PROGRESSIVE CODING IN JPEG2000 – IMPROVING CONTENT RECOGNITION PERFORMANCE USING ROIs AND IMPORTANCE MAPS

Anthony Nguyen[†]*, Vinod Chandran*[†]*, Sridha Sridharan*[†] *and Robert Prandolini*[‡]

[†] Research Concentration in Speech, Audio and Video Technology
Queensland University of Technology, GPO Box 2434, Brisbane QLD 4001, Australia
an.nguyen@qut.edu.au, v.chandran@qut.edu.au, s.sridharan@qut.edu.au

[‡] DSTO C3 Research Centre
Fern Hill Park, Department of Defence, Canberra ACT 2600, Australia
robert.prandolini@dsto.defence.gov.au

ABSTRACT

This paper presents a system, which permits arbitrary regions of importance and corresponding degrees of importance to be assigned in an image within the JPEG2000 image coder. The image can be reconstructed progressively in layers prioritised by importance values rather than the objective criterion of overall PSNR optimisation. Importance progressive coding is very desirable in surveillance applications. The system, IMP-J2K, aims to improve the interpretability or content recognition performance of compressed images, by selecting and distributing the bit allocation to contents of importance in an image and embedding these earlier in the code-stream. IMP-J2K makes use of JPEG2000's data packing sub-system, which can selectively encode data in layers.

1. INTRODUCTION

Importance progressive coding for interpretability is a very desirable feature in a number of applications, especially surveillance where particular regions of an image are of higher importance than others. The term interpretability refers to the content recognition performance by trained image analysts. One application within surveillance may be for the support of image compression in tactical domain military applications. Operations may have users to take an image, mark on the image a number of Regions of Interest (ROI) and be able to adjust the fidelity of each ROI and the background region. The purpose of this activity is to maximise the compression of the image whilst maintaining hi-fidelity for the relevant ROIs. A recipient of this compressed file would thus receive the image at the quality selected by the sender. This is very advantageous for trained analysts, who are required to utilise the imagery to support decision processes in strategic, operational and tactical tasks.

Importance progressive coding aims to optimise the interpretability of an image for a given bit rate by selecting and distributing the bit allocation to contents of importance. The coding strategy attempts to improve the image understanding for users of low quality images by optimising the amount of useful visual information encoded for a given bit-rate. The JPEG2000 standard is a state-of-the-art progressive image coding system that has capabilities for importance progressive coding.

The paper provides an overview of the JPEG2000 progressive coder in Section 2. Section 3 describes the proposed importance progressive coder and discusses a system for the application briefed earlier. Results are presented in Section 4, followed by discussion and conclusion in Section 5 and 6.

2. JPEG2000

JPEG2000 is a new international standard for image compression, coding and decompression [1]. The coding scheme uses a wavelet-based technology in its compression scheme (see Figure 1(a)). The coder is essentially a bit-plane coder, based on the Embedded Block Coding with Optimised Truncation (EBCOT) coding engine originally proposed by Taubman [2]. This section describes pertinent mechanisms of JPEG2000 that are used to develop IMP-J2K. Further details on JPEG2000 can be found in [1].

The wavelet decomposition produces sub-bands of transform coefficients at a number of scales. JPEG2000 divides each sub-band into blocks of coefficients, called code-blocks, which are independently bit-plane coded. Code-blocks are coded a bit-plane at a time starting from the most significant bit-plane to the least significant bit-plane. The individual bit-planes are coded in only one of three sub-bit-plane coding passes. The decision as to which pass a given bit is coded in is made based on the significance of that bit's location and the significance of neighbouring locations.

Once all the code-blocks have been compressed, a post-processing operation passes over all the compressed code-blocks to generate the final bit-stream. The bit-stream is organised as a succession of layers. Each layer contains the additional contributions from each code-block as shown in Figure 2. The coding passes can be put into any coding layer such that each layer will incrementally improve the overall image quality for the entire image at full resolution. There is no restriction on the number of sub-bit-plane coding passes contributed by each code-block to a given layer. The operation is based upon the rate-distortion curve and determines the extent to which each code-block's bit-stream should be truncated in order to achieve a particular target bit-rate. The first, lowest quality layer is formed from the optimally truncated code-block bit-streams. Then each subsequent layer is formed by optimally truncating the code-block bit-streams to achieve higher target bit-rates.

