Interscale Predictive Wavelet Coding with Huber Markov Random Field

Jin Li
Sharp Laboratories of America, Digital Video Division
5750 NW Pacific Rim Blvd., Camas, WA 98607, USA.
Tel. (360) 817-8486, Fax. (360) 817-8436, Email: lijin@sharplabs.com

ABSTRACT: *An interscale predictive wavelet coding (IPWC) scheme is proposed in this paper. The image is encoded top-down, from the coarsest scale to the finest scale. At each scale, the current scale coefficients are predicted from the coded coefficients at the coarser scale and the prediction residue is encoded. The interscale prediction is based on MAP estimation with Huber Markov random field statistical image model. The efficiency of IPWC is supported by experimental results.*

Keywords: wavelet coding, interscale prediction, MAP, Markov random field

1. INTRODUCTION

Wavelet transform provides an ideal tool for multiscale image analysis. An interesting theoretical problem in wavelet analysis is the image characteristics across wavelet scales, which has important applications in wavelet sampling and extrapolation, resolution enhancement technique and high efficient coding algorithms, etc.. Computer vision researchers, such as Mallet [1] and Berman [2] pioneered the cross-scale image characteristic research. In their work, the characteristics of edge across scales were investigated and multiresolution edge detection algorithm in the wavelet domain was proposed. Researches on the relationship between fractal and wavelet [3] [4] revealed that fractal could be considered as an interscale prediction that used the coefficients at coarse scale to predict those of the finer scale. Since fractal was well known to be efficient in exploring the self-similarity inherited in the textures, the result suggested that there were cross scale correlation in the texture regions of the image. Several state-of-the-art image coding algorithms also utilized the cross scale characteristics of wavelet for high efficient compression. For example, the scheme of embedded zerotree wavelet (EZW) coding proposed by Shapiro [5] and the scheme of set partitioning in hierarchical trees (SIPHT) proposed by Said and Pearlman [6] investigated the correlation of insignificant coefficients across wavelet scales. The scheme of layered zero coding (LZC) proposed by Taubman and Zakhor [7] further exploited the correlation of significant coefficient across wavelet scales.

In this research, we propose a scheme to predict the fine scale coefficients based on the coded coefficients at the coarse scale. The interscale prediction is based on the Huber Markov statistical image model without any overhead information. We then apply it for image coding applications. The prediction is subtracted from the origin and the resultant residue is encoded. We show experimental results which demonstrate the effectiveness of the algorithm.

The paper is organized as follows. The theory of interscale prediction is explained in Section 2. The implementation of interscale predictive wavelet coder (IPWC) is described in Section 3. We show the experimental results and compare IPWC with several other coders in Section 4. A concluding remark is given in Section 5.

2. THEORETICAL ASPECT FOR INTERSCALE PREDICTION

In this section, we describe the theoretical framework of interscale wavelet prediction, i.e., the prediction of the current scale coefficients from the coded coefficients at coarse scale. Assume depth d wavelet transform has been applied. Let the finest scale be scale 1, and the coarsest scale be scale d. Let the original and the predicted coefficients at scale s be $\mathbf{x_s}$ and $\mathbf{z_s}$. Let the coding bitstream, the direct decoded image, and the width of the quantization bin at scale s be denoted by $\mathbf{y_s}$, , $\mathbf{m_s}$ and $\mathbf{q_s}$, respectively. We assume if the quantized coefficients are entropy encoded, they are already decoded by the corresponding entropy decoder. Let Q and Q^{-1} be the quantizer and dequantizer, and H and H^{-1} be the forward and backward transform operator.

When applying the interscale prediction on scale s, IPWC first enlarges the processing image region from the coarse scale s+1 to scale s and fill the scale s coefficients with 0. It then predicts the scale s coefficients by the maximum *a posteriori* (MAP) optimization based on the available coding bitstream of scale s+1 and up:

$$\hat{\mathbf{z}}_\mathbf{s} = \arg\max_{\mathbf{z_s}} \Pr(\mathbf{z_s} \mid \mathbf{y_{s+1}}) \qquad (1)$$

We define the minus likelihood function as potential U(x):

$$U(x) = -\Pr(x) \qquad (2)$$

Therefore, the MAP optimization in (1) is equivalent to minimize the potential:

$$\hat{\mathbf{z}}_\mathbf{s} = \arg\min_{\mathbf{z_s}} U(\mathbf{z_s} \mid \mathbf{y_{s+1}}) \qquad (3)$$

