Improvements to JPEG-LS via diagonal edge-based prediction

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


Additional Information:

- Copyright 2002 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/10.1117/12.453103

Metadata Record: https://dspace.lboro.ac.uk/2134/6513

Version: Published

Publisher: © 2002 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/
ABSTRACT

JPEG-LS is the latest pixel based lossless to near lossless still image coding standard introduced by the Joint Photographic Experts Group (JPEG). In this standard simple localized edge detection techniques are used in order to determine the predictive value of each pixel. These edge detection techniques only detect horizontal and vertical edges and the corresponding predictors have only been optimized for the accurate prediction of pixels in the locality of horizontal and/or vertical edges. As a result JPEG-LS produces large prediction errors in the locality of diagonal edges. In this paper we propose a low complexity, low cost technique that accurately detects diagonal edges and predicts the value of pixels to be encoded based on the gradients available within the standard predictive template of JPEG-LS. We provide experimental results to show that the proposed technique outperforms JPEG-LS in terms of predicted mean squared error, by a margin of up to 8.51%.

Keywords: JPEG-LS, Lossless Image Coding, Predictive Coding, Edge Detection

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, lossless/near-lossless compression schemes such as DPCM, FELICS, LOCO-I, have by now evolved into an international standard, JPEG-LS, with the major contribution coming from LOCO-I (Low Complexity, Context-Based Lossless Image compression algorithm).

Lossless image compression schemes can be divided into two main components, namely modeling and coding. In the modeling part the pixels are coded in a raster scan order. If $x_i$ is the pixel under consideration at time $i$, it is referred to by assigning a conditional probability distribution of $p(x_i | x^{i-1})$ where $x^{i-1} = x_1, x_2, ..., x_{i-1}$. Thus the code length contributed by $x_i$ is given by $- \log_2 p(x_i | x^{i-1})$. Ideally this value averages to the entropy of the probabilistic model being used. Usually the above probability assignment is divided into three further steps. In the first step, prediction, a value for $x_i$ is guessed based on a subset of the past sequence $x_{i-1}$. In reality this subset is chosen from the immediate neighbourhood of already coded pixels. The second step is context modeling, which is the determination of the context (smooth, textured, edgy etc.) in which $x_i$ occurs. This is also based on the past sequence of pixels, mostly pixels in the immediate neighbourhood of $x_i$. Finally a probabilistic model is produced for the prediction error, $e_i \Delta(x_i - \hat{x_i})$, based on the context of $x_i$. The fact that the prediction error $e_i$ is dependant on context modeling clearly signifies the importance of accurate context determination.

* Correspondence: Email: E.A. Edirisinghe@ljboro.ac.uk Tel: +44 (0)1509 228234 Fax: +44 (0)1509 211586
LOCO-I\textsuperscript{2}, the compression algorithm that mostly contributed to the JPEG-LS standard, follows the above-mentioned traditional structure, namely, predictor-modeler-coder. However, in the context modeling section of the JPEG-LS algorithm, only horizontal and vertical edges, in the vicinity of a pixel $x_i$, are considered in the determination of its context. Unfortunately this means that in the presence of a diagonal edge, its erroneous detection as a horizontal or a vertical edge would result in a wrong context classification for the pixel $x_j$. This would result in a higher prediction error, $e_j$, and an ultimate reduction of the rate-distortion performance of the JPEG-LS algorithm.

Even though one could argue that the likelihood of a diagonal edge is minimal as compared to a horizontal or a vertical edge and that the errors resulting from the wrong context classifications could be neglected in exchange of a lower complexity context modeling and prediction approach, existing research\textsuperscript{14} indicates that this is not always the case. It has been shown that, by considering the relative gradients present within the so-called predictive template of JPEG-LS (Fig.1) it is possible to detect these diagonal edges and establish ways to accurately predict pixels in the vicinity of such edges. J.Jiang et. al.\textsuperscript{14}, proposed a method that is based on six thresholds to detect diagonal edges and perform accurate prediction. However the use of a large number of thresholds in context determination increases the complexity of the approach and also makes it rather difficult to determine the best set of six parameters that could be applied for the efficient coding of a wider range of images. In addition to this, experiments carried out by us showed that certain constraints that have been used in this work are ‘too tight’ and would result in the elimination of certain diagonal edges from the selection process. In particular it is in fact possible to more accurately predict near single pixel width diagonals. In order to solve the above issues we propose a low complexity approach that is based on two, easily deterministic thresholds (parameters). The flexibility/simplicity and the increased accuracy of the proposed algorithm in the context determination process and in prediction, is shown to clearly outperform the performance of the previously published techniques\textsuperscript{14}.

