Proceedings of the 2007 IEEE Symposium on Computational Intelligence in Image and Signal Processing (CIISP 2007)

Progressive Image Reconstruction based on Multi-scale Edge Model

Paul Bao, *Senior Member*, *IEEE*
†University of South Florida, USA, pbao@lakeland.usf.edu
Xianjun Zhang
‡East China Normal University, Shanghai, China, xjzhang@sei.ecnu.edu.cn

Abstract - **In this paper, we present a progressive image reconstruction scheme based on the semantically scalable multi-scale edge representation of images, with the resolution and visual quality scalable to various bitrate requirements. In the multi-scale edge representation an image is decomposed into its multi-scale primal sketch and the background where the multi-scale primal sketch preserves the structural semantics of images, and the background represents the smooth locale. Edge compensation is performed to smoothly remove edges at each scale. The multi-scale edges are then embedded encoded using the GFA modeling. The image reconstruction is progressively achieved by synthesizing multi-scale edges on the reconstructed image obtained from previous scale. As edge synthesis is performed at consecutive scales, the visual quality of the reconstructed image is progressively enhanced. Experiment shows that the proposed scheme performs well at low bit-rate multiresoultion representation and progressive reconstruction.**

I. INTRODUCTION

Wavelet-based image multiresolution representation and reconstruction [1-3] is advantageous compared with other methods such as the block-based ones, due to its unique joint space-frequency characteristics. However, the *structural aliasing* caused by the subsampling in wavelet transforms leads to seemingly inevitable artifacts and blurring around edge areas since the structural information sampled off could not be effectively compensated. This instigates the design of other forms of semantic representations of images.

The multi-scale edge representation was studied by Mallat *et al.* [4] in the framework of wavelet theory. It shows that multi-scale Canny edge detection is equivalent to finding the local maxima of a wavelet transform and the evolution of wavelet local maxima across scales characterizes the local shape of irregular structures. An algorithm that reconstructs a close approximation of image from its multi-scale edges is also presented [4]. However, the reconstruction method is complex and computationally expensive. Moreover, the quality of reconstructed image decreases significantly as the number of edge scales in the model drops.

Considering the drawbacks of pure wavelet-based coding and the essence of edges for visual perception, a number of papers proposed to integrate edge information into the coding streams. Schilling *et al.* [5] and Zhu *et al.* [6] presented, respectively, the schemes where edges are extracted and encoded separately while the original images are encoded by wavelets. Then, the coding of edges is multiplexed with the wavelet coefficients of the original images. Obviously, the performance improvement at visual quality is traded off by longer coding streams. Li [7] proposed another approach which transforms the original image to scale space by the forward diffusion with a *Gaussian* kernel at the chosen scale. The diffused image is then encoded by wavelet. When the image is decoded, an inverse diffusion is used to reconstruct the image. But this approach is computationally expensive and suffers the following that the inverse diffusion only works for a specific class of signals: ideal step edge whereas ramp edges and pulse edges are present in most natural images.

Xue *et al.* [8] proposed a multi-scale edge model which formulates a semantics-driven and directly operable image representation, aimed at supporting many common operations in visual computing and communications. In the multi-scale edge model, an image is decomposed into its multi-scale primal sketch (MSP) and the background. Multi-scale primal sketch is the union of the pulse and the ramp edges extracted and organized at consecutive scales and the background is the small image survived after a hierarchical scale transform of edge-removed image. The sample-based representation, such as DCT, wavelet and VQ, directly operate on low-scale image models. The object-based representation, as exemplified by MPEG4, facilitates some high-scale image operations such as scene composition of objects and background. Compared to these two representations, multi-scale edge model is a versatile hybrid semantic-statistic image representation, compromising between low-scale and high-scale image representations.

II. MULTI-SCALE EDGE MODEL

In this section, we briefly discuss the multi-scale edge model and the scale transform. A detailed description of the model can be found in [8].

