Region-Based Guaranteed Image Quality in JPEG2000

T M Trenschel, T R Bretschneider, G C Leedham
School of Computer Engineering
Nanyang Technological University

Abstract

This paper describes a spatially variant method that optimises the compression rate within a given image dynamically depending on the application specific requirements and the actual image content. The concept utilises the JPEG2000 standard, however, a generalised region-of-interest (ROI) map is used in this paper together with a different post-compression rate distortion optimisation. The generalised ROI map comprises a multi-level priority scheme as well as continuously adjustable qualities for each individual region to reflect the application specific requirements for a particular image. The key point is that these pre-defined quality constraints are always met for each individual region, thus making the approach suitable for data with inherent reliability requirements. The implementation of the proposed system is transparent and concerns only the encoder side while any standard JPEG2000 decoder can be used for decompression. Results have shown that for the utilised imagery, i.e. satellite and medical images, compression rate improvements of up to 32% in comparison to the JPEG2000 compressor fulfilling the same quality needs are achievable.

1. Introduction

The storage and bandwidth requirements of nowadays compute systems cannot keep up with the exponentially growing data flood from the variety of acquisition systems. In particular this holds true for all imaging devices that produce data of high spatial, spectral, radiometric and/or temporal resolution. Typical examples can be found in spaceborne remote sensing and medical imaging. Since the data has to be archived for many years, i.e. for monitoring purposes and legal matters, compression is mandatory. High quality demands without introducing distortions are required by the corresponding applications. Often this only allows the utilisation of lossless compression techniques. But these cannot achieve the same compression rates as lossy compression, and thus result in larger storage volume and increased bandwidth demands. However, if application relevant areas can be identified then the output of the compression can be optimised by using a spatially variant compression rate for different image locations. Thus image parts that can tolerate more distortion are compressed more than others with less tolerance.

In numerous applications JPEG2000 has become the new standard for image compression since it solves many problems and shortcomings of its predecessor JPEG, i.e. the support of an arbitrary number of image bands with user-selected resolutions, a better performance at very high compression rates, regions-of-interest etc. The most common use for JPEG2000’s lossy compression mode is to produce for a pre-defined bandwidth an image that is optimised with respect to the human perception. This subjective compression approach has an implication on all applications that incorporate image analysis, i.e. compression inherent artefacts can affect the result. These distortions can make highly valuable data like satellite or medical imagery worthless. However, most applications have certain foci on specific aspects within an image. If these are detected and described in terms of regions a-priori then the actual compressor can incorporate this knowledge and adapt the compression. The JPEG2000 standard allows the implementation of ROIs a framework for this approach, but the distinction between foreground and background is not sufficient for many real-world applications since they have a finer definition for the objectives in certain image regions. Thus it is required to extend the concept towards a method that guarantees continuously adjustable lower image quality bounds for arbitrary parts of the image.

This paper describes a novel technique that implements the concept of guaranteed image quality in the JPEG2000 standard. The approach proposes a generalised concept that relates a quality requirement to each pixel in the image. Therefore it is possible to fulfil given constraints from image analysis algorithms with respect to the maximum tolerable distortion for proper processing. Moreover the concept is extended by incorporating priorities. This aspect is of particular interest for applications that have a well-defined order of objectives with respect to the image regions but a limited bandwidth, which might not allow the transmission of the entire image in every case or at the demanded quality. An example is given by low earth orbiting remote sensing missions where the satellite is only for a short period of time in contact with the ground receiving station. The quality and priority information together form the user’s interest and are transparently incorporated in the JPEG2000 framework. Hence, any JPEG2000 decoder can be used for the decompression of the image. To enable this functionality a different post-compression rate allocation scheme was implemented. Essentially, the contribution of each data block is determined by selecting the truncation
point – namely the rate of compression – in such a way, that all requested qualities above a chosen priority threshold are met. If there is still bandwidth left, then the threshold can be lowered to include additional objectives, i.e. further image data. Thereby all image sections meet the required quality for all priority levels supported by the given bandwidth if any.

A schematic overview of the data flow for the entire system is given in Figure 1. Firstly, the image is analysed, whereby the actual processing depends on the particular objective of the user. The availability of a database, which for example stores a catalogue of representative artefacts like tumours, can support the analysis. The extracted information is then used to compute a generalised ROI map, the so-called compression map. This map $C$ represents the application related varying value of different areas in the image, i.e. it is a quantitative representation with respect to the underlying image $S^N$ and is described by $\Phi: S^N \rightarrow \mathbb{C}^{2N}$, where $N$ is the number of radiance values per pixel. Note that any meaningful image analysis function $\Phi$ can be used. Furthermore the compression map describes the priority of the related areas in terms of progressive coding. Thus the mapping function is described by $\Phi: S^N \rightarrow \mathbb{C}^{2N}$. Often actual applications do not make use of the full $2N$ values but rather a smaller number if this is sufficient. The successive image compression step computes a compressed image – some parts possibly lossless, others lossy – as defined by the compression map, which guarantees an optimised storage / bandwidth usage.