For clarity of presentation this paper is divided into five main sections. Apart from section 1, which is an introduction to the research context discussed in this paper, section 2 introduces the reader to the details of the prediction scheme used by JPEG-LS/NLS standard and identifies associated problems using well-supported examples. Section 3 describes the proposed solution, which is based on the accurate detection of diagonal edges and accurate prediction of pixels in the vicinity of such edges. Section 4 provides the experimental results in support of the proposed modification and a detailed analysis of the results. Finally section 5 concludes with suggestions for future work and improvements.

\section*{2. JPEG-LS CONTEXT CLASSIFICATION \& PREDICTION SCHEME}

In the following discussion we limit our analysis \& modifications to the lossless-mode operation of JPEG-LS. However, the ideas presented below could be easily extended to the near-lossless mode of operation.

\subsection*{2.1 Context Classification}

The context classification process of JPEG-LS is based on the simple causal template depicted in fig. 1. Here, $x$ is the pixel to be predicted and $a$, $b$, $c$ and $d$ are previously encoded pixels in the immediate neighbourhood of `$x$'. For the purpose of simplicity, we have ignored the time index $i$ from the above notation.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{template.png}
\caption{JPEG-LS Predictive Template}
\end{figure}
According to the JPEG-LS context classification procedure, pixel 'x' is defined to be in a smooth region if all four neighbouring pixel values, a, b, c and d are equal. Under such conditions, run-length coding \(^{11,12}\) is applied in encoding the pixel values. If this condition is not satisfied the context is classified as non-smooth and simple edge detection based predictive coding is used in coding 'x'.

2.2 Predictive Coding

For non-smooth areas, JPEG-LS uses the following simple edge detection/prediction scheme, to determine the predictive value of each pixel, x.

\[
\begin{align*}
\text{if } (c \geq \max(a, b)) & \quad \text{/* Edge */} \\
\hat{x} &= \min(a, b) \\
\text{else} & \quad \text{if } (c \leq \min(a, b)) \quad \text{/* Edge */} \\
\hat{x} &= \max(a, b) \\
\text{else} & \quad \text{/* Non _ Edge */} \\
\hat{x} &= a + b - c;
\end{align*}
\]

where, \(\max(.,.)\) and \(\min(.,.)\), respectively denotes the maximum and minimum valued elements from the set \((.,.)\) and \(\hat{x}\) is the predicted value of pixel x. Further analysis of the above pseudo-code reveals the following facts:

An edge would be detected among the three pixels a, b and c, when either \(c \geq \max(a, b)\) or \(c \leq \min(a, b)\) is satisfied. When \(c \geq \max(a, b)\), the presence of a horizontal edge is signified by \(\max(a, b) = a\) and the presence of a vertical edge is signified by \(\max(a, b) = b\). Similarly when \(c \leq \min(a, b)\) the presence of a horizontal edge is signified by \(\min(a, b) = a\) and a vertical edge is signified by \(\min(a, b) = b\).

![Figure 2. Edge detection for (a) c ≥ max(a, b) ⇒ \(\hat{x} = \min(a, b)\) (b) c ≤ min(a, b) ⇒ \(\hat{x} = \max(a, b)\),](image-url)
Figures 2(a) and 2(b) show such an edge detection process. [Note: The highlighted pixels values correspond to higher valued pixels.] Careful analysis of these figures show that a logical selection of the predictive value for pixel $x$ would be the pixel that is on the same side of the detected edge. Thus under the above conditions, the predictors proposed by JPEG-LS standard are clearly justified, under the assumption that the edges involved are either vertical or horizontal.