A. Aliasing Effect of Wavelet Transform

The edges correspond to the variation of intensity values and two types of edge waveforms: pulse edge and ramp edge correspond to two different scene characteristics: short-term transients and object boundaries, respectively (figure 1). The multi-scale edge model recognizes the difference between these scene characteristics. The scalability is an intrinsic property of edges and, with a fixed image resolution, represents the spatial distance across which the intensity variation happens. Within this spatial distance, the gray of pulse edge varies approximately quadratically, and that of the ramp edge varies linearly. The scalability can be quantitatively described by edge scales [8], defined as $s = \lfloor \log_2 k \rfloor$, where k is the number of transfer pixels. If k = 0, the scale of the edge is defined to be s = 0.Pulse edges and pulse edges with scale 0, 1, 2 are illustrated in Figure 1. Edges of different scales may not be well preserved in wavelet transform due to subsampling in each scale. It may be observed

- edges of scale 0 at even or odd locations are poorly approximated by even- or odd-phase subsampling, respectively, causing edge aliasing
- the approximation improves drastically as edge scale increases to 1 or 2

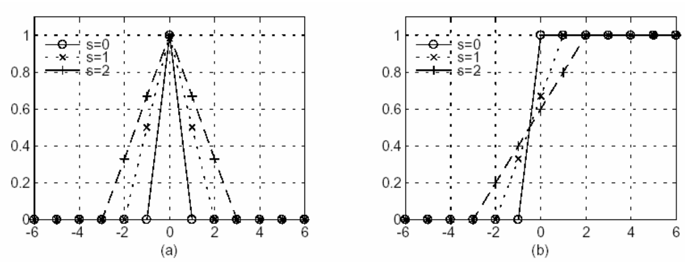


Figure 1: Edge Type and Scale. (a) pulse edges with scale 0, 1 and 2, which has 1, 3 and 5 transfer pixels respectively. (b) ramp edges with scale 0, 1 and 2, which has 0, 2 and 4 transfer pixels respectively.

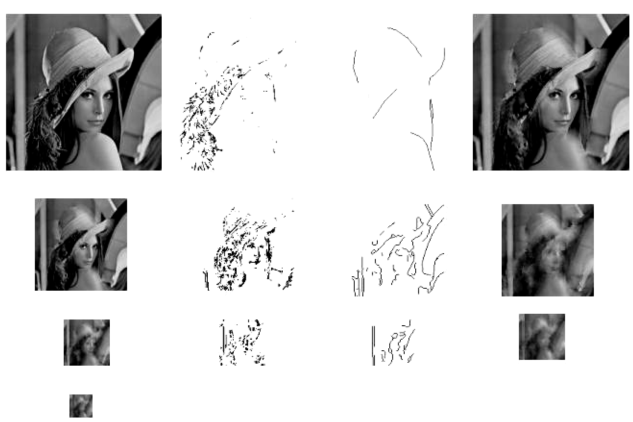


Figure 2. An example of multi-scale edge model

B. Multi-scale Edge Model

Multi-scale edge model decomposes an image into its multi-scale primal sketch (MPS) and the background [8]. An illustration of the multi-scale edge model is given in Figure 3. The left-upper image is the original input image. The pulse and ramp edges of scale 0 are extracted and represented as primal sketches, shown in the middle of the first row. The image is smoothed at the locations of the removed edges to minimize the discontinuity. This process is called *edge compensation*. Wavelet transform is then performed on the edge-compensated image and only the baseband image is retained as a new input image to the next scale of edge modeling. The edge modeling is repeated recursively for other scales.

C. Scale Transform

The scale transform, defined in [8], along with edge compensation and synthesis formulates the transform between images and their multi-scale edge representation. Scale transform can be derived from a wavelet transform as follows [8]. Consider the (biorthogonal) wavelet transform $\{h_n\}, \{g_n\}, \{\tilde{h}_n\}, \{\tilde{g}_n\}$. A sequence $\{c_n^0\}$ is dyadic decomposed to its low- and high-frequency components $\{c_n^1\}$ and $\{d_n^1\}$, respectively

$$c_n^1 = \sum_k h_{2n-k} c_k^0, \qquad d_n^1 = \sum_k g_{2n-k+1} c_k^0 \tag{1}$$

The input sequence may be reconstructed using only the low-frequency component

$$\tilde{c}_l^0 = \sum_n \tilde{h}_{2n-l} c_n^1 \tag{2}$$

Equation (1) is called the forward scale transform which along with the edge compensation transforms an image into its multi-scale edge model and (2) the inverse scale transforms which with the edge synthesis reconstructs the image from its multi-scale edge model. Integer scale transforms may be designed to facilitate the efficient and progressive reconstructions on limited devices. The generation of the multi-scale edge model consists of edge classifications and edge compensation and synthesis. An efficient algorithm for the multi-scale edge representation generation is discussed in [8].

