Can Region of Interest Coding Improve Overall Perceived Image Quality?

Andrew P. Bradley
Centre for Sensor Signal and Information Processing (CSSIP),
School of Information Technology and Electrical Engineering,
The University of Queensland,
St Lucia, QLD 4072, Australia.
a.bradley@cssip.uq.edu.au

Abstract

In this paper we investigate the possibility of improving the overall perception of image quality by preferentially coding certain regions of interest (ROI) in an image. Experiments are conducted utilising an automated algorithm for visual attention (VA) to detect the primary ROI(s) in an image, and then encoding the image using the maxshift algorithm of JPEG 2000. The results from a 2 alternative forced choice (2AFC) visual trial show that, while there is no overall preference for the ROI encoded images, there is an improvement in perceived image quality at low bit rates (below 0.25 bits per pixel). It is concluded that a perceived increase in overall image quality only occurs when the increase in quality of the ROI more than compensates for the corresponding decrease in quality of the image background (i.e., non ROI).

Keywords

JPEG 2000; Region of Interest Coding; Image Quality.

Introduction

With the introduction of third generation (3G) mobile devices there will be an increasing demand for the efficient transmission of multi-media data, such as speech, audio, text, images, and video. Of these multi-media data types image and especially video data will provide the toughest challenges because of their high bandwidth and user expectations in terms of high quality of service. Therefore, to enable the successful adoption of the multitude of 3G applications, the transmission of multi-media data must be at high compression ratios and be of a perceptually high quality. In this paper we shall investigate the suitability of the region of interest (ROI) coding feature provided by JPEG 2000 (JP2K) to improve the perceptual quality of compressed images, where the ROI has been automatically extracted from the image by using an algorithm that simulates visual attention (VA).

JPEG 2000 (JP2K) is the emerging image and video compression standard developed by the International Organisation for Standardisation/International Telecommunications Union (ISO/ITU-T). JP2K has been designed to complement the current JPEG standard by providing improved compression performance and a rich set of new functionalities [7]. JPEG 2000 Part I, the core-coding algorithm, became an international standard in December 2000 and further work is ongoing to tailor the standard to specific applications, such as medical imaging and video coding. JP2K provides, in a single bit-stream, a broad set of functionalities, such as: progressive transmission by resolution, quality, component, or location; random access; lossless to lossy compression; and error tolerance [3]. The specific functionality investigated in this paper is the ability of JP2K to encode a region of interest (ROI) in an image with more detail than the background. In this paper JP2K ROI coding is used in combination with an algorithm for visual attention (VA) [5, 6], to provide a progressive bit-stream where the regions highlighted by the VA algorithm are presented first in the bit-stream. This results in an *interest ordered* bit-stream where any valid bit-stream termination results in an image where the ROI is coded to a higher quality than the background. The efficacy of this technique is then evaluated using a visual trial to determine under what conditions it provides an increase in overall perceived image quality compared to conventional JP2K at the same bit-rate.

As has been demonstrated in previous rate distortion experiments [1], it is important to reduce the overhead associated with ROI coding in order to ensure maximum coding efficiency. Briefly, this can be achieved by:

1. Ensuring the ROI is < ¼ of the area of the whole image;
2. Reducing the number of regions of interest (to two or less); and
3. Ensuring region boundaries are reasonably regular (smooth).

The first constraint ensures that there are enough available bits in the background to be able to preferentially encode the foreground ROI. Whilst the last two constraints ensure that the overhead associated with ROI coding is minimised, e.g., for maxshift coding it minimises the number of code-blocks that contain coefficients from both the ROI and the background. The algorithm for processing the VA map produced by [5, 6] to meet the above constraints is detailed in [2]. An example image, with the ROI selected by the VA algorithm highlighted, is shown in Figure 1.

Figure 1. Cycles image with VA ROI highlighted.