When $\min(a, b) < c < \max(a, b)$, there is not enough justification to determine whether there is an edge or not. However contrary to the observations in [14], it is possible to show that the pixel value $x$ is not essential in the determination of the context. We observe that under this condition a vertical/horizontal edge may be still detected if the value of $c$ is sufficiently close enough to either $\min(a, b)$ or $\max(a, b)$ and the difference between $a$ and $b$ is sufficiently large. [we address this issue in more detail in section 3]. However the important fact here is that the predictor $\hat{x} = a + b - c$, used by JPEG-LS under this condition, is very well balanced to accurately predict $x$, regardless of the fact whether the context is smooth or contains a vertical/horizontal edge. This is easily proved by rearranging the terms of $\hat{x} = a + b - c$ and further analyzing the results.

One could re-arrange the terms of $\hat{x} = a + b - c$ as:

$$\hat{x} - a = b - c$$

(1)

$$\hat{x} - b = a - c$$

(2)

Both equations above depict pixel gradient equalizations across boundaries (or more precisely, vertical/horizontal edges). When eq.(1) is satisfied one could expect a horizontal edge to be present and when eq.(2) is satisfied one could expect a vertical edge to be present. Both equations support the presence of a smooth area, as such an area could be characterized by small differences between adjacent pixel values, $a$, $b$ and $c$. In other words if $a \equiv c \Rightarrow \hat{x} \equiv b$ and if $b \equiv c \Rightarrow \hat{x} \equiv a$, and either way since $a \equiv b$, $\hat{x}$ would be approximately the same.

The above discussion clearly proves the excellent performance of the predictor used by JPEG-LS under the condition $\min(a, b) < c < \max(a, b)$.

Unfortunately there is one more type of edge that is possible: i.e. a diagonal edge, which has been ignored by the JPEG-LS algorithm. In support of the idea of disregarding the presence of diagonal edges one may claim that a combination of many vertical and horizontal edges can produce a line with arbitrary shape or edges in an image. However, what matters here is whether the description by the two edges (horizontal/vertical) will be accurate enough to minimize the predictive errors. In cases where a diagonal edge exists inside the predictive template, it can be expected that the errors produced can be unnecessarily high.

Thus, in order to find a suitable low-cost solution to the above problem, we propose the introduction of simple diagonal edge detection based prediction, for those pixel locations that are in the vicinity of diagonal edges. The presence of three types of edges is expected to increase the prediction accuracy of arbitrary shaped object boundaries/edges, leading to improvements in rate-distortion performance of JPEG-LS image compression standard.
3. DIAGONAL EDGE DETECTION & PREDICTION

3.1 Diagonal Edges & Their Detection

Figures 3(a), (b), (c) and (d) show the four main orientations of diagonal edges that may be considered for further improvement of the JPEG-LS prediction scheme. The highlighted rectangles represent the opposite type (high/low) of pixel intensity to that of the non-highlighted rectangles. A closer investigation of these patterns shows that it is not possible to accurately identify the diagonal edge orientations indicated in figs. 3(a) and 3(b), by means of low cost techniques limited to the standard predictive template of JPEG-LS (fig.1). The reason for this is that the value of the pixel to be predicted $x$ cannot be involved in the decision making process and due to this fact it can not be guaranteed whether the edge is actually a diagonal edge or rather a horizontal (fig. 3(a)) / vertical edge (fig. 3(b)).

However, in the case of orientations illustrated in figs. 3(c) and 3(d) it is possible to accurately establish the presence of a diagonal edge. Figure 3(c) illustrates a diagonal edge that separates two areas of differing intensity and fig. 3(d) illustrates a diagonal line of single pixel width. A close comparison of the pixel intensities of figs. 3(c) and 3(d) indicates that a diagonal edge is characterized by the fact that $a \equiv b$ and pixel $c$ being greatly different from both $a$ and $b$. Interestingly we find that these conditions would not be worthwhile testing under the condition, $\min(a,b) < c < \max(a,b)$. This is due to the fact that under this condition when $a \equiv b \Rightarrow c \equiv a$ and $c \equiv b$. This actually depicts a smooth texture rather than any kind of edge as discussed previously. Thus our search for diagonal edges should happen when the condition $c \geq \max(a,b)$ or the condition $c \leq \min(a,b)$ is satisfied.