By applying the Bayes rule, we have:

$$U(\mathbf{z_s} \mid \mathbf{y_{s+1}}) = U(\mathbf{y_{s+1}} \mid \mathbf{z_s}) + U(\mathbf{z_s}) - U(\mathbf{y_{s+1}}) \qquad (4)$$

Term $U(\mathbf{y_{s+1}})$ can be dropped in optimization for it is constant with respect to $\mathbf{z_s}$. In the prediction of scale s coefficients, the presense of scale s+1 coding bitstream $\mathbf{y_{s+1}}$ indicates that the original image at scale s and above is within the set:

$$\mathcal{G}_s = \left\{ \mathbf{x_s} \;\middle|\; \begin{matrix} QH[\mathbf{x_{s+1}}] = \mathbf{y_{s+1}} \\ QH[\mathbf{x_s}] = \text{arbitrary} \end{matrix} \right\} \qquad (5)$$

For scalar quantizer Q, set $\mathcal{G}_s$ can be equivalently written as:

$$\mathcal{G}_s = \left\{ \mathbf{x_s} \;\middle|\; \begin{matrix} |H[\mathbf{x_{s+1}}] - \mathbf{m_{s+1}}| \le \mathbf{q_{s+1}} \\ H[\mathbf{x_s}] = \text{arbitrary} \end{matrix} \right\} \qquad (6)$$

It is apparent that the set $\mathcal{G}_s$ is convex. Since only image in set $\mathcal{G}_s$ is compressed to bitstream $\mathbf{y_{s+1}}$, we can formulate $\Pr(\mathbf{y_{s+1}}|\mathbf{z_s})$ as:

$$\Pr(\mathbf{y_{s+1}}|\mathbf{z_s}) = \begin{cases} 1 & \mathbf{z_s} \in \mathcal{G}_s \\ 0 & \mathbf{z_s} \notin \mathcal{G}_s \end{cases} \qquad (7)$$

Applying the definition of potential in (2), we have:

$$U(\mathbf{y_{s+1}}|\mathbf{z_s}) = \begin{cases} 0 & \mathbf{z_s} \in \mathcal{G}_s \\ \infty & \mathbf{z_s} \notin \mathcal{G}_s \end{cases} \qquad (8)$$

Substitute (4), (8) into (3), the MAP estimation can be converted to a constraint optimization problem:

$$\hat{\mathbf{z}}_\mathbf{s} = \arg \min_{\mathbf{z_s} \in \mathcal{G}_s} U(\mathbf{z_s}) \qquad (9)$$

The potential $U(\mathbf{z_s})$ relates to the *a priori* model of the image. In this research, we use the Huber Markov random field (HMRF) statistical image model [5] , which has been shown to successfully model both the smooth regions and the discontinuities in the image. Since the likelihood function of Markov random field can be explicitly written as Gibbs function, the *a priori* probability $\Pr(\mathbf{z_s})$ can be represented as:

$$\Pr(\mathbf{z_s}) = \frac{1}{A_0} \exp\left\{ -\sum_{c \in \mathcal{C}} V_c(\mathbf{z_s}) \right\} \qquad (10)$$

where c is a pair of two neighbor pixels in the space doamin, and $\mathcal{C}$ is the set of all such neighboring pixel pairs. A_0 is a normalizing constants. Substitute the definition of the potential (2) and eliminate the constant A_0 in optimization, we conclude that:

$$U(\mathbf{z_s}) = \sum_{c \in \mathcal{C}} V_c(\mathbf{z_s}) \qquad (11)$$

$V_c(\cdot)$ is a local potential function in the form:

$$V_c(\mathbf{z_s}) = \rho(d_c^t \mathbf{z_s}, T_c) = \rho(x_1 - x_2, T_c) \qquad (12)$$

and

$$\rho(u, T_c) = \begin{cases} u^2, & |u| \le T_c \\ T_c^2 + 2T_c(|u| - T_c), & |u| > T_c \end{cases} \qquad (13)$$

where d_c^t is the differential operator of the pixel pair c, and T_c is the continuity factor of the pixel pair. We plot function $\rho(\cdot)$ in Figure 1 with solid line. Compared with function x^2, which is plotted with dashed line in the same figure, the slope of function $\rho(\cdot)$ is limited to T_c so that when the value of two neighbor pixels differs more than T_c, an edge is considered existing between the pixel pair. Function $\rho(\cdot)$ is known as the Huber minimax function, and for that reason this statistical image model is called the Huber-Markov random field (HMRF) model.