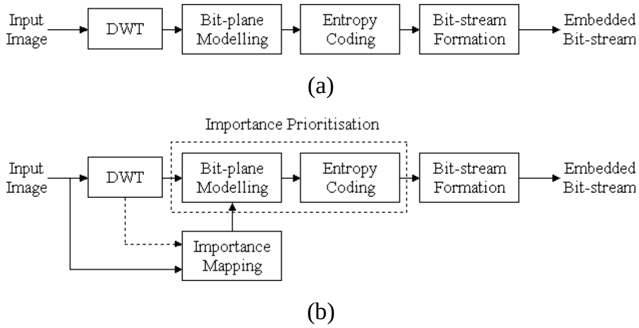

Figure 1. (a) JPEG2000 and (b) IMP-J2K encoder block diagram

The default implementation of the JPEG2000 encoder computes the order in which code-block sub-bit-planes should appear in the layered hierarchy of the overall bit-stream based upon rate-distortion criteria. The implementation places in each layer the sub-bit-planes that correspond to the steepest rate-distortion slope as estimated by the encoder. The implementation aims to minimise the MSE at each point in the embedded bit-stream and improve the MSE performance.

The flexibility in the formation of code-block contributions to layers is left to the implementer of the encoder. Thus implementations may modify the distortion estimates of each sub-bit-plane to make it consistent with *a priori* knowledge of region importance. The flexible ordering of the different code-block contributions allows for different types of scalability such as importance scalability. A more complete description of the layered bit-stream formation can be found in [2].

3. IMP-J2K

The research and development in this paper continues previous efforts on importance progressive image coding [3-5], whose goal is to develop a JPEG2000 based importance progressive image coder. The paper presents an Importance Progressive JPEG2000 Image Coder (IMP-J2K), which allows ROI coding within the JPEG2000 framework. The block diagram of the IMP-J2K encoder is shown in Figure 1(b). IMP-J2K makes use of the data-packing-sub-system from JPEG2000 and modifies the embedded coding order by selectively packing data into the progressive layers of JPEG2000 with priority set according to a

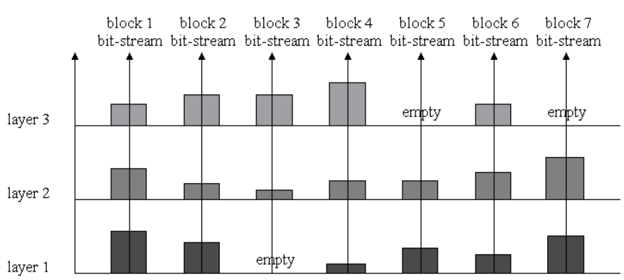

Figure 2. Code-block contributions to bit-stream layer. Contributions may vary from layer to layer. Importance mapped scores can influence this and provide importance scalability.

scale-space importance map. Multiple arbitrary ROIs can be defined at different spatial locations, sub-band orientation and scale using arbitrary importance scores. No importance map information or any importance scores need to be signalled to the decoder. The proposed coder generates legal JPEG2000 code-streams and standard decoders may decode IMP-J2K code-streams. An advantageous feature associated with IMP-J2K is that ROIs need not be coded with a single importance score for all ROIs. A gradual finer embedding of the ROIs with the background may be also achieved. The proposed system provides an effective framework for further importance progressive coding research.

3.1. Importance prioritisation

IMP-J2K takes advantage of the fact that an arbitrary ordering of code-block sub-bit-planes to a quality layer is possible. Different sub-bit-planes from code-blocks can be ordered in an application dependent manner, thus providing importance scalability in its bit-stream formation.

The distortion cost function used to allocate code-block contributions to quality layers is modified according to the region characteristics provided by a scale-space importance map. Importance scores are input to the post-processing rate-distortion optimisation routine, which effectively controls the relative significance of including a different number of sub-bit-planes from the embedded bit-stream of each code-block.