![Figure 1: Dataflow for content-based compression](image)

The remaining part of the paper is organised as follows: Section 2 provides a brief overview on the terminology in JPEG2000 and introduces the bit allocation mechanism that is modified in the subsequent Section 3. The emphasis is on the compression map concept and the modified bitstream truncation scheme. Actual results for satellite imagery and medical images are presented and discussed in Section 4. Finally, Section 5 concludes this paper.

## 2. Overview and Analysis of JPEG2000

The actual compression of the proposed approach is based on the JPEG2000 standard [1] since it provides a framework for supporting multispectral imagery and uses an efficient wavelet compression scheme, which can utilise lossless and lossy compression seamlessly within an image. It can also divide the image into tiles and compress them independently.

One of the major processing steps in JPEG2000 is the embedded block coding with optimised truncation (EBCOT) [2] to produce resolution or quality progressive bitstreams. EBCOT consists of two different stages. Firstly, it divides the image into code blocks of typically $32 \times 32$ or $64 \times 64$ pixels. The second step performs the so-called post-compression rate-distortion optimisation (PCRD-opt) that minimises the total distortion of all code blocks described by

$$ D = \sum z_i D^{(z_i)}, $$

whereby $z_i$ indicates the optimal choice of where to truncate the embedded bitstream of code block $i$. The corresponding distortion for the computed $z_i$ is described by $D^{(z_i)}$. Note that each code block is encoded independently. The optimisation is performed with respect to the user-defined compression rate, i.e. the maximum bitstream length $L_{\text{max}}$. Thus the constraint

$$ \sum L^{(z_i)} \leq L_{\text{max}} $$

has to be met, where $L^{(z_i)}$ is the length of the bitstream for the code block $i$ with the truncation point $z_i$. By relating the distortion $D$ and the bitstream length $L$ (basically a cost function), the distortion-distortion-rate slopes $\lambda_i$ defined by

$$ \lambda_i = \frac{\Delta D}{\Delta L} = \frac{D^{(z_i+1)} - D^{(z_i)}}{L^{(z_i+1)} - L^{(z_i)}} $$

are described, where the individual $D_i$ and $L_i$ are the distortions and bitstream lengths for the code block $i$, respectively. The superscripted indices symbolise possible truncation points. Thus a measure of the distortion reduction $\Delta D$ contributed by adding $\Delta L$ symbols of the code block $i$ in order to transmit its wavelet coefficients more accurately is provided. Figure 2 shows the distortion-rate slopes as solid lines and the corresponding convex hull by a dashed line for an arbitrary code block. Note that the first one is the result of the discrete nature of truncating the bitstream, while the latter plot assumes continuous truncation points. For a given $L_{\text{max}}$ the PCRD-opt algorithm determines a minimal threshold $\lambda$ for all code blocks $i$ to find the truncation points $z_i$ according to

$$ \lambda^{(z_i+1)} < \lambda \leq \lambda^{(z_i)}. $$

If the distortion-rate slope is steeper, then Equation (4) selects a later truncation point $z_i$ and allocates more symbols for the respective code block. Further details on EBCOT and its PCRD-opt can be found in the work of Taubman [1].
3. Guaranteed Image Quality Model

The ROI map concept in JPEG2000 needs to be extended to support the priorities and distortion limits of the compression map. Note that in many applications the required data representation accuracy and the priority are correlated, but this is not a general requirement [4]. In fact, the maximal tolerable distortion is completely independent of the priority. This becomes apparent if the objective to transmit the complete image lossless \((D_{\text{limit}} = 0)\) is associated with the lowest possible priority. In this case if and only if the available bandwidth allows, the original image is received. Thus the proposed framework gives an additional flexibility over the existing approach. Moreover the structure of a compression map caters for a variety of foci that enable the straightforward extension to new requirements, e.g. adaptive channel coding schemes with respect to the data content.

The separation of different emphasises is a major change to the ROI concept of JPEG2000, which comprises only the binary fore- and background information. For the implementation of ROIs the code block based approach was used, since the alternative of using pixel-based coefficient scaling beyond binary ROIs is not supported by the current JPEG2000 Part 1 standard.

The following sub-sections describe the propagation of the compression map through the different levels of the wavelet decomposition and the chosen avenue of modifying the post-compression rate allocation scheme of EBCOT.