3. INTERSCALE PREDICTIVE WAVELET CODING (IPWC)

The coding procedure of the interscale predictive wavelet coding (IPWC) can be shown in Figure 2. IPWC encodes the image top-down, from the coarsest scale to the finest scale. The coarsest scale coefficients are encoded directly. As coding proceeds from scale s+1 to scale s, IPWC first extends the processing region to scale s, and then MAP predicts the scale s coefficients from the coding bitstream at scale s+1 and up. Next, the prediction is subtracted from the original image coefficients. And finally, the prediction residue is encoded. Because the prediction is based on the already coded coefficients at scale s+1 and up, the decoder can calculate the prediction the same way as the encoder, and thus reconstruct the original image by adding the encoded residue to the prediction.

The detail implementation of IPWC can be described below:

Step 1: Wavelet decomposition

The image is decomposed by a 2-D pyramidal wavelet transform with the biorthogonal 9-7 tap spline filter [9] . The symmetric boundary extension is used to reduce the signal discontinuity at the boundaries. We apply a depth 5 wavelet transform for an image of size 512×512.

Step 2: Direct coding the coarsest scale

The coarsest scale wavelet coefficients are encoded directly with a layered zero coder (LZC) developed by Taubman and Zakhor [7] . LZC is selected for its good R-D performance and its simplicity in implementation. However, we by no means restrict the wavelet coder to be LZC, other coding technologies, such as the embedded zerotree wavelet coding (EZW) [5] proposed by Shapiro or the set partitioning in hierarchical trees (SPIHT) proposed by Said and Pearlman, can be used as well.

Step 3: Interscale wavelet prediction with constraint optimization

As coding proceeds from scale s+1 to scale s, IPWC predicts the coefficient at scale s with MAP estimation (1), or equivalently the constraint minization of the potential (9). Since function $\rho(\cdot)$ is a convex function, and the potential $U(\mathbf{z_s})$ is formed by linear combination of $\rho(\cdot)$, it can be proven that $U(\mathbf{z_s})$ is also convex. The solution of the constraint optimization problem (9) in the convex set $\mathcal{Z}_s$ can thus be searched through gradient projection. As shown in Figure 3, IPWC sets the initial guess of $\mathbf{z_s}$ to be scale s+1 decoding $\hat{\mathbf{z}}_{\mathbf{s+1}}$, and fill scale s wavelet coefficients with 0:

$$H[\mathbf{z_s}] \leftarrow \begin{cases} H[\hat{\mathbf{z}}_{\mathbf{s+1}}], & \text{scale s+1 and up} \\ 0 & \text{scale s} \end{cases} \tag{14}$$

The prediction $\mathbf{z_s}$ is then updated with gradient projection, with detail procedures follows:

Step 3a: Gradient minimization

The steep descent algorithm is used to find the updated prediction $\mathbf{z'_s}$:

$$\mathbf{z'_s} = \mathbf{z_s} - \alpha_S\, \mathbf{g_s} \tag{15}$$

$\mathbf{g_s}$ is the gradient which can be calculated by:

$$\mathbf{g_s} = \nabla U(\mathbf{z_s}) = \sum_{c \in \mathcal{C}} \rho'(d_c^t\, \mathbf{z_s}, T_c) d_c \tag{16}$$

where $\rho'(\cdot)$ is the first derivatives of the Huber function. α_S is the stepsize calculated via:

$$\alpha_S = \frac{\nabla U^t(\mathbf{z_s}) \cdot \nabla U(\mathbf{z_s})}{\nabla U^t(\mathbf{z_s}) \cdot \Delta U(\mathbf{z_s}) \cdot \nabla U(\mathbf{z_s})} \tag{17}$$

with

$$\Delta U(\mathbf{z_s}) = \sum_{c \in \mathcal{C}} \rho''(d_c^t\, \mathbf{z_s}, T_c) d_c d_c^t \tag{18}$$

Step 3b: Projection

Since the updated image may fall outside of the coding constraint set $\mathcal{Z}_s$, the result of gradient minimization $\mathbf{z'_s}$ is projected back to $\mathcal{Z}_s$ via:

$$\mathbf{z''_s} \leftarrow \mathcal{P}_{\mathcal{Z}_s}(\mathbf{z'_s}) \tag{19}$$