By considering the above, we now formulate the formal conditions for the accurate detection of diagonal edges of the nature illustrated in figs. 3(c) and 3(d), as follows:

$$((c - \max(a,b)) > T_1 \quad \text{OR} \quad (\min(a,b) - c) > T_1) \quad \text{AND} \quad (\abs(a-b) \leq T_2)$$

In eq. (3), OR and AND represent the standard Boolean operations, whereas abs represents the absolute value of the element quoted within parenthesis. $T_1$ and $T_2$ are predefined positive thresholds (ideally $T_1 > T_2$) that could be used in deciding the contrast of pixel values present across the detected diagonal edges. Further analysis of the above gradients reveal that it is the conditions that are opposite to what has been stated above, that would be satisfied when a horizontal or a vertical edge is present. Thus it provides an accurate means of differentiating diagonal edges from vertical/horizontal edges under the condition $c \geq \max(a,b)$ or $c \leq \min(a,b)$.

Accurate detection of diagonal edges only amounts to a more accurate context classification and is not the only factor that results in the ultimate aim of achieving better prediction accuracy. In the following section we introduce the reader to a modified prediction strategy, which increases accuracy of prediction near diagonal edges.
3.2 Predicting

In order to arrive at a suitable predictor for the pixel $x$, in the presence of a diagonal edge of the nature illustrated in figs. 3(c) and/or 3(d), we propose to equalize the gradients, $(a - d)$ and $(b - \hat{x})$. Thus,

$$
(a - d) = (b - \hat{x})
$$

$$
\Rightarrow \quad \hat{x} = b + d - a
$$

(4)

The above gradient equalization could be justified due to two reasons. Firstly, it is highly unlikely for these gradients to have opposite signs between two consecutive scan-lines. Secondly, it is highly likely that the pixel values change in an equivalent way between two scan-lines when moving away to the right, from the diagonal edge. An interesting observation here is that the above predictor is well suited to cater for the needs of diagonal edges with width of a single pixel (fig. 3(d)), unlike the predictor proposed in [14]. In addition to the above observations the proposed predictor is also able to equalize gradients in a direction parallel to the detected diagonal edge. This is due to the fact that $(a - d) = (b - \hat{x}) \Rightarrow (a - b) = (d - \hat{x})$. Thus it is expected that eq.4 provides a good predictor in the presence of diagonal edges.

[Note: In order to keep the predicted pixel value $\hat{x}$ within the range $0 \leq \hat{x} \leq 255$, if we find that $\hat{x} < 0$ we make $\hat{x} = 0$, and if we find that $\hat{x} > 255$ we make $\hat{x} = 255$.]

Considering the diagonal edge detection and prediction strategies proposed above, the JPEG-LS pseudo-code for edge detection/prediction could be modified as bellow:

```plaintext
if (c ≥ max (a, b))
    if ((c - max (a, b)) > T1 AND (abs (a - b) ≤ T2))
        \hat{x} = b + d - a;
    else
        \hat{x} = min (a, b);
    
else
    if (c ≤ min (a, b))
        if ((min (a, b) - c) > T1 AND (abs (a - b) ≤ T2))
            \hat{x} = b + d - a;
        else
            \hat{x} = max (a, b);
    
else
    \hat{x} = a + b - c;
endif
```

Note that according to the above pseudo-code the diagonal edges are detected from among edges that would otherwise be categorized as vertical/horizontal edges and not from areas, which would otherwise be considered as non-edge detected. The reasons for this have already been discussed.
In order to check the validity of proposed scheme, we performed experiments on the standard test image, Test8 (fig. 4), which is widely used within the image compression research community as a test image. This image has a special property in that the four constituent sub-images it represents, consist of images with widely different texture / statistics, being from natural, synthetic, text and dot-matrix form. Thus any performance results of an algorithm obtained based on the testing on this image could easily be generalized.

The performance is measured in terms of the Mean Squared Error of the Predicted image (PMSE), which is defined as follows:

\[ PMSE = \frac{1}{N} \sum_{i=1}^{N} (x_i - \hat{x}_i)^2 \]  

where, \( N \) is the total number of pixels in the image and \( x_i \) and \( \hat{x}_i \) are respectively the original and predicted value of the \( i \) th pixel. Thus, better prediction would lead to a lower value of the above performance measure.

