Rate-controlled, region-of-interest-based image coding with JPEG-LS

This item was submitted to Loughborough University’s Institutional Repository by the author.


Additional Information:

• Copyright 2003 Society of Photo-Optical Instrumentation Engineers. One print or electronic copy may be made for personal use only. Systematic electronic or print reproduction and distribution, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited. This paper can also be found at: http://dx.doi.org/10.1117/12.476372

Metadata Record: https://dspace.lboro.ac.uk/2134/6512

Version: Published

Publisher: © 2003 SPIE

Please cite the published version.
This item was submitted to Loughborough’s Institutional Repository (https://dspace.lboro.ac.uk/) by the author and is made available under the following Creative Commons Licence conditions.

For the full text of this licence, please go to:
http://creativecommons.org/licenses/by-nc-nd/2.5/
Rate Controlled, Region of Interest Based Image Coding with JPEG-LS

E.A.Edirisinghe *, S.Bedi
Department of Computer Science, Loughborough University, UK

ABSTRACT

Since the standardization of JPEG-LS, several improvements and variations to the original algorithms have been proposed. In this paper we propose a Region of Interest (ROI) based coding strategy for JPEG-LS that has the additional ability of providing effective rate controlled image compression. Given a single ROI or multiple ROIs of a fixed or arbitrary shape, the scheme we propose is able to compress a given image by a required ratio, whilst maintaining the subjective image of the ROIs at either lossless or at a quality specified by a Target Compression Ratio (TCR) of the ROI. We provide experimental results to compare the performance of the proposed rate-control algorithm with the state of the art near lossless rate control schemes. We show that the proposed scheme is able to achieve much higher TCRs, at increased accuracy and better objective image quality, using a less computationally intensive rate control algorithm. Finally we demonstrate that the proposed ROI based coding scheme can be used to extend the applicability of JPEG-LS to medical and satellite imaging applications and provides a useful alternative to JPEG-2000 based ROI coding.

1. INTRODUCTION

High quality, high compression rates, and low computational cost are important factors in many areas of digital imaging, ranging from digital photography to advanced consumer electronic applications. However the relative importance of these factors is application dependant. For example, in video telephony applications, low computational cost and high compression rates are imperative, whereas in medical imaging devices, high quality and low computational cost are a priority. Due to the high demand of applications that require coding schemes that satisfy the latter set of conditions [3,11], lossless/near-lossless compression schemes such as FELICS [5] and LOCO-I [21] have by now evolved into an international standard, JPEG-LS [11], with the major contribution coming from LOCO-I (Low Complexity, Context-Based, Lossless Image compression algorithm).

JPEG-LS consists of two modes of operation namely, lossless mode and near lossless mode. In its near-lossless (NLS) compression mode JPEG-LS introduces a small amount of information loss, which is solely dependant on a single constant parameter, NEAR (NEAR>0 for near-lossless mode and NEAR=0 for lossless mode). In other words, regardless of the nature of the local texture, the information loss introduced throughout the image dimension fall within the same order of magnitude. This could cause a conflicting problem between the compression efficiency and quality requirements, since the information loss introduced to smooth areas is the same as that applied to highly textured/edgy, noisier areas. Yet it has been shown that human visual perception is more sensitive to smoother areas than to noisier areas [17,19,20]. Considering these drawbacks, in [17], authors proposed the replacement of the single information loss parameter NEAR, by three independent information loss parameters namely to determine the information loss in differently textured areas within an image. In [19,20] they extended the use of this scheme to rate-controlled image compression, aimed at extending JPEG-LS’s use to application areas such as digital photography that requires images to be compressed at a pre-specified rate, to be stored within a given space of memory or storage media.

A detailed investigation carried out by us in [22] revealed that the rate control strategy proposed in [19] is not dynamic enough to follow a given Target Compression Ratio (TCR) and thus fails in achieving higher TCRs with the required accuracy. This is due to the fact that it only changes one information loss parameter per compression assessment point and would on some occasions end up changing an information loss parameter that does not correspond to the local texture within the area under consideration, thus making the rate control inefficient. In [20] a rate-control strategy that is based on a highly modified version of JPEG-LS has been proposed. In this work the JPEG-LS’s probabilistic modeling
part has been replaced by a much simpler look-up table, to provide statistical estimation for Golomb coding \[^{12,13}\]. A line-fitting technique based on a careful selection of a number of parameters has been used in predicting the change of the so-called On-line Compression Ratio (OCR), which is controlled to change towards the specified TCR. The algorithm has been optimized for obtaining a good subjective reconstructed image quality. Unfortunately due to the nature of its design, the objective image quality of reconstructed images is poor and the complexity reduction introduced by novel probabilistic modeling process is largely offset by its sub-optimal efficiency. In order to address the problems of the above rate control schemes, within the research context of \[^{22}\], we proposed an improved rate control strategy that is less computationally costly as compared to the methods in \[^{19,20}\] and results in a more accurate rate-control and a better rate-distortion performance. In addition we have shown that the rate-control scheme of \[^{22}\] has improved flexibility in terms of operational range and adaptability as compared to the above \[^{19,20}\] techniques, which enables its efficient use in ROI based coding. Within the research context of this paper we apply this rate control scheme to ROI based coding, demonstrating its practical usefulness in real world applications.