III. MULTIRESOLUTION IMAGE RECONSTRUCTION

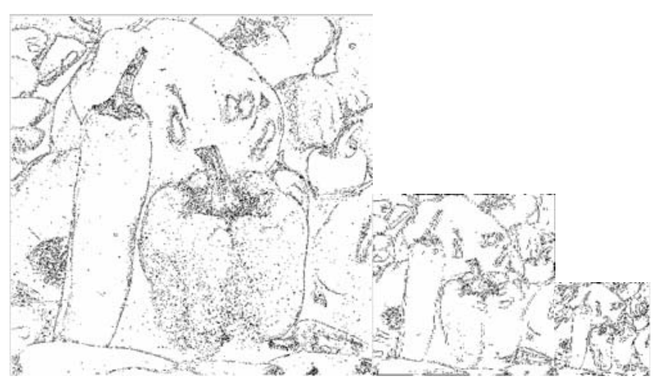


Figure 4. Detected edges of scale zero, scale one and scale two, respectively

In the multi-scale edge representation, the image is progressively described by the background and the multi-scale primal sketch, where the multi-scale primal sketch is scalable semantically, equivalent to reconstructing edges from high scales (dull structures) to low scales (sharp structures). As the progression of the reconstruction, higher frequency information is superposed onto the reconstructed image and the visual sharpness and resolution of reconstructed image are progressively enhanced. Figure 4 shows the extracted images of the multi-scale edges for image *Pepper*. Figures 5-7 show the progressive reconstruction sequence from the background and the primal sketches of scales 2 to 0, recursively, with their PSNR values for *Pepper*. For the reconstruction from the three-scale primal sketch, the edge blurring is observed, which is mainly due to the edge compensation and synthesis. An edge-compensated deblurring method may be designed.

Figure 5. Progressive reconstruction from one-scale primal sketch (a) background; (b) with scale-zero primal sketch (PS) at PSNR=29.97

Figure 6. Progressive reconstruction from two-scale primal sketch (a) background; (b) with scale-zero PS (c) with scale-one PS at PSNR=27.36

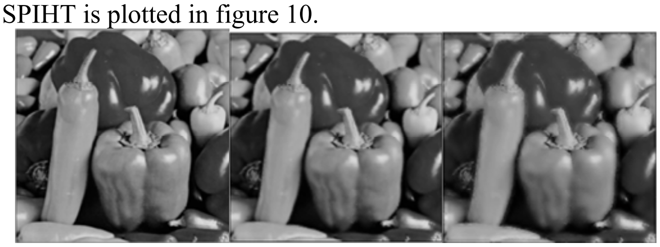
Figure 7. Progressive reconstruction from three-scale primal sketch (a) background; (b) with scale-zero PS; (c) with scale-one PS; (d) with scale-two PS at PSNR=25.16

The reconstruction is noticeably related to the contents in the images. Some images may contain very complex structures or greatly varying illuminations. A large number of short and dense edges may be extracted for such images and as a result, the bit rates increase accordingly. However, by properly thresholding the edge detection, the bit rates can be controlled within the budget of coding, traded-off by slight degradations of quality due to absence of some high frequency information. In other images, the background is rather simple and the foreground is well structured with a few large smooth regions, resulting in a few sparse and long edges

IV. EXPERIMENTAL RESULTS

We apply the multi-scale edge model based progressive image reconstruction on various benchmark images aimed at showing the effectiveness and rate-distortion performance of the scheme. Since at each scale some high frequency information may be misrepresented by multi-scale primal sketch, the loss of information is proportionally correlated to the scales of wavelet transforms. As the number of scales increases, the bitrate decreases but the reconstruction error increases, and vice versa. Therefore the scheme may be optimized and compromised on the multi-scale edge representation, visual quality and resolution scalability. Conventionally, a typical 512×512 image should be transformed to three or four scales for the multi-scale edged model based reconstruction.

Since the reconstruction is progressive and scalable at resolution and visual sharpness, the scheme is adaptive to various multimedia computing and communications. Figure 8 shows the reconstructed *Peppers* using one-, two and three-scale primal sketches. The PSNR measures of the reconstructed images from three-scale PS are around 25 dB, comparable to the result in [9]. Figure 9 shows the visual effect comparison for *Lena* at bitrates 0.01 and 0.04 between the proposed scheme and the well-known SPIHT scheme [3]. As expected, at the extreme low bitrates, since the proposed scheme allocates most of bit budget to the background and the high-scale structural constructs, only smooth background and the dull structures (high-scale sketches) are reconstructed. As the bitrate increases, sharper structures will be reconstructed. The rate-distortion comparison between the scheme and the SPIHT is plotted in figure 10.