Experimental Methodology

As discussed in [1] selection of the most appropriate JP2K ROI encoding methodology for a particular application is dependent upon a number of factors: the desired bit-rate; relative ROI/background importance; the shape and size of the ROI; and whether the ROI is fixed or is to be selected by the user. For client/server applications it is essential to be able to extract any ROI from an encoded image, in which case code-block selection is the best method to use [4]. However, in the type of applications we are considering the ROI can be calculated directly from the image and is fixed. Therefore, it is desirable to have the ROI embedded in the bit-stream using coefficient scaling [1]. In addition, the ROI is assumed to be of primary importance and so we desire to receive it as early as possible in the bit-stream. Therefore, we shall use the method of coefficient scaling provided in Part I of the JP2K standard, the max-shift algorithm. In addition, we shall tailor the JP2K ROI coding to our particular requirements by using: small (16x16) code-blocks for fast ROI refinement; a 5 level irreversible (bi-orthogonal spline 9/7) wavelet transform for high compression (lossy) efficiency with the lowest level of the DWT defined to be part of the ROI; and an increased quantisation step size (of 0.03125 which is four times the default) to prevent ROI over-coding.

The visual trial was based upon six images, namely: boat, cycles, beach, helicopter, land, and road sign. These images were chosen to have a reasonably varied content, whilst still containing one or two primary objects that could be considered to be more important (visually interesting) than the background. The images selected for the visual trial are not intended to be representative of any particular potential application, but were chosen solely to judge the efficacy of ROI coding in JP2K.

The purpose of the visual trial was to directly compare images encoded to a specified bit-rate using standard JP2K and JP2K ROI coding, where the ROI is determined using the VA algorithm [2]. The comparisons were made at four logarithmically spaced bit-rates (and hence varying image qualities) of 0.125, 0.25, 0.5, and 1 bits per pixel (bpp). A two alternative forced choice (2AFC) methodology was selected because of its simplicity, i.e., the observer views the two images and then selects the one preferred, and so there are no issues with scaling opinion scores between different observers. There were ten observers (8 male and 2 female) all with good, or corrected, vision and all observers were non-experts in image compression. The viewing distance was approximately 40cm (i.e., a normal PC viewing distance) and the image pairs were viewed one at a time in random order. The observer was free to view the images multiple times before making a decision, however a buzzer sounds after 20 seconds to indicate that they should make a decision. In addition, a blank mid-grey image is shown between each image (for 2 seconds) to prevent observers switching between the two images to find insignificant differences. Each image pair was viewed twice, giving ($6\times4\times2$) 48 comparisons, which means that each observer takes approximately 10 minutes to view all of the images. Images were viewed on a 12.1" Thin Film Transistor (TFT) display, in a darkened room (i.e., daylight with drawn curtains). The test images were displayed on a mid-grey background to a maximum size of 410×600 pixels. Prior to the start of the visual trial all observers were given a short period of training on the usage of the visual trial software and they were told to select they image they preferred assuming that it had been downloaded over the internet or wireless network.

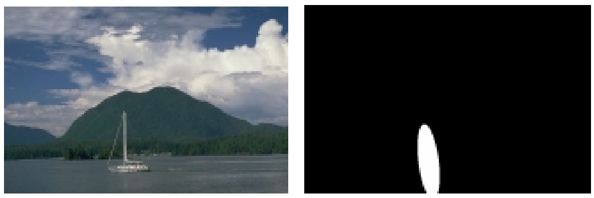

Figure 2. Boat image and VA ROI mask.

Figure 3. Beach image and VA ROI mask.

Results

Table 1 shows the overall preferences, i.e., independent of (summed over) image and bit-rate, for standard JP2K and JP2K ROI coding with the ROI determined using the VA algorithm. Table 1 also shows the standard errors associated with the preferences assuming a Gaussian approximation to the Binomial distribution. From Table 1 it can be seen that standard JP2K is preferred over ROI coding approximately 65% of the time. This shows that standard JP2K produces good quality images over a wide range of bit-rates and indicates that ROI coding may not be suitable as a general-purpose image coding technique. Therefore, we will have to examine the results in more detail to identify the conditions to which the ROI JP2K coder is best suited.