3.1. Compression Map Concept

The wavelet transform calculates wavelet coefficients, which again are derived from several coefficients of the previous resolution level [5]. Hence, a coefficient \(c_r(n,m)\) in resolution level \(r\) belongs to the ROI if any of its predecessor coefficients in resolution level \(r-1\) were part of the ROI. Technically, the ROI mask in JPEG2000 is calculated using the binary conjunctive operator on the ROI information of all parent coefficients. In contrast to the binary ROI concept of JPEG2000, the compression map consists of two continuous components for each coefficient, namely the priority \(P_r(n,m)\) and the distortion limit \(D_{\text{limit},r}(n,m)\). To maintain the correct priority and distortion in the next lower resolution level the conjunctive operator has to be replaced. The priority \(P_r(n,m)\) becomes the maximum of the parent pixel priorities \(P_{r-1}\), and the correct \(D_{\text{limit},r}(n,m)\) quality is met by the minimum of the parent \(D_{\text{limit},r-1}\).

Due to this propagation of the minima the achievable compression in the lower resolution levels is reduced, i.e. information that normally would require a shorter bitstream has to be encoded with a higher precision. In this context it is worthwhile mentioning that most of the time the proposed approach results in a better performance for smaller code blocks while generally JPEG2000 achieves better compression ratios for larger code block sizes. However, two different aspects ease the increased bitstream drawback of the proposed block-based quality adjustment: Firstly, low resolution levels need substantially less symbols to be represented in a high quality, because a lower resolution means less coefficients for the same image area. Secondly, it is possible to limit the propagation of very small \(D_{\text{limit}}\) by dividing the image into tiles [6]. The JPEG2000 standard supports compressing tiles of an image independently and the ROI propagation is blocked at tile borders.

3.2. Content-based Compression

The JPEG2000 PCRD-opt algorithm is very suitable with respect to a human observer, since for a given compression rate it allocates for each code block as many symbols as required to give a homogeneous level of quality improvement. Obviously, this falls short if a defined level of quality for each code block, as it was defined by the compression map, is required.

After the propagation of the compression map on a pixel level, the coefficients are grouped into code blocks, which are compressed separately. Hence, each code block \(i\) has a
unique spectrum of tolerable distortions \((p)D_{\text{limit,}i}\) for a requested priority \(P\). This spectrum \((p)D_{\text{limit,}i}\) is derived by taking the minimum of \(D_{\text{limit,}i}\) at each available priority of a code block \(i\). The distortion rate curve is calculated the same way as in EBCOT, but the selection of truncation points is modified, i.e. the appropriate truncation point \(zi\) is selected from the distortion spectrum. For a particular code block \(i\) it gives the information how much distortion could be tolerated if compressed for a particular priority level \(P\). The truncation point \(zi\) is therefore the smallest \(n\) that fulfils the following equation:

\[D_i^{(n)} \leq (p)D_{\text{limit,}i}\]  

(5)

The new post-compression rate allocation, which is replacing JPEG2000’s PCRD-opt algorithm, works as follows:

1. Start with priority \(P\), which is the highest available priority level
2. For each code block select a truncation point \(zi\) to fulfil the equation \(D_i^{(zi)} \leq (p)D_{\text{limit,}i}\)
3. If \(\Sigma L_i^{(zi)} > L_{\text{max}}\) then produce the JPEG2000 bit-stream with the next higher priority and stop
4. Decrease the priority \(P\) to the next available objective and continue with step 2

After termination of the algorithm it is possible to further improve the utilisation of the available bandwidth by filling the remaining bandwidth by partly transmitting the information requested by the next lower priority level.

4. Results

Actual measurements using the proposed content-based compression technique were carried out for two different application fields, which highly depend on an accurate representation of the original information while at the same time impose a great burden on the bandwidth and storage, respectively.

4.1. Application in Remote Sensing

For a first application with content-based compression a binary prioritisation scheme was investigated. The scenario assumes that the user requires the original data, but some parts of the multispectral satellite image are of no value. As an example the monitoring of landmasses was selected, whereby water areas are regarded as noise with respect to the application. For the extraction of relevant areas global thresholds were applied to the quotient images of the spectral bands since this approach is robust against illumination changes. After additional morphological erosion the result is a binary image indicating water bodies as application noise. Then the compression-map was computed by assigning an infinite tolerable distortion to pixels that were identified to be part of the water area and all other pixels received a distortion limit of zero. Figure 3 shows a sub-section of the original image and the computed compression map, which for the reason of simplicity, reflects the correlated quality and priority description.

