ENHANCEMENT OF JPEG CODED IMAGES BY ADAPTIVE SPATIAL FILTERING

Amlan Kundu

US WEST Advanced Tech.,
Boulder, CO 80303
email: akundu@advtech.uswest.com

ABSTRACT

The JPEG coder has proven to be extremely useful in coding image data. For low bit-rate image coding (0.75 bit or less per pixel), however, the block effect becomes very annoying. The edges also display 'wave-like' appearance. In this paper, an enhancement algorithm is proposed to enhance the subjective quality of the reconstructed images. First, the pixels of the coded image are classified into three broad categories [1]: a) pixels belonging to Quasi-constant regions where the pixel intensity values vary slowly, b) pixels belonging to dominant-edge (DE) regions which are characterized by few sharp and dominant edges and c) pixels belonging to textured regions which are characterized by many small edges and thin-line signals. An adaptive mixture of some well-known spatial filters which uses the pixel labeling information for its adaptation is used as the adaptive optimal spatial filter for image enhancement. Some experimental results are also provided to demonstrate the success of the proposed scheme.

1. INTRODUCTION

Recently, a number of schemes for post-processing of still and video coded images are proposed [1, 2, 3]. The need for such post-processing arises from the annoying visual artifacts present in low bit-rate image coding. For instance, in low bit-rate JPEG still coding, the block effect over the quasi-constant regions in the image are very annoying. The edges, on the other hand, display 'wave-like' appearance. In low bit-rate vector quantized images, similar block effects are seen while the edges look jagged after reconstruction [4]. The block effects are also prominent in low bit-rate fractal coding. The post-processing algorithms are essentially smoothing techniques which attempt to remove these artifacts for better visual perception. Since there is no well-defined mathematical representation for visual perception, the post-processing algorithms are best designed by proper understanding of the ad-hoc nature of the problem. For the problem at hand, i.e., artifacts present in low bit-rate JPEG coding, we need considerable smoothing to remove block effects. On the other hand, the block effects are somewhat muted over the textural regions. The edges require special filtering that preserves the edges. Besides, the degradation of edges are spatially somewhat broadened after reconstruction. Thus, we need to treat the edgels, texels and the pixels belonging to quasi-constant regions differently. In the scheme described in this paper, the pixels in the coded image are labeled with three distinct labels. This pixel labeling is then used to adapt different type of spatial filtering schemes to different type of pixels.

2. THREE-WAY PIXEL LABELING

Our objective is to classify each pixel as a member of one of the three kind of regions – the DE regions, the textured regions, and the QC regions – but not to explicitly assemble the pixels into regions.

Edgel Detection: First, the edgels are detected. Instead of well-known but computationally expensive operators, we have used the edge operator described in [5] as this operator can be easily tuned to detect the strong edges with very little computation. The edge operator is defined by

$$edge\_map \;=\; H \circ T \circ R \; (input\_image). \qquad (1)$$

The operator R, for each sliding window, computes a measure of the range of pixels inside the window. If this range is bigger than a threshold τ_1, the central pixel of the window is considered likely to be an edgel. The operator T then divides the pixels centered around this likely edgel into two groups (groups of higher intensity and lower intensity values). This stage is controlled by another threshold τ_2. Let N be the number of pixels belonging to the group of higher intensity values. Under some mild assumptions, it can be shown that N follows standard normal distribution. The operator H uses a hypotheses testing paradigm to make the final determination whether or not the central pixel is an edgel [5].

Texel Detection: The pixels belonging to the QC and the textural regions are separated from each other using a simple measurement - the number of zero-crossings - and a thresholding operation. The number of zero crossings are calculated as follows: In an nxn sliding window, the mean is calculated over all the pixels and then subtracted from each pixel. Let the intensity value of such a pixel be T_i. As a result, some elements will be positive, and some elements will be negative. The change of sign in the intensity values of two consecutive pixels in any particular direction is defined as a zero-crossing in that direction. The number of zero-crossings (NZC) are computed along each row, along each column and along the two diagonal directions. Both the QC and the textural regions have a rather large value for the NZC parameter. So, the objective is to use the NZC

0-8186-7310-9/95 $4.00

parameter along with some LSD (local standard deviation) measure to discriminate between the QC and the textural regions. In our work, a simple parameter reflecting LSD is proposed. After the subtraction of local mean, if the absolute value of any pixel inside the window is less than a preselected threshold value (β), that pixel value is set to zero. Otherwise, the pixel value remains unchanged. The NZC parameter is recomputed after this thresholding operation. In a QC region, this new NZC value is much less than its original NZC value if β is properly chosen. In a textural region, on the other hand, this NZC value is only slightly less than its original value. Consequently, this new recomputed NZC (RNZC) parameter could discriminate between the QC and the textural regions.