Table 1. Overall preferences (independent of image and bit-rate)

Compression Method	Number of Preferences	Standard Error
JP2K	311	± 12.3
JP2K ROI	169	± 12.3

Figure 4 shows that preferences vary both across the image set (independent of bit-rate) and with bit-rate (independent of image). Standard JP2K is shown with red standard error bars (on the left) whilst JP2K ROI coding is shown with blue standard error bars (on the right). From Figure 4 it can be seen that there is a large variation in preferences across each of the images in the test set. For example, standard JP2K is preferred at every bit-rate on the boat image, whilst the two methods are equivalent on the cycles and beach images (within 1 standard error). However, the second, and more important, trend that can be observed in Figure 4 is an increase in preferences for ROI coding as the bit-rate decreases. At the lowest bit-rate tested (0.125 bpp) the preferences for ROI coding are 68, with a standard error of ± 5.8, and 52 (± 5.8) for standard JP2K. This indicates a clear preference (i.e., statistically significant) for the JP2K with ROI coding at this bit-rate.

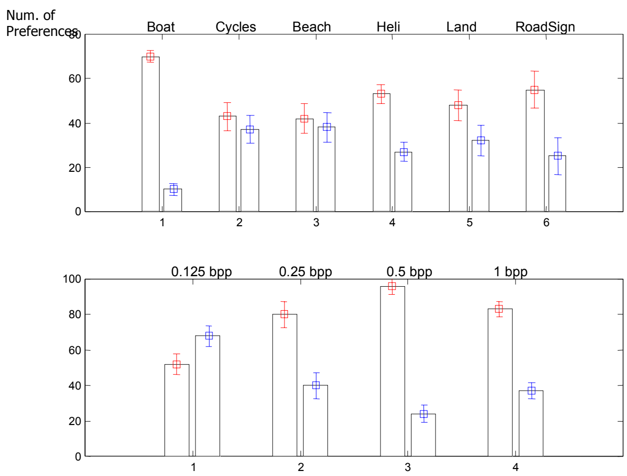

Figure 4. JP2K (left) and ROI JP2K (right) preferences for each image (independent of bit-rate) and preferences at each bit-rate (independent of image).

Note: the decrease in preferences for standard JP2K at 1 bpp in Figure 4 is due to the two methods producing image that look increasingly similar. Therefore, preferences between the two methods will tend to random (i.e., 50/50) selection.

Discussion

As illustrated in Figure 4, there are two main sources of variation that can explain the differences in preferences: variation with image and variation with bit-rate (variation with observer being indicated by the standard errors in the results). The variation across the images in the test set shows that for an image that has an ROI and a background of little importance, such as the beach image (see Figure 3), the ROI coding works well. However, for an image that has an ROI and also some visually important contextual details in the background, such as the boat image (see Figure 2), the ROI coding works less well. The increase in preferences for the ROI coding as bit-rate decreases, illustrated in Figure 4, is undoubtedly the most significant and consistent effect observed in the visual trial (being apparent in 5 of the 6 images in the test set).

It is worthwhile noting that the performance of ROI coding on the boat image was degraded by the fact that the ROI found did not enclose the whole region of primary interest. However, in general it is probable that most images that have a primary ROI will also have some important contextual details in the background and so ROI coding is unlikely to provide an overall improvement in image quality at all bit-rates.

Figure 5. Cycles image JP2K (1bpp).

Figure 6. Cycles image JP2K ROI (1bpp).

At low bit-rates (< 0.25 bpp), having the ROI encoded first in the bit-stream can significantly improve the visual quality of the ROI compared to standard JP2K. In addition, the background (non-ROI) areas are not of significantly poorer visual quality and are often of preferable visual quality as they contain less (wavelet) compression artefacts. At the low bit-rates the background regions tend to contain only coefficients from the lowest level of the wavelet transform rather than sporadic coefficients from higher levels of the DWT (as in standard JP2K). This results in a background that is uniformly blurred, which is often preferable to a less blurred background that also has wavelet artefacts.

At the higher bit-rates (> 0.25 bpp) the ROI is often not of significantly better visual quality than standard JP2K. This combined with the fact that the background areas are often more blurred and pixelated than standard JP2K results in lower preferences (see Figures 5 and 6). This effect should come as no surprise as once the ROI is coded to a visually acceptable level it takes a significant number of bit refinements (of the high entropy least significant bits) to get a visible improvement in image quality. In addition, because small code-blocks were used at all bit-rates for the ROI coding (16×16 compared to 64×64) there is a reduced compression efficiency, especially with the entropy coding of the code-blocks. This reduction in compression efficiency is particularly apparent at the higher bit-rates due to the increased number of significant coefficients [1]. However, using small code-blocks reduces the ROI coding overhead and therefore ensures that the complete ROI appears as early as possible in the bit-stream.