![Figure 3: Multispectral false colour satellite scene (24 bits) and corresponding quality compression map](image)

The compression results in bits per pixel (BPP) for the entire scene of 1500x1500 pixel are shown in Table I. Since a valid approach for the depicted scenario is to blank out all pixels that the user is not interested in, the table also shows the compression performance for replacing water pixels with zeros. Note that in all cases EBCOT is bounded by the minimal tolerated distortion.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Block size</th>
<th>8×8</th>
<th>16×16</th>
<th>32×32</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contend-based comp.</td>
<td></td>
<td>12.64</td>
<td>11.93</td>
<td>11.90</td>
</tr>
<tr>
<td>EBCOT comp.</td>
<td></td>
<td>15.69</td>
<td>14.46</td>
<td>14.02</td>
</tr>
<tr>
<td>EBCOT-blank comp.</td>
<td></td>
<td>12.48</td>
<td>11.57</td>
<td>11.25</td>
</tr>
</tbody>
</table>

When examining the results of the content-based compression one observes, that the water pixels are not completely lost, but the information is slowly degrading with increasing distance to the coast line. Thus, the content-based compressed image contains significantly more information, while its compression ratio is very close to a lossless compressed image where all water pixels have been blanked out. The information is only slowly degrading, because some information of code blocks next to the region-of-interest is preserved in the next higher decomposition level. Another advantage of the content-based compression is its robustness against classification errors. When ROI pixels and non-ROI pixels are aligned within the same code block, the quality requirement of the ROI pixel determines the truncation point. Therefore any misclassifications like observable around the southern river delta in Figure 3 have no impact.

The main performance gain of content-based compression becomes evident if a non-binary interest in terms of tolerable distortion is considered. Three different applications for the three ground cover types ‘urban’, ‘agriculture’ and ‘water’ were selected. Individual distortion tolerance values were assigned to the corresponding regions of a multispectral scene obtained by the SPOT satellite whereby the
urban parts always had to be compressed lossless. Table II provides a summary of the results in terms of BPP. Taking from each approach the optimum value the saved amount of data lies for the given example in the range of 4.46 Mbit and 30.64 Mbit.

Table II: Comparison of compression achievements in BPP

<table>
<thead>
<tr>
<th>Application</th>
<th>Algorithm</th>
<th>Block size</th>
<th>8x8</th>
<th>16x16</th>
<th>32x32</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cont.-based</td>
<td>7.54</td>
<td>8.63</td>
<td>9.44</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Cont.-based</td>
<td>11.88</td>
<td>11.02</td>
<td>10.76</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Cont.-based</td>
<td>11.38</td>
<td>10.71</td>
<td>10.59</td>
<td></td>
</tr>
<tr>
<td>1, 2, 3</td>
<td>EPCOT</td>
<td>11.80</td>
<td>11.60</td>
<td>11.11</td>
<td></td>
</tr>
</tbody>
</table>

4.2. Application in Medical Imaging

In a second investigation the content-based compression is applied to a magnetic resonance image. In the given example in Figure 4 two multiple sclerosis plaques are identified that have to preserved together with their immediate surrounding to enable a later assessment of the treatments effects.

Figure 4: Magnetic resonance image [7] with highlighted multiple sclerosis and corresponding compression map

With the a-priori knowledge that the problem is most likely located more towards the centre of the skull the quality descriptors were selected according to the increasing brightness levels. The successive compression yielded to a decreased storage requirement of 26.2% in comparison to the standard JPEG2000 compression. Note that the white areas were compressed lossless while the light grey regions exhibit a difference in absolute values between the original and the compressed version of less than one coding value.

5. Conclusion

An extended framework for JPEG2000 was proposed that introduces the concepts of multi-level priorities and upper bounds on the compression related distortions. In contrast to the original standard the suggestions enable an arbitrary quality for any region within the image. In simple words the scheme employed by JPEG2000 to allocate code symbols for each code-block is based on the idea to minimise the global distortion. Therefore EBCOT allows distortions, if for the saved symbols a higher distortion in another code-block can be avoided. For content-based compression the aim is to guarantee a certain spatially varying level of quality and the truncation points are chosen accordingly for each code-block independently. It is obvious, that EBCOT is superior compared to the content-based compression scheme when compressing an image with a constant quality level, which does not allow the distribution of the unavoidable distortion in an optimal fashion. For lossless compression both strategies perform equally, since the truncation point for each code-block is pre-defined by the constraint to be represented error free. However, for content-based compression the proposed flexible coding scheme results in significant larger compression rates than JPEG2000 and thus helps to conserve storage and bandwidth. One of the major advantages is that the approach is transparent to the decoder.

References