PSNR=29.97 (1-scale) PSNR=27.36 (2-Scale) PSNR=25.16 (3-scale)

Figure 8. Reconstruction with bitplanes 7-0 (hybrid GFA and run-length code)

Table 1 gives their respective PSNR values. Note that the performance degrades as the complexity of the contents increases. However, the global structures and local smoothness are well preserved. The PSNR comparisons between the PS scheme and SPIHT at various bitrates are also given in Table II. Experiments show that our approach is comparable with SPIHT at rate-distortion performance but offers a semantically-embedded code stream for the progressive and scalable reconstruction at resolution and visual quality.

(a) SPIHT at 0.04 bpp (b) SPIHT at 0.1 bpp

(c) PS at 0.04 bpp (d) PS at 0.1 bpp

Figure 9. Visual effect comparison of the PS and SPIHT scheme

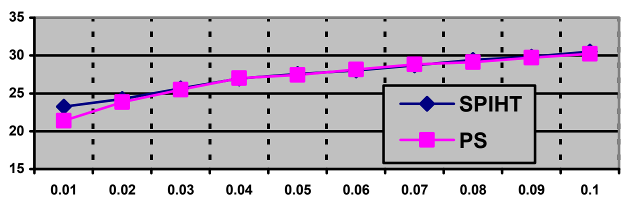


Figure 10. Rate-distortion of PS versus SPIHT

TABLE I. THE BIT BUDGETS OF MULTI-SCALE EDGE MODELS WITH DIFFERENT SCALES

Images	Bit budgets			
	1-scale	Base	Total	bpp
Peppers	9545	59340	68885	0.2627
Lena	9830	66187	76017	0.289
	2-scale	Base	Total	bpp
Peppers	11890	14480	26370	0.1046
Lena	13107	16538	29645	0.113
	3-scale	Base	Total	bpp
Peppers	14239	3650	17889	0.07
Lena	16384	4153	20537	0.078

TABLE II: PSNR VALUE COMPARISON WITH SPIHT FOR *LENA* AND *PEPPERS*

Images	SPIHT			PSNR/bpp		
	0.2	0.1	0.08	1-scale	2-scale	3-scale
Peppers	32.73	29.84	28.89	29.97/0.26	27.36/0.1	25.16/0.07
Lena	33.15	30.22	29.35	31.70/0.28	27.10/0.11	25.13/0.07

V. CONCLUSION

We propose a progressive image reconstruction scheme scalable semantically on the multi-scale edges and resolutions, based on a low bit-rate multi-scale primal sketch representation. The scheme is asymmetric with a computational expensive multi-scale primal sketch modeling but a very efficient reconstruction process. The reconstruction scheme is progressive and scalable at resolution and edge semantics (scales) with the GFA-embedded code stream, enabling numerous multimedia communication applications over heterogeneous networks. The reconstruction scheme is comparable with the SPIHT at rate-distortion performance quality at very low bitrate. The design of the reconstruction based on fixed-point scale transform is also given to facilitate multimedia computing and communications on limited devices such as PDAs and mobile devices.

REFERENCES

[1] S. G. Mallat, "Multifrequency channel decompositions of images and wavelet models," in *IEEE Transactions Acoustics, Speech, and Signal Processing*, vol. 37, no. 12, pp. 2091- 2110, Dec. 1989.

[2] A. Said and W.A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans Circuits Syst. Video Technol*. vol. 6, no. 3, pp. 243-250, June 1996.

[3] Z. Xiong, K. Ramchandran, and M. T. Orchart, "Wavelet packet image coding using space-frequency quantization," *IEEE Transactions Image Processing*, vol. 7, no. 6, pp. 892-898, June 1998.

[4] S. Mallat and S. Zhong, "Characterization of signals from multi-scale edges," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14, No. 7, pp. 710-732, July 1992.

[5] D. Schilling and P. Cosman, "Edge-enhanced image coding for low bit rates," *Proc. IEEE International Conference Image Processing*, vol. 3, pp. 747-751, 1999.

[6] B. Zhu, A. H. Tewfik, M. A. Colestock, O. N. Gerek, and A. E. Cetin, "Image coding with wavelet representations, edge information and visual masking," *Proc. IEEE International Conference Image Processing*, vol. 1, pp. 582-585, 1995.

[7] X. Li, "Low bit rate image coding in the scale space," in *Proc. 2002 IEEE International Conference Data Compression*, pp. 33-42, 2002.

[8] X. Xue and X. Wu, "Directly operable image representation of multi-scale primal sketch", *IEEE Transactions on Multimedia*, vol. 7, no. 5, pp. 805-816, Oct. 2005

[9] J. Elder, "Are edges incomplete?", *International Journal of Computer Vision*, vol. 34, no. 2/3, pp. 97-122, 1999.