Another reason for the reduction in preferences at bit-rates > 0.25 bpp is due to the inherently uneven image quality in the majority of ROI coded images. This results in images that do not appear *natural* as the ROI is in sharp focus, whilst the background appears blurred. A more gradual change in image quality between ROI and background would, however, increases the size of the ROI, which has a negative impact on ROI coding efficiency.

There is an anecdotal explanation for the reduction in preferences at bit-rates > 0.25 bpp by considering the rule of thumb that to observe a significant increase in the visual quality of an image you have to (approximately) double the bit-rate. This means that the ROI has to be coded to twice the bit-rate of the background to observe a significant improvement in perceived visual quality. Therefore, if we assume the ROI is ¼ of the image area, then to code an image to the same (target) bit-rate as standard JP2K, the background can only be coded to half the target rate to allow the ROI to be coded at twice the target rate. For example, if the target bit-rate is 0.5 bpp, then we can either code to this bit-rate using standard JP2K, or code the ROI to 1 bpp and the background to 0.25 bpp using JP2K ROI coding. Therefore, ROI coded images will invariably have an ROI that looks better, but a background that looks worse, than standard JP2K images coded to the same bit-rate. Results observed in this visual trial indicate that the ROI encoded images only score an overall improvement in image quality at target bit-rates less than 0.25 bpp (i.e. 0.125 bpp). At bit-rates greater than 0.25 bpp the increase (if any) in quality of the ROI does not compensate for the decrease in background quality when observers judge overall image quality.

Conclusions

Results from the visual trial indicate an overall preference for standard JP2K independent of image and bit-rate. However, the proposed VA ROI JP2K coding method was clearly preferred at the lowest bit-rate tested (0.125 bpp). This indicates that, when observers judge overall image quality, it is only at this bit-rate that the visible increase in quality of the ROI more than compensates for the decrease in quality of the background. Therefore, it can be concluded that ROI coding in JP2K will only produce an overall increase in perceived image quality when: the image contains a small number (≤ 2) of regions of interest; these regions are relatively small (< ¼ of the total image area); and the bit-rate is low enough to produce visible compression artefacts (< 0.25 bpp).

Acknowledgements

This work was undertaken whilst the author was a visiting research fellow at Btexact Technologies, Adastral Park, Martlesham Heath, Ipswich, UK. All experiments were performed with JJ2000, version 4.1 (the current version is available from http://jj2000.epfl.ch).

References

[1] A. P. Bradley and F. W. M. Stentiford, JPEG 2000 and Region of Interest Coding, *Digital Image Computing Techniques and Applications* (DICTA), Melbourne, Australia, January 2002

[2] A. P Bradley and F. W. M. Stentiford, Visual Attention for Region of Interest Coding In JPEG 2000, Submitted to *Journal of Visual Communication and Image Representation*, January 2002.

[3] C. Christopoulos, A. Skodras, and T. Ebrahimi, The JPEG 2000 Still Image Coding System: An Overview, *IEEE Transactions on Consumer Electronics*, Vol. 46, No. 4, pp. 1103-1127, November 2000.

[4] D. Santa-Cruz, T. Ebrahimi, M. Larsson , J. Askelof, and C. A. Christopoulos, Region of Interest Coding in JPEG 2000 for Interactive Client/Server Applications, *IEEE 3rd Workshop on Multi-media Signal Processing* (MMSP), pp.389-394, September 1999.

[5] F. W. M. Stentiford, An Evolutionary Programming Approach to Simulation of Visual Attention, *Congress on Evolutionary Computation*, Seoul, Korea, pp. 851-858, May 2001.

[6] F. W. M. Stentiford, An Estimator for Visual Attention Through Competitive Novelty with Application to Image Compression, *Picture Coding Symposium*, Seoul, Korea, pp. 101-104, April 2001.

[7] D. S. Taubman and M. W. Marcellin, *JPEG 2000: Image Compression Fundamentals, Standards, and Practice*, Kluwer Academic Press, 2001.