To evaluate the performance of the proposed technique, we fix the threshold \( T_2 = 10 \) and vary the threshold \( T_1 \) between 40-255. The PMSE graph obtained for the proposed scheme is shown fig. 5. The PMSE value (1075.535) obtained when using the standard JPEG-LS prediction scheme is illustrated as a straight line.
The graph in fig. 5 clearly shows that the proposed prediction algorithm, outperforms the original JPEG-LS prediction algorithm in terms of the mean squared error of the predicted image, within the full range of the threshold $T_1$, particularly performing best when $T_1 = 70$. At this setting the total improvement of the predicted mean squared error is 8.51%. Further analysis revealed that 1147 pixels (1.75% of the total) were classified as being on a diagonal edge at this setting. Given the fact that an improvement of 8.51% in terms of the total predicted mean squared error was obtained in an image where only 1.75% are categorized as diagonal edges (at the optimum parameter setting) proves that better results could have been obtained in images where there are more diagonally oriented edges. When $T_1 = 255$, none of the pixels are classified as diagonal pixels and thus the performance of the proposed coder converges to that of JPEG-LS. Generally speaking, the performance of the proposed coder approaches the performance of the JPEG-LS coder with increasing $T_1$. This is because, larger the value of $T_1$, lesser and lesser amounts of edges would be classified as diagonal edges, as tighter conditions apply in the classification of an edge as a diagonal. Similarly, a decrease in the threshold $T_2$ would also result in lesser amount of edges being categorized as diagonals. Our experiments showed that for a large set of test images, the best settings for the two threshold values would be $T_1 = 60$ and $T_2 = 8$.

A more detailed performance assessment of the proposed algorithm indicated that in the four constituent sub-image sections of the test image (Test8), the best performance is in fact in the bottom right hand corner sub image, which represents a dot matrix image. The reason for this is that this sub-image consists of a large amount of single pixel width diagonal edges, which are very accurately detected and predicted by the proposed scheme. In such an area, the prediction techniques adopted by JPEG-LS and the method of [14] would perform poorly. Another advantage of the proposed scheme over the prediction scheme proposed in [14] is the lower complexity of prediction. Given the fact that the diagonal edge prediction with the proposed scheme only needs one addition and one subtraction, as against one addition and one division (a division is approximately six times computationally expensive as compared to a subtraction) proves the validity of this claim.
An interesting aspect, which can be observed from the results of the above experiments is that, although the proposed algorithm improves the accuracy of prediction, measured in terms of MSE values, the compression ratios stay the same (1.915:1). This is expected due to the fact that the entropy coding length is determined by statistical modelling and yet this part of operation has not been revised in the light of the proposed diagonal edge detection. In this sense, smaller predictive errors may not necessarily produce higher compression efficiency. The decisive factor for compression efficiency is how accurate could the statistical modelling produce statistical information to drive the entropy coding. However, when predictive errors are minimized, the direct and positive effect upon statistical modelling would be significant, since the statistical distribution of those errors would become more focused around its mean value as opposed to being scattered. The specific advantage with JPEG-LS is that the number of context quantization regions could be reduced and more probabilities could be assigned around the center of the statistical distribution. Another impact upon compression efficiency by smaller predictive errors can be seen with the near lossless compression mode in JPEG-LS. The near lossless mode requires a small quantization of those predictive errors in order to produce higher compression ratios. To this end, the reconstructed image quality could benefit by introducing smaller quantization steps corresponding to the smaller predictive errors.

5. CONCLUSIONS

In this paper we have proposed an improvement to the JPEG-LS prediction algorithm by introducing accurate diagonal edge detection and a modification to the prediction scheme. We have shown that diagonal edges with certain orientation could be accurately detected using the pixels within the JPEG-LS predictive template and that the resulting context knowledge could be used in the accurate prediction of pixels in the vicinity of such edges. It has been shown that the proposed prediction scheme outperforms JPEG-LS by margins of up to 8.51% in terms of the mean squared error of the predicted image. More importantly we show that these improvements to the prediction have been achieved with no additional computational cost. Currently further research is underway to modify the context modelling part of the JPEG-LS scheme so that the proposed modification could also result in an overall gain in compression performance.

REFERENCES