A simple implementation for changing the embedded coding order for importance prioritisation is to change the distortion metric based on the importance scores for each code-block during the bit-stream formation. The distortion contributed by each sub-bit-plane pass for a particular code-block is weighted by the square of the code-block's importance score. In this method of importance weighting, regions with higher importance would effectively contribute a higher distortion, while regions with lesser importance would contribute less distortion for the corresponding regions. Since the bit-stream formation is driven by a post-compression rate-distortion optimisation, importance scores effectively control the embedded coding order of the code block sub-bit-planes while coding is in process. That is all code-blocks whose samples contribute in any way to the reconstruction of the ROI will have their distortion metrics scaled for the purpose of rate allocation.

This implementation has little effect on the lossless file size (bit-rate), but can have dramatic impact on the interpretability for partial decoding at lower rates. It is this flexibility that makes importance progressive coding promising.

3.2. Importance mapping

An importance map is an image map describing the relative weights of importance of regions in an image. An extension of this concept is to define importance maps at a number of scales and sub-band orientations similar in structure to the wavelet decomposition produced by the JPEG2000 coding engine. A scale-space pyramid for importance maps may be systematically generated in several ways as in [3-5], but for applications requiring the hand marking of ROIs such as that discussed earlier, a simple and effective method is proposed:

1. Hand mark a number of arbitrary polygonal shaped ROIs from an image.
2. Assign importance scores to regions according to some *a priori* knowledge of region importance. All other regions (i.e. background) are assigned a lesser score.
3. Determine the block approximation of the ROIs to determine which blocks have at least a threshold percentage of their pixels lying within the polygonal ROI. The block approximation identifies which code-blocks belong to which ROI.
4. Each block approximation is assigned the corresponding importance score for the region it belongs to. This forms the importance map at the highest resolution.
5. Generate the scale-space importance map pyramid from the importance map at the highest resolution. This can be obtained by taking the maximum of 2-by-2 block neighbourhoods. As such every successive lower resolution will consist of half the number of blocks horizontally and vertically, which corresponds to the code-blocks from JPEG2000 at that same scale. Other sampling techniques may be used but this approach was found to be effective.
6. The scale-space importance map pyramid is then input into IMP-J2K for importance progressive coding.

The proposed schema uses the same importance map for all directional sub-bands for the corresponding resolution. But the user may identify ROIs for different scales and sub-band orientations if desired. It is important to note that the size of the blocks of importance scores, say 16-by-16 or 32-by-32, are constant and are defined at every scale and sub-band to correspond to the relevant code-block entity in JPEG2000.

4. RESULTS

The following results for the IMP-J2K coding system described above are implemented using the Kakadu software package version 3.1, of JPEG2000 [6]. A 1024-by-1024 partition from a much larger "factory" aerial image sourced from DSTO (see Figure 3(a)) was used to generate an IMP-J2K code-stream, using the ROI coding scheme described above. A 5 level wavelet transform using the default Daubechies (9,7) filter was performed on the original image. Code-block sizes of 16-by-16 samples and 30 quality layers were used to generate the code-stream. A code-stream output from the default JPEG2000 using the same constrained parameters as the IMP-J2K was also generated to compare the two results.

Three arbitrary polygonal shaped ROIs covering approximately 16% of the image area were hand marked and were arbitrary assigned importance scores according to *a priori* knowledge of the region importance for the image as shown in Figure 3(a). All other regions (i.e. the background regions) are assigned an importance score of 1. The block approximation of the ROIs was then applied to determine the code blocks, which represents the polygonal ROIs. A threshold of 0.2 was used for all ROIs. The scale-space importance map pyramid is then derived from the importance map at the highest resolution. The scale-space importance map pyramid is then input into IMP-J2K for importance progressive coding.

