both directions before a lower rate of compression (achieving the same target compressed image size) and then intelligently up-sampling the down-sampled image after decompression yields far better results - often comparable to JPEG 2000. It is an interesting fact that, even without this obvious improvement, the original JPEG algorithm did not finish last in the large scale testing of candidate algorithms carried out for JPEG 2000!

Protection of IPR in JPEG DCT and JPEG 2000 images

Many techniques have been proposed and are used to allow metadata to be included in digital images. Sometimes this metadata is provided to secure parts of the image, and to establish its credentials. In principle, there are three major reasons for adding security information to the image:

- to define the ownership of the image, and protect its usage
- to prevent access to all of the image data, for instance to prevent its effective use at higher resolutions
- as an indication that the image has not been tampered with

In looking at image protection, we are concerned with both 'labelling', in which conventional metadata is added to the image, and watermarking, where artefacts are added to the actual image data. Watermarking can be further subdivided into visible (where the intention is to degrade the image and make presence of the watermark obvious to a viewer), and invisible (sometimes requiring the data to be extracted by reference to the genuine original).

JPEG DCT images have traditionally been labelled both in the file format and sometimes in the codestream in order to address ownership. Unintentionally (or some might feel intentionally…), images have also been created with parts of the information required for image decoding removed, such as the quantisation matrix definitions. This is permissible in a closed application where encoder and decoder can assume a set of defaults, and indeed can save 700 bytes or so in the compressed image size. The effect is to allow image viewing, but only of an inaccurate facsimile. Normally however, the file header contains information about many aspects pertaining to the image - its time and date of creation, the capturing device or application, its size, colour space and other technical details, and where relevant, its ownership.

In the past, applications, such as those developed through the European funded MENHIR program, have concentrated on the idea of a 'licence plate' which links a unique registration identity to a record in a database provided by a Registration Authority. The SPIFF file format defined in IS 10918-3 and the registration mechanism allowed for in IS 10918-4 are designed to accommodate this type of usage, and the JPEG Registration Utilities Authority web site (*http://jura.jpeg.org*) was established to prototype such operation.

What is the purpose of such labelling? Typical examples have been provided in a contribution to JPEG (1677) by Claude Rollin of the French organisation SACD. They include:

- For the author
  - protection of his 'moral rights' to be acknowledged as the author and not to have it defaced
  - preventing others from using it without his permission
  - allowing collection of any revenues accruing from its licensed use
- For the 'image broker'
  - to show images with identification and protection (promotion & dissemination)
  - to sell images (content trading)
  - to control the use (piracy)
  - to collect the economic rights (exploitation fees)

These are inherently the needs of the artistic community as they relate to digital imaging. Other interest groups add to these needs - for example the medical community has requirements to preserve anonymity, prevent misuse and recirculation, except to agreed viewers, link to data records and other imagery from many modalities, and to indicate when images have been modified or tampered with.

In general, the use of metadata within the file format can meet many of these needs, but often an additional layer is required to provide a fuller level of protection. This is particularly true of imagery that is destined for the World Wide Web (WWW).

Encryption can be used, either of all or parts of the image, and here the well defined and many optional structures within the JPEG 2000 specification, it is felt, can be put to good effect. Images that can be readily viewed in a thumbnail format, or perhaps partially magnified might be encrypted to avoid misuse. To some extent, there is legal protection as well, as under the WIPO treaty on copyright (to which most major countries are signatories) there is scope for legal action if parts of a document relating to its copyright protection are removed. A further option has been discussed within the JPEG 2000 initiative - that of restricting implementers access to the separate technical IPR necessary for creation of their products if they fail to provide a mechanism to keep IPR information embedded in image metadata. This is somewhat similar to respecting the font embedding licensing information added to TrueType fonts. It essentially attempts to prevent major suppliers and organisations from abusing a well-defined and specified mechanism to preserve IPR claims.

For some time, there has been discussion about how and whether watermarking should form part of the standardisation process for JPEG 2000. Despite general agreement that it would be useful to at least provide the tools and the hooks to allow some generic watermarking schemes to be included, unfortunately the initial drafts of the JPEG 2000 standard have no mention at all of watermarking as yet. There are many blanket type patents applying in the arena, and the need for a standard which can be implemented without payment of license and royalty fees, as with JPEG DCT is an important driver.

One of the major contributors to JPEG 2000 work, Professor Touradj Ebrahimi of EPFL, Switzerland is chairman of the JPEG 2000 requirements group, and his work and that of his staff and pupils is well know to all specialists in watermarking. Diego Santa Cruz and David Vanrell in particular have looked specifically at the effect of JPEG 2000 compression on watermarking techniques. Their conclusion is that the use of a technique which modifies coefficients in the wavelet domain not only out performs conventional watermarking in the spatial domain, but is far easier to integrate into the JPEG 2000 system.

Does watermarking have a future in JPEG 2000?

This paper has looked at how the JPEG 2000 work has developed, starting with the original groundwork on the first JPEG standard. When JPEG 2000 started to evolve, it was very noticeable that not only did JPEG's technical membership increase, so did participation by many organisations with a close interest in the protection of IPR. In some areas, such as medical imaging, there is very clearly concern that any technique that can introduce artefacts, especially those that might not be expected as typical of a specified lossy compression process, is not acceptable. On the other hand, a watermarking scheme, which deliberately distorts the image to make it unusable for diagnosis unless the correct user's identity is provided, might well be attractive. It could help prevent the well-reported tragedies where imaging results have led to medical error. In other markets, the very features of a given watermarking scheme are likely to offer designers an opportunity to integrate watermarking technology into an overall solution which meets the objectives of such important groupings as Internet based image brokers.

The likelihood is therefore that watermarking technology will be used in conjunction with JPEG 2000, but at present it seems unlikely that this will come about directly through the standardisation process.

It is clear than many of the advanced features inherent in JPEG 2000 will only evolve through active participation by users of the standard. As a result of widespread development of JPEG related tools through the EEC funded MENHIR and OCTALIS projects, a new project called MIGRATOR 2000 has been agreed under the 5th Framework program to try and migrate those tools towards the new environment offered by JPEG 2000. As UK partners in this collaboration we look forward to working with other European partners with extensive experience derived from MENHIR, OCTALIS and other work to examine how watermarking might be best applied within the confines of JPEG 2000. Initial results from this work are hoped to be available towards the end of this year - hopefully in time to meet the requirements of the standard's user community. Current partners in the project (as well as Elysium Ltd) are NetImage (France), ADETTI (Portugal), UCL (Belgium), Fratelli Alinari (Italy), Thomson CSF (France) and SACD (France). All of these organisations are well acquainted with image security and watermarking issues, and hope as one of the objectives to deliver tools into the public domain which will have general applicability to extending the uses and options for image authentication and protection.