For clarity of presentation, this paper is divided into five sections. Apart from this section, which is an overview to the proposed work, in section 2, we introduce the reader to the original JPEG-LS context modeling and prediction scheme and summarize the subsequent modifications proposed in \[^{17}\] for information loss re-distribution aimed at improving subjective image quality. In section 3, we summarize the design and performance of our low cost rate control scheme of \[^{22}\]. In section 4 we make use of this rate control scheme in further extending the applicability of JPEG-LS to ROI based coding. In section 5, we provide extensive experimental results and a detailed analysis to prove the effectiveness of the proposed ROI based coding strategy. Section 6 concludes, with an insight into the future directions of research.

## 2. INFORMATION LOSS DISTRIBUTION & RE-DISTRIBUTION IN JPEG-LS

JPEG-LS compression algorithm specifically comprises of the following sections; run-length Coding, non-linear prediction, context based statistical modelling and Golomb-Rice coding \[^{12,13}\]. In this paper we limit our discussion to run-length coding and non-linear prediction stages and propose the use of the remaining stages, unaltered.

The prediction and modelling stages of JPEG-LS are based on the causal template depicted in figure 1, where \(x\) denotes the current pixel, and \(a, b, c, d\) are reconstructed neighbouring pixels in the relative positions shown in figure 1.

![Figure 1. JPEG-LS predictive pattern](image)

The context that conditions the encoding of the current prediction residual \(x - \hat{x}\) (where \(\hat{x}\) is the predicted value of \(x\)) is built out of the differences \(\Delta_1 = d - b\), \(\Delta_2 = b - c\), and \(\Delta_3 = c - a\). If \(\Delta_1, \Delta_2, \Delta_3 < N\), the texture is assumed to be smooth and the run-length coding mode is used to code the residual. Here \(N\) is a predetermined information loss parameter represented by an integer constant \(NEAR\). When \(\Delta_1, \Delta_2, \Delta_3 < N\) is not satisfied the non-run length coding mode, which consists of predictive coding, quantization controlled by the \(NEAR\) parameter and entropy coding is used. Hence the above coding mode classification can be summarized as follows:

\[
coding \mod e = \begin{cases} 
\text{run-length} & \text{if } \Delta_1, \Delta_2, \Delta_3 < NEAR \\
\text{predictive-coding} & \text{otherwise}
\end{cases}
\]  

Once the coding mode is found to be predictive (i.e. non-run length coding), using eq. (1), a simple local texture analysis among the context pixels \(a, b\) and \(c\) is further used to explore the existence of vertical and/or horizontal edges.
and determine a best possible prediction value for the current pixel $x$. If an edge is detected the pixel that is not on the edge is taken as the predictive value. Otherwise the predictive value is assumed to be $\hat{x} = a + b - c$. For more detailed understanding we provide the pseudo-code of the above procedure as follows:

```c
if (c ≥ max (a, b))
    \[ \hat{x} = \min (a, b) \]
else
    if (c ≤ min (a, b))
        \[ \hat{x} = \max (a, b) \]
    else
        \[ \hat{x} = a + b - c \]
```

In JPEG-LS, the information loss that is introduced to all types of texture areas is the same. This is due to the fact that subsequent to the coding mode selection procedure discussed above, the same NEAR parameter is used for the quantization of all prediction residuals, $x - \hat{x}$. In order to solve this problem, in [17], the authors proposed the re-distributed of information loss in a manner more suitable for human perception, thus resulting in improved subjective image quality. The basic idea of this technique is the introduction of three different levels of information loss $(\text{NEAR}_L, \text{NEAR}_M, \text{NEAR}_H)$ corresponding to the three modes in which JPEG-LS operates, namely, the run-length coding mode, non-edge detected mode and the edge detected mode. The three levels, $\text{NEAR}_L, \text{NEAR}_M$ and $\text{NEAR}_H$ indicate low-level, medium-level and high-level information loss respectively and replaces the single level of information loss parameter $(\text{NEAR})$ of JPEG-LS. As areas that fall into the run-length coding mode indicate the smoothest texture, any excessive information loss in these areas would be highly intolerable for comfortable human perception. Thus $\text{NEAR}_L$, is used to quantize the prediction errors of such areas. In edge-detected areas that are capable of accommodating the highest level of information loss, the highest parameter, $\text{NEAR}_H$ is used as the quantizer. For the remaining areas, i.e. non-edge detected areas, the medium parameter, $\text{NEAR}_M$ is used. The above selection can be summarized as follows:

$$\text{NEAR} = \begin{cases} 
\text{NEAR}_L & \text{for mode selection & run-length coding} \\
\text{NEAR}_M & \text{for non-edge detected prediction} \\
\text{NEAR}_H & \text{for edge detected prediction}
\end{cases}$$

(2)

The above multi-level information loss re-distribution strategy is useful in applying rate-control to the original JPEG-LS algorithm. Two attempts [19,20] have been made by the authors of [17] to this effect. In the following section we summarize the design and performance of an improved rate control strategy proposed by us in [22]. Readers interested in the detail design details and an in depth performance analysis of this rate control scheme is referred to [22].

### 3. LOW COST RATE CONTROL

In the rate control strategy proposed in [22], the On-line (i.e. cumulative) Compression Ratio (OCR), is calculated at the end of each scan line and is controlled to follow the TCR as close as possible. The necessary change in compression is obtained by controlled changes in the three parameters, $\text{NEAR}_L$, $\text{NEAR}_M$ and $\text{NEAR}_H$. For a general scanline ‘i’ of the image, the step sizes used to change the values of the three parameters, $\Delta L$, $\Delta M$ and $\Delta H$ are given by:

$$\Delta N_i = \left[ \beta_{N_i} P_{N_i-1} \left[ \frac{\text{TCR}}{\alpha OCR_{i-1}} \right] \right]$$

(3)
where, $N$ represents one of the three contexts, represented by parameters, $\text{NEAR}_L$, $\text{NEAR}_M$ or $\text{NEAR}_H$ and
\[ \Delta N = \{\Delta L, \Delta M, \Delta H\}. \]
$\alpha$ is a constant that limits the step sizes, $\Delta N_i$ within reasonable limits ($<4$) when
\[ OCR_{i-1} << TCR. \]
$\beta_N > 1$ is a constant dependent on human visual perception that is used to make the step sizes associated with smooth areas (i.e. $\Delta L$) relatively smaller as compared to non-edge-detected ($\Delta M$) and edge-detected ($\Delta H$) areas, especially when $OCR_{i-1} << TCR$. Ideally, $\beta_L < \beta_M < \beta_H$. This is due to the fact that human eye would be sensitive to sudden large changes in smooth areas as compared to similar changes in edgy areas.

$P_{N_{i-1}} (0 \leq P_{N_{i-1}} \leq 1)$ represent the fraction of pixels belonging to the context $N$, within the previously coded line, $i-1$, and $OCR_{i-1}$ represents the corresponding OCR. The use of $P_{N_{i-1}}$ related to line $i-1$, in determining $\Delta N$, which in turn is used in coding the pixels belonging to line $i$ is justified due to the fact that it could be assumed that within the next line (i.e. line $i$) the fraction of pixels that belongs to each of the three contexts would be similar.

It should be also noted that in the above scheme, all three parameters are incremented or decremented together, provided they fall within their respective ranges and all three $\Delta N$ values calculated by eq. (3) are non-zero. In the event of a $\Delta N$ value being equal to zero it implies that either the TCR has been achieved (i.e. $OCR_{i-1} = TCR$) or $P_{N_{i-1}} = 0$, which implies that no pixels have been found in the previous line to fall within the context denoted by $N$. In the latter case one could assume that it is likely that in the subsequent line of pixels a similar distribution occurs and thus no steps are taken to make any unnecessary change in the corresponding information loss distribution parameter.

In addition to the above, under the proposed rate-control scheme, the parameters $\text{NEAR}_L$, $\text{NEAR}_M$, and $\text{NEAR}_H$ are allowed to change by magnitudes higher than 1 (up to 4 in our experiments), when $OCR_{i-1} << TCR$. This increases the speed at which a given TCR is achieved and thus improves the consistency of the subjective quality of the images, specifically at the top edge of the reconstructed images. It also aids in achieving higher TCR’s with increased accuracy as compared to the benchmark techniques.