The IMP-J2K and JPEG2000 code-streams were decoded at layered bit-rates, which are logarithmically spaced and PSNR values were calculated for each one of these, for each ROI and background as shown in Figure 3(c) and (d). In addition a plot of the distortion for the ROIs and background versus quality layer was also generated to illustrate the operation of the data packing sub-system for each successive layer for the three ROIs (Figure 3(b)). The distortion, D, at quality layer, q, for region, p, is expressed as a percentage of the total image distortion given by:

$$D_{q_p} = \frac{\sum\left(\tilde{x}_p(n) - x_p(n)\right)^2}{\sum x^2(n)}$$

where $\tilde{x}(n)$ and $x(n)$ are the reconstructed and original pixel values respectively. The results are a demonstration of the capabilities of IMP-J2K. It provides a framework for an open area of research to be done in this field.

5. DISCUSSION

By allowing the background to be coded with lower fidelity than ROIs, significant gains can be achieved by transmitting a better interpretable image for the same number of layers or equivalently at the same cost in storage or transmission time. As can be seen from Figure 3(c), IMP-J2K encodes ROIs in order of importance. Regions of higher importance are, in general, encoded with higher PSNR for a given bit-rate than regions of lesser importance. Note the bit-rate delay in encoding regions of lesser importance with respect to the region of highest importance. There is however a trade-off as to how the ROIs and background are encoded. That is, if a region has been heavily emphasised then all other regions of importance including the background are correspondingly de-emphasised. However, Figure 3(d) illustrates that no such emphasis on the order of importance is present, indicating that ROIs and background are treated equally important in the default JPEG2000 schema.

Figure 3(b) shows the ROI distortion for each of the quality layers encoded. The figure illustrates how region distortions are gradually improved as a function of the quality layer. ROIs can be seen to improve with increasing quality layer, with the most important region reconstructing near perfect first then the next, in order of importance as defined by the user-defined importance map. The background however in this case, at low quality layers (i.e. layers 1 to 5) is significantly better than regions 2 and 3. This is a consequence of code-blocks from coarser scales being assigned a corresponding high importance due to the fact that a region of high importance was located within the respective block. As a result, if two ROIs begin to overlap at coarser scales then the corresponding code-block at that scale would emphasise both ROIs with the higher of the two importance scores. This is not a drawback for ROI coding but may influence the rate of reconstruction of the ROIs.

The IMP-J2K has been designed for full access on a code-block-by-code-block basis and gives the encoder full control over the assignment of ROIs. The method even allows ROIs to be defined at different scales and sub-band orientations.

6. CONCLUSIONS

ROI coding for interpretability is a very desirable feature in a number of applications, especially surveillance. The paper has