The projection is performed in the wavelet domain as:

$$H[\mathbf{z''_{s+1}}] \leftarrow \begin{cases} \mathbf{m_{s+1}} - \mathbf{q_{s+1}}, & H[\mathbf{z'_{s+1}}] < \mathbf{m_{s+1}} - \mathbf{q_{s+1}} \\ H[\mathbf{z'_{s+1}}], & |H[\mathbf{z'_{s+1}}] - \mathbf{m_{s+1}}| \le \mathbf{q_{s+1}} \\ \mathbf{m_{s+1}} + \mathbf{q_{s+1}}, & H[\mathbf{z'_{s+1}}] > \mathbf{m_{s+1}} + \mathbf{q_{s+1}} \end{cases} \tag{20}$$

No change is needed for scale s coefficients.

Step 3c: Update the continuity factor

Before the next gradient projection, the continuity factor T_c for each pixel pair $x_1, x_2 \in c$ is adjusted. The difference before the steepest descent minimization x_1-x_2, after the steepest descent minimization x'_1-x'_2, and after the projection x''_1-x''_2 is checked. The continuity factor T_c is reduced by half for those pixel pair satisfy the conditions:

$$\begin{cases} |x_1 - x_2| > 0.7|x'_1 - x'_2| \\ |x''_1 - x''_2| > 0.7|x'_1 - x'_2| \end{cases} \tag{21}$$

The rationale of the above rule is that if the edge over the pixel pair c is at first smoothed by the steepest descent algorithm (15) and then pulled back by the projection operation (19), it is concluded that the continuity factor T_c is too strong and should be reduced.

Step 3d: Iteration

Step 3a to step 3c is iterated until the decrease in potential $U(\mathbf{z_s})$ is smaller than a predetermined threshold.

Step 4. Coding the prediction residue

As illustrated by Figure 3, after the interscale prediction, the prediction residue is calculated by subtracting the predicted wavelet coefficients from the original. The prediction residue is then encoded again by LZC. Step 3 and 4 is repeated until scale 1 wavelet coefficients have been coded.

4. EXPERIMENT RESULTS

We tested the interscale predictive wavelet coder (IPWC) on the Lena image of size 512 × 512. IPWC is compared with two other coders, the JPEG and the layered zero coder (LZC) [7] . We show JPEG as a reference. Because LZC is used to encode the prediction residue of IPWC, we compare IPWC with LZC to demonstrate the merit of interscale prediction. The original, JPEG coded, LZC coded and IPWC coded Lena and Barbara image are shown in Figure 4(a), (b), (c), and (d), respectively. For clarity, only the center 256 × 256 region of Lena is shown. Compared with MPEG4, the two wavelet coders - the LZC and IPWC, offer substantial improvement in both PSNR and subjective quality. Though the PSNR performance of IPWC is only slightly better than that of LZC, the subjective quality of the IPWC coded image is much better than that of LZC. The patch pattern surrounding the hat, cheek of Lena is gone; the ringing artifact around the eye of Lena is also greatly reduced. The result shows that IPWC provides effective MAP prediction for the small wavelet coefficients around the large isolated coefficients in the decoded image, and greatly reduces the patch patterns surrounding the slant edges and the ringing artifacts.

5. CONCLUSIONS

We propose an interscale predictive wavelet coding (IPWC) scheme. IPWC encodes the image top-down, from the coarsest scale to the finest scale. At each scale coding, it MAP predicts the current scale coefficients from the coding bitstream of the coarse scale, and then encodes the prediction residue. Experimental results show that IPWC is comparable in R-D performance with other state-of-the-art image coders. Moreover, viewing of the IPWC coded image shows that it has less ringing artifacts and better subjective quality.

6. ACKNOWLEDGEMENTS

This research was done when the author was at the University of Southern California. It was funded by the Integrated Media System Center, a National Science Foundation supported Engineering Research Center with additional support from the Annenberg Center for Communication at the University of Southern California and the California Trade and Commerce Agency.

7. REFERENCES

[1] S. G. Mallet and S. Gong, "Characterization of signals from multiscale edges," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 14, No. 7, pp. 710--732, Jul. 1992.