This formalization for finding the texels is similar to a-level crossing problem in stochastic process theory. Since the zero crossings, or rather the β-crossings, are counted along 1-D, we postulate that the intensity values T_i along 1-D is the realization of a stochastic process where each random variable has a zero mean. We further assume that the process is normal and differentiable. Then the level-crossing density (for level a) is given by

$$\lambda_a = 1/\pi\sqrt{(-R''(0)/R(0))}e^{-a^2/2R(0)} \tag{2}$$

where $R(\tau)$ is the autocorrelation of the process. Our objective is to find a (or β) such that, under mild assumptions, the number of level crossings for textural regions is guaranteed to be much higher than that of quasi-constant regions. Let $R_1(\tau)$ be the autocorrelation of the process when the region under consideration is a QC region. Similarly, let $R_2(\tau)$ be the autocorrelation of the process when the region under consideration is a textural region. We assume that $R_1(\tau)$ has the form $c_1 f_1(\tau)$ where $f_1(0) = 1$, $f_1(\infty) = 0$ and $f_1\prime(\tau) = -ve$. We further assume that $R_2(\tau)$ has the form $c_2 f_2(\tau)\cos(w\tau)$ where $f_2(0) = 1$, $f_2(\infty) = 0$ and $f_2\prime(\tau) = -ve$. To be valid autocorrelation functions, $f_1(.)$ and $f_2(.)$ approach zero with $|\tau| \to \infty$. The periodic component of $R_2(\tau)$ is a simplistic assumption but understates the textural characteristics quite well because the autocorrelation in a textural region does not fall off monotonically as it does in a QC region. Let us define

$$\Lambda = \frac{\lambda_a|texture}{\lambda_a|QC\ region} \tag{3}$$

It can be shown that $a - level$ crossings of the textural region will be much more when the following constraint is satisfied.

$$\beta^2 = a^2 > 3c_1/2 \tag{4}$$

Here, c_1 is essentially the variance of the QC region. If we assume that the variance of QC regions never exceeds 4% of the dynamic range, then for the dynamic range of 0-255, β is approximately found to be 12. To summarize, the pixel classification algorithm is executed as follows:

- The edge operator is first used to locate the dominant edges.
- If the $RNZC$ is greater than η (see Eqn. (5)) at any of the **remaining** pixel positions, that pixel is defined as a texel. Otherwise, the pixel belongs to a QC region.

η is calculated from the following consideration. Over a QC region, each pixel value, by assumption, follows a Gaussian distribution $N(0, \sigma)$. Since β is approximately equal to 1.22σ, the probability that the pixel is not set to zero is approximately 0.125. For each nonzero pixel, the maximum contribution to $RNZC$ is 4 as this pixel could be counted 4 times as part four directional countings (horizontal, vertical and two diagonal). So, a good first approximation value of η is

$$\eta = n^2 \cdot 4 \cdot 0.125 \tag{5}$$

Here, n^2 is the window size. For 5x5 windows used in our experiments, an η in the range $(10 - 12)$ is adequate.

3. OPTIMAL ADAPTIVE SPATIAL FILTER FOR ENHANCEMENT

Over the QC regions, the Hodges-Lehman D filter is used for smoothing the reconstructed image as this filter is known to be very efficient [6] in smoothing out noise with a short-tailed distribution. The coder noise over QC blocks usually appears as a short-tailed distribution.

3.1. Hodges-Lehman D Filter:

Let X_i $1 \le i \le n$ be a sample from a population with distribution $F(x, \theta)$ and density $f(x, \theta)$ where f (.) is symmetric about zero, continuous and strictly positive on the convex support of $F, [x : 0 < F(x) \le 1]$. Denote $X_{(1)}, .., X_{(n)}$ as the order statistics of the sample. Let n=2m or 2m-1. In either case, we define the D filter output as:

$$D_n(x_1, .., x_n) = median_{\ 1 \le i \le m}\ (X_{(i)} + X_{(n-i+1)}) \cdot 1/2. \tag{6}$$

Thresholded D filter: First, the pixels with intensity values in the range $(c - Range)$ and $(c + Range)$ are selected. Then, the D filter is applied only to the pixels within this range. Usually, c is the intensity value of the central pixel in the window.

Over a DE region, the noise profile appears to be a long-tailed distribution. The median filter is known for its efficiency in smoothing out such noise distribution. Also, **the median filter preserves the edges** . Our experimental findings have shown that, for regions around the dominant edges, a 5x5 or a 7x7 median filter is generally required for smoothing. Another filter we have considered is the Multistage Median Filtering. Consider a nxn window with four 1-pixel wide subwindows along horizontal, vertical and two diagonal directions. Let these windows be designated as $W_1, .., W_4$. Let z_i be equal to $median$ ($all\ pixels\ in\ W_i$). The maximum and the minimum of these four median output are computed as $y_{(max)}(n)$ and $y_{(min)}(n)$. Let $a(n)$ be the central pixel. Then the output of the multistage median filter is defined as:

$$output = median\ (\ y_{(max)}(n),\ y_{(min)}(n),\ a(n)\ ) \tag{7}$$

Next, we describe the most effective filtering combinations as determined by the experiments.