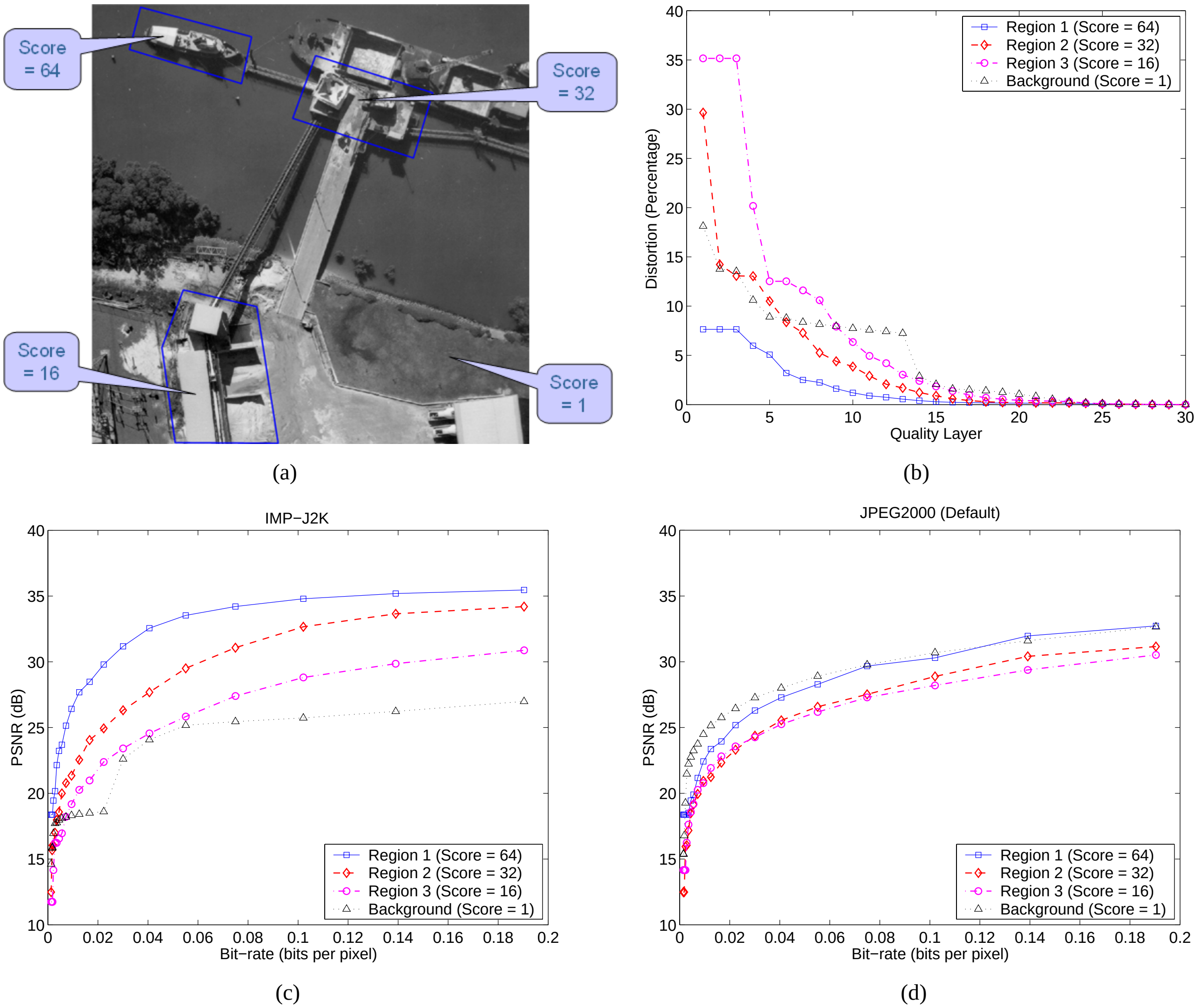

Figure 3. (a) Selected ROIs from a 1024-by-1024-pixel partition of "factory" image. Also indicated are user-defined importance scores for the ROIs and the background [64,32,16,1]. (b) Region distortion versus quality layers for IMP-J2K. (c)-(d) PSNR plots for ROIs and background for IMP-J2K and JPEG2000 for the first 20 layers.

proposed an importance progressive coding scheme, IMP-J2K, with application to tactical domain military image compression. Multiple arbitrary ROIs can be defined at different spatial locations, sub-band orientation and scale using user-defined importance scores. The regions are prioritised in order of importance by the proposed system, while maintaining a high level of compression. IMP-J2K is a highly flexible system, which can provide an effective framework for further importance progressive coding research.

7. REFERENCES

[1] C. Christopoulos, A. Skodras, and T. Ebrahimi, "The JPEG2000 Still Image Coding System: An Overview," *IEEE Trans. on CE*, vol. 46, pp. 1103-1127, 2000.

[2] D. Taubman, "High performance scalable image compression with EBCOT," *IEEE Trans. on IP*, vol. 9, pp. 1158-1170, 2000.

[3] R. Prandolini, "Coding of surveillance imagery for interpretability using local dimension estimates," VCIP, pp. 516-526, 2000.

[4] A. Nguyen, V. Chandran, S. Sridharan, and R. Prandolini, "Importance coding of still imagery based on importance maps of visually interpretable regions," ICIP, pp. 776-779, 2001.

[5] A. Nguyen, V. Chandran, S. Sridharan, and R. Prandolini, "Importance coding in JPEG2000 for improved interpretability," IVCNZ, pp. 339-344, 2001.

[6] D. Taubman, "Kakadu: A comprehensive, heavily optimized, fully compliant software toolkit for JPEG2000 developers," http:/maestro.ee.unsw.edu.au/~taubman/kakadu/, 2002.