In section 2 it was mentioned that in JPEG-LS, the prediction of pixel $x$ at a given location is based entirely on the pixels, $a$, $b$, $c$ and $d$, located within the predictive template illustrated in fig.1. These pixels are in the so-called ‘reconstructed’ domain, which implies that both in the encoder and decoder, the context modeling and prediction processes are identical. As a result of all improvements, variations and applications proposed in sections 2 and 3 being based on pixels situated within this predictive template, there is no additional need of having to transmit any of the dynamically changing parameters produced at the encoder side, to the decoder, for accurate reconstruction of the decoded image. Perfect reconstruction of these parameters is possible at the decoder.

In the following section we apply the above rate control strategy to ROI based image coding.

### 4. REGION-OF-INTEREST BASED CODING

An ROI is a part of an image that is considered more significant in terms of perception, usability and/or analysis point of view by a human or a machine. Thus an ROI is normally coded at a better quality as compared to the region outside it. The latest transform based still image compression standard, JPEG-2000 [21], has an extensive, flexible and efficient framework to support ROI based coding (Annex H of [21]). Within our current research context we only attempt to introduce ROI based coding to JPEG-LS to extend its usability to low cost ROI based coding. We do not propose our work here as an alternative or replacement to the ROI based coding with JPEG-2000. However we show that the proposed extension has additional features, which are not available directly from the ROI framework that is present in JPEG-2000.
In order to apply the rate-control procedure of \cite{22}, which is summarized in section 3, to ROI based coding, we consider the following cases:

1. ROI is coded losslessly and the overall image undergoes rate-controlled image compression.
2. Both ROI and non-ROI regions undergo rate-controlled image compression, but the loss introduced to the ROI is upper-bounded in order to maintain a good visual quality.
3. Both ROI and non-ROI regions undergo rate-controlled image compression, but the loss introduced to the ROI is fixed, regardless of local texture variations.
4. ROI is coded near-losslessly, with no rate control, but the overall image undergoes rate-controlled image compression.
5. ROI region is compressed with the idea of maintaining a good visual image quality and the non-ROI region is ‘blanked-out’ completely, therefore no rate-controlled compression of the overall image.

The special cases provided above are aimed at meeting the requirements of various application areas in digital imaging ranging from digital photography to advanced medical imaging. As an example, surgeons would be interested in visually inspecting a ROI of a medical image for diagnostic purposes, where the ROI needs to be coded in a manner that would preserve its visual quality to the highest possible level (Cases 1 & 2). The overall image would still need to be compressed to a specific file size for efficient storage within a database. For some diagnostic applications, ROI would need to be of sufficiently good quality, but the compressibility of the image would take precedence over subjective quality of the image (Cases 3 & 4). In all above examples it may be the case that the non-ROI region is also important to give a suitable context to the ROI, whereby their coding even at a lower quality is essential. However, there may be applications, such as in digital photographic filtering, where, the non-ROI is not necessary. For such applications, case 5 above would be an ideal set-up.

In order to implement the above cases we use a ROI template of equal dimensions to the image, consisting of value 1 for all spatial locations within the ROI and value 0 for all locations outside. This template was subsequently used in order to activate different combinations of algorithms for the ROI and non-ROI regions as described above.

5. EXPERIMENTAL RESULTS & ANALYSIS

To demonstrate the use of the rate control scheme of \cite{22} (summarized in section 3) in ROI based coding, as depicted in figure 2 (a), we extracted two arbitrary shaped ROIs, consisting of the upper part of the hat and the face of Lena image. This extraction was done manually as would be required in most practical applications involving human beings.

![Figure 2](image_url)

Figure 2. ROI coding of “Lena” image (a) ROI template with two arbitrary shaped regions (b) ROI coded image with no rate control, CR = 17.672:1, with ROI coded losslessly
Figure 2 (b) illustrates the Lena image, coded with the ROI specified in fig 2 (a) (white region) losslessly compressed and the non-ROI blanked out to contain a constant value (i.e. 0, black). The compression ratio achieved is 17.672:1. No rate-control is applied in this case. The high compression ratio achieved here is due to the fact that the major portion of the image contains a constant pixel value.

Figure 3. Visual inspection of ROI based coding results for “Lena” image, with the outside of ROI coded with L, M and H varying from 3 to 12, and the ROI coded with

(a) L=0, M=0, H=0 fixed, i.e. losslessly, TCR = 4:1, RCCR=3.989:1, PSNR = 34.107dB
(b) L_MAX : M_MAX : H_MAX = 2:9:12, TCR=8:1, RCCR=7.973:1, PSNR = 32.848 dB
(c) L_MAX : M_MAX : H_MAX = 5:5:5, TCR=8:1, RCCR=7.998:1, PSNR=33.225 dB
(d) L=6, M=6, H=6 fixed, i.e. no rate control within ROI, TCR=8:1, RCCR=7.994:1, PSNR=33.469 dB
In figure 3, we apply various combinations of coding styles to the ROI and non-ROI regions, as briefly discussed in section 4. In all these cases, we attempt rate controlled image compression by setting a realistic TCR that depends on the natural compressibility of the image.