[2] Z. Berman and J. S. Baras, "Properties of the multiscale maxima and zero-crossings representations," *IEEE Trans. on Signal Processing*, Vol. 14, pp. 3216-3231, Dec. 1993.

[3] G. Caso and C. -C. J. Kuo, "New results for fractal/ wavelet image compression," in *SPIE Visual Communications and Image Processing '96*, vol. 2727, (Orlando, FL), pp. 536-547, Mar. 1996.

[4] J. Li and C. -C. J. Kuo, "Fractal wavelet coding using a rate-distortion constraint," in *IEEE International Conference on Image Processing*, (Lausanne, Switzerland), pp. II: 81--84, Sept. 1996.

[5] J. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. on Signal Processing*, Vol. 41, No.12, pp. 3445-3462, Dec. 1993.

[6] A. Said, and W. Pearlman, "A new, fast and efficient image codec based on set partitioning in hierarchical trees", IEEE Trans. On Circuit and System for Video Technology, Vol. 6, No. 3, Jun. 1996, pp. 243-250

[7] D. Taubman and A. Zakhor, "Multirate 3-D subband coding of video," *IEEE Trans. on Image Processing*, Vol. 3, No. 5, pp. 572--588, Sep. 1994.

[8] R. L. Stevenson and S. M. Schweizer, "A nonlinear filtering structure for image smoothing in mixed noise environments," *J. Mathematical Imaging and Vision*, Vol. 3, pp. 137-154, Nov. 1992.

[9] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. on Image Processing*, Vol. 1, pp. 205-230, Apr. 1992.

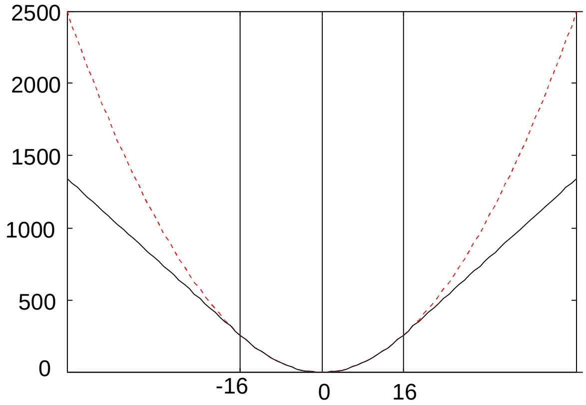

Figure 1. The plot of (a) Huber minimax function (solid line), (b) function x^2 (dashed line).

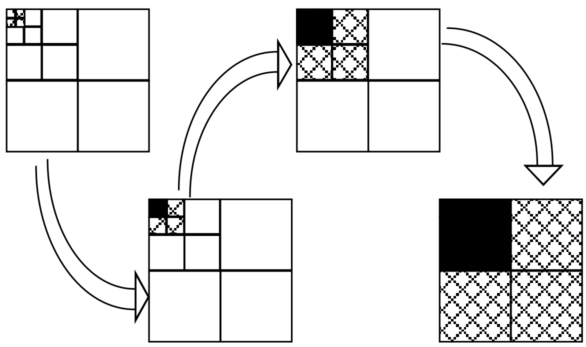

Figure 2. The coding procedure of interscale predictive wavelet coding (The black region has already been coded, the shaded region is the current scale under coding.)

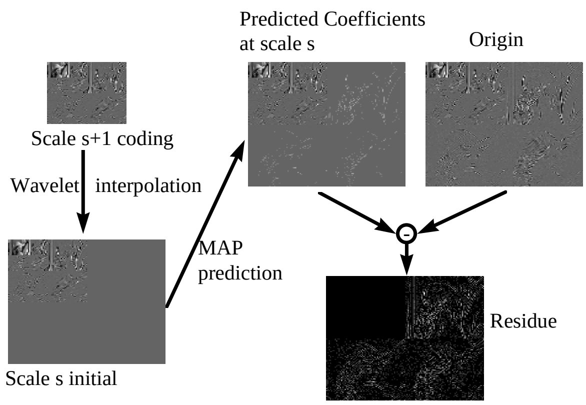

Figure 3. Illustration of interscale prediction

(a) (b)

(c) (d)

Figure 4 Experimental results of Lena image, the (a) original, (b) JPEG coded (0.17bpp, 27.3dB), (c) LZC coded (0.17bpp, 32.2dB), (d) IPWC coded (0,17bpp, 32.3dB).