3.2. Adaptive Filter Combination

1. A 5x5 or 7x7 median filtering is used to smooth the edge points where the preceding and the succeeding points (in a raster scan sense) of any edge point are also considered as edge points. This 'flattening' of edge can be easily realized by a morphological dilation operation using a horizontal line element with three elements. The filtering is followed by one or two passes of 3x3 D filtering over the QC regions. Filtering over the textural regions shows no visual improvement. It should be noted that the texels could easily appear as 'too much smoothed'.
2. One pass of 3x3 D filtering over the QC regions is followed by 5x5 multi-stage median filtering to smooth the edge points where the preceding and the succeeding points of any edge point are also considered as edge points. The filtering is followed by one pass of 3x3 thresholded D filtering over the entire image.

For color images, the individual R, G and B components are filtered using the schemes outlined above. The 3-way pixel classification information is obtained from the 3-way pixel classification of Y image where the greyscale image Y is defined as $Y = 0.2999 \cdot R + 0.587 \cdot G + 0.114 \cdot B$. The filtering on Y, U and V images (as opposed to R, G, B images) show disturbing color variations. Such observation is previously made in [3].

4. VISUAL COMPARISON TEST AND EXPERIMENTS

Since no universally accepted quantitative criterion exists to express the visual quality of images, a subjective visual comparison test has been designed. A pair of images are displayed on the screen side by side. The two images are created by two different filtering schemes (one image could be the original coded image), and they are combined in a random fashion. The subjects, i.e. human observers, are instructed to choose one image over the other. If this experiment is carried out for a sufficiently large number of human subjects, the image with the better visual quality is likely to be chosen more often than the other; and the experiment is very likely to reveal which filtering schemes provide the maximum improvement in visual quality.

The segmentation based enhancement algorithm is tried on one coded image (0.625 bit/pixel) as depicted in Fig. 1(a). Figures 1(b) and 1(c) show the enhanced images using the filtering schemes **(2) and (1)** as described before. Figure 1(d) shows 3-way pixel classification of Fig. 1(a). By means of visual comparison test as described above, it is corroborated that the enhancement schemes do improve the visual quality of the images. This experiment is done on a few other greyscale images with similar observations. The visual improvement, particularly in removing blocking effects, lends itself to some form of quantitative representation as shown below.

4.1. M_1 Measure of Difference Image

Consider the difference image obtained by subtracting the original image from the coded or the post-processed image. Let the difference signal image be $s(i,j)$ over the QC regions and the local average of $s(i,j)$ be $r(i,j)$. Indicate the set of (i,j) which belongs to the QC regions by R. The measures, M_1 is now defined as follows:

$$M_1 = \frac{\sum_{(i,j)\in R} |r(i,j) - s(i,j)|}{N} \tag{8}$$

where N is the number of points in R. M_1 measures the average absolute deviation of the intensity of the difference or error image. For ideal filter performance, M_1 should be zero, i.e., the difference image is constant. It is relevant to note here that the addition of a constant to its pixel values does not impair the visual quality of the image in any noticeable way as long as the constant is small compared to the dynamic range of the pixels. For the coded image (Fig. 1(a)), M_1 is found to be 1.44 while the post-processed images using both schemes (1) and (2) yielded M_1 of 1.32. Similar improvement in M_1 is found in other greyscale coded and post-processed images. This measure, however, is not meaningful for the texels and the edgels.

For color images with bright colors, the blocking effect and the edge degradation are not often visible until the bit rate is very low (1:30 or even less). The bright colors tend to conceal the coding artifacts. The same filtering scheme is found effective on very low bit-rate JPEG coding of color images, but filtering is done separately on individual R, G and B frames as mentioned before.

As JPEG coding is gaining wide popularity, this enhancement technique will find applications as a value-added technique particularly for low bit-rate coding situations. Some preliminary experiments are also performed on low bit-rate fractal coded greyscale images using similar filtering schemes. Once again, distinct visual improvement is noticed in reducing blocking effects.

5. REFERENCES

[1] R.A. Gopinath *et al.* , "Wavelet based post-processing of low bit rate transform coded images", Proc. of ICIP, Austin, Texas, pp. 913-917, 1994.

[2] T. -S. Liu and N. S. Jayant, "Adaptive post-processing algorithms for low bit rate video signals", Proc. of ICASSP '94, Adelaide, Australia, pp. V-401 – V-404.

[3] V. Ramamoorthy, "Removal of staircase effects in coarsely quantized video sequences", Proc. of ICASSP '92, San Francsisco, CA, pp. III-309 – III-312, 1992.

[4] "Enhancement of Vector Quantized Images by Adaptive Nonlinear Filtering," patent awarded (#5218649) to US WEST with A. Kundu, V. Ramamoorthy and M. Terry as the authors.

[5] A. Kundu and S. K. Mitra, "Image Edge Extraction Using A Statistical Classifier Approach," IEEE Trans. on PAMI, PAMI-9, 4, 569-577, 1987.

[6] P. J. Bickel and J. L. Hodges, "The Asymptotic Theory of Galton's Test and a Related Simple Estimate of Location," Annals of Mathematical Statistics, 38, pp. 73-89, 1967.


(a)


(b)

Figure 1: (a) Original Image. (b) Enhanced image using scheme (2).


(c)


(d)

Figure 1: (c) Enhanced image using scheme (1).
(d) 3-way pixel labeling of Fig. 1(a).