Figure 3 (a) illustrates the Lena image compressed at a TCR of 4:1, with the ROI coded losslessly. A Rate Controlled Compression Ratio (RCCR) of 3.989:1 is achieved at a PSNR of 34.107 dB. As the ROI is coded losslessly, rate control is not applicable to this area and therefore under this condition the size of the ROI (area) would limit the maximum achievable TCR. For the ROI under consideration the maximum achievable TCR with an acceptable accuracy was 4:1.

Figure 3 (b) illustrates the Lena image compressed at a TCR of 8:1, with the ROI coded near losslessly, with rate control in this area aimed at maximizing the subjective image quality of the ROI (i.e. using the so-called operation mode-2 of rate control scheme of [22]). This is achieved by limiting the MAX_L (maximum value of NEAR_L allowed) to 2, in order to safeguard the subjective quality of the smooth, HVS-sensitive areas, within the ROI and setting the MAX_M and MAX_H parameters rather high (9 and 12 respectively) to aim at achieving a higher target compression ratio. A RCCR of 7.973:1 is achieved at a PSNR of 32.848 dB. Thus as compared to the case depicted in fig. 3 (a), fig. 3 (b) provides a visually equivalent image, but at almost twice the compression ratio.

Figure 3 (c) illustrates the Lena image compressed at a TCR of 8:1, with the ROI coded near losslessly, with rate control in this area set to operation mode-1 of the rate control scheme of [22], in which a single level information loss distribution controlled by single parameter, N (as in original JPEG-LS) is assumed. A slightly higher RCCR of 7.998:1 is achieved in addition to a clearly better objective image quality of 33.225 dB. The better objective quality is mainly due to the fact that in the edgy areas, the compression parameters are now limited by 5, as compared to 9 and 12 in the condition illustrated in fig. 3(b). However a close visual inspection proves that the ROI of fig. 3(b) is of better subjective quality as compared to the ROI of fig. 3 (c). A closer investigation of the quality on the cheeks of the figures proves this fact.

Figure 3(d) illustrates the Lena image compressed at a TCR of 8:1, with the ROI coded near losslessly, but with no rate control in this area. Here constant parameter of NEAR=6 is used in coding all texture areas within the ROI. Even though this achieved RCCR of 7.994:1 and PSNR of 33.469 dB, shows a competitive alternative to cases illustrated by fig. 3(b) and 3(c) at an equivalent TCR, the ROI shows easily visible subjective quality degradation.

Thus even though various combinations are possible in coding the ROI, under the proposed scheme, the one illustrated in figure 3(b) is recommended for use in application areas, where subjective image quality of the ROIs are significantly important over other parameters, such as accuracy of achieving the TCR and the objective quality measure, PSNR. By adapting a similar procedure one is easily able to modify the proposed ROI based coding technique so that multiple ROI could be coded under varying requirements/conditions, depending on their practical usage. For example in some areas of medical imaging and in some machine vision applications, edgy texture could be more important than the smooth texture and in such ROI's the parameters MAX_M and MAX_H could be set lower than MAX_L, to inverse the effects described above. This type of a flexible coding approach is not possible using the current JPEG-2000 standard.

6. CONCLUSIONS

In this paper we have proposed a variation of JPEG-LS to effective, low cost effective rate control. The proposed rate-control algorithm has the flexibility of operating under two operational modes, one aimed at obtaining the best possible objective rate-distortion performance and the second aimed at obtaining the best possible subjective image quality. We have provided extensive experimental results and analysis to compare the performance of the proposed algorithm with the state of the art near lossless rate control schemes. The results clearly indicate the proposed scheme’s superior performance as compared to these algorithms. The proposed rate control algorithm is able to achieve much higher TCRs, at increased accuracy and better objective image quality, using a less computationally intensive rate control strategy. We have finally applied the proposed rate control scheme to develop a flexible, interactive, ROI based coding scheme that could be effectively used in medical, satellite and photographic imaging application areas.

Proc. of SPIE Vol. 5022     649
REFERENCES


* Correspondence: E.A.Edirisinghe@lboro.ac.uk; Telephone: +44 (0)1509 228234; Fax: +44 (0)1509 211586