A Three-Level Hierarchical Encoder Using Shape Independent Transform

Martin MOJZEŠ, Jaroslav POLEC

Dept. of Telecommunications, Slovak University of Technology, Ilkovičova 3, 812 19 Bratislava, Slovakia

mojzes@ktl.elf.stuba.sk, polec@ktl.elf.stuba.sk

Abstract. *In this paper a scheme for utilizing shape independent basis functions for a hierarchical multiresolution image compression is shown. First, an image is segmented and its segments' boundaries are polygon approximated, thus achieving an image mask. Second, this image mask and the image are used as an input of a three-level hierarchical encoder. The hierarchical encoder subsamples the image and the image mask and encodes them shape independently; it produces an output bit stream on a respective level that is also used on lower level(s) for further coding. On the base level a triangulation of the image mask is performed for superior performance. Another compression mode is, hence, introduced for the shape independent transform coding.*

Keywords

Image compression, hierarchical multiresolution encoding, shape independent transform.

1. Introduction

Modern communication technologies develop very fast. Transmitted information needs to be displayed to a user. In general, different compression modes can be used – sequential, progressive or hierarchical. They define the way the most important information from the data is chosen and the order in witch it is processed.

Different users use different display devices, ranging from cell phone displays to high resolution LCDs, to portray image information. If one source image was available and straightforward encoding method was used, full decoding would have been necessary to be performed for each display resolution although many displays require only the lower resolution. Thus we recognize a need for multiresolution processing which would ensure different resolutions in one bit stream that does not have to be fully decoded to receive the lower resolution image. Multiresolution output bit stream is achieved by hierarchical schemes, e.g. [1]. In the hierarchical scheme, an image is coded as a sequence of layers in pyramid. A decoder then processes only the pyramid layers necessary to decode to specific resolution. Hence one bit stream contains all the data needed to decode to multiple resolutions. Another asset of the multiresolution approach can be observed when errors occur in transmission. Even if it is impossible to resend the image data and if the error is not in the beginning of the bit stream, low resolution image is already correctly received. Moreover, user may want to receive the low resolution image first and then decide if he or she wishes to receive the full resolution image, too.

We are not aware of any hierarchical compression method using the shape independent transform. In [6], a quasi-progressive compression mode although labeled as progressive mode is presented. [14] introduces progressive segmentation-based image coding.

1.1 Prior Work

Our scheme is based on a region-oriented image coding algorithm employing orthonormal basis functions [10] improved on by a segment-oriented shape independent orthogonal transform texture approximation [6] in sequential mode. We showed that the same basis functions can be used also for a segment-oriented interpolation of images [11] and that we can use different basis functions [7]. For faster processing we presented smaller regions based on a triangulation of a segmented image [12] and further we showed that the missing data can be extrapolated with the existing set of basis functions of the selected orthogonal transform [13]. Progressive and hierarchical modes should be also of our interest. Progressive compression mode processes spectral coefficients from each region in order of their significance. It would require additional information about coefficient's position to be coded and also add complexity to the processing. Thus, we focus only on the hierarchical mode. Crucial prerequisite for the hierarchical compression mode is an ability to perform interpolation of a respective region with the chosen basis functions that we showed it holds [11]. No significant additional data reduction is expected, but the hierarchical compression and transfer mode may be of importance for specific applications as we outlined above.

In this paper a scheme for the hierarchical encoding in three levels using the shape independent transform is presented. In Section 2 description of the processes applied to the image before entering the hierarchical structure itself is given. Section 3 describes the hierarchical encoder with results in Section 4 and conclusion drawn in Section 5.

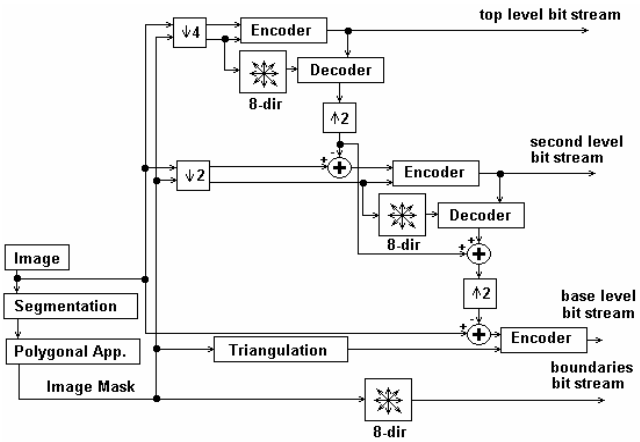


Fig. 1. Block diagram of the three-level hierarchical encoder.

2. Preprocessing

Digital image is an input of the proposed scheme. Concerning advantages of the segment-oriented approach in the image processing [2] this approach is used. The image hence needs to be prepared to be encoded using the shape independent transform. In our scheme, the image is segmented and the segments' boundaries are approximated before it enters the hierarchical processing (see Fig. 1).

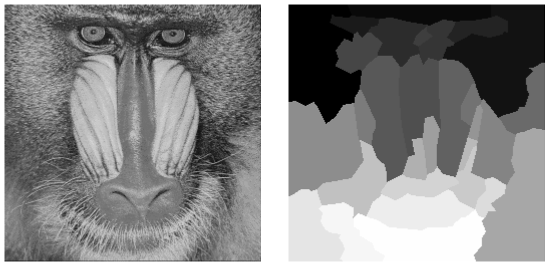
An unsupervised segmentation method for color-texture regions [3] is used. It tests for a homogeneity of the given color-texture pattern. The result of the segmentation can be an image where each region is defined by its unique gray or color value. Segment's texture and boundaries are processed separately, thus next boundaries are found using the 8-directional algorithm [4]. These boundaries are then approximated with polygons. The polygonal approximation supports the bit rate reduction, because there is no need to encode all the directions between the endpoints of the boundary segment. Hence an image mask is formed.

On the base level the highest resolution (the largest image) is being encoded. Here, the shape independent transform is applied to the largest segments in the process. This implies that a time needed for the transform is the highest. To provide faster processing, smaller regions entering the shape independent transform are necessary. Polygon decomposition technique [12] that introduces triangular mesh of the image mask [5] is, therefore, used.

3. Hierarchical Encoder

The image and its mask are an input into the hierarchical encoder (see Fig. 1). Both are sub-sampled by the factor of 4 and 2 in each dimension on the top and the second level, respectively. Each lower size (higher level) image provides prediction for the next level [1]. Except for the top level of the pyramid, the difference between the source image and the reference reconstructed image is coded. On the top level, the subsampled image is coded; and on the base level, the triangulated mask replaces the image mask as an input into the level coder. The boundaries of the image mask encoded using the 8-directional algorithm are transmitted in the separate boundaries bit stream. The level texture coding uses basis functions of 2D shape independent orthogonal discrete cosine transform II [6]. Shape independent transforms with different basis functions are also possible [7], [8]. The decoder operates in the reversed order.

Fig. 2, 3, 4 and 5 show the performance of the proposed scheme in preprocessing, on the top level, the second level and the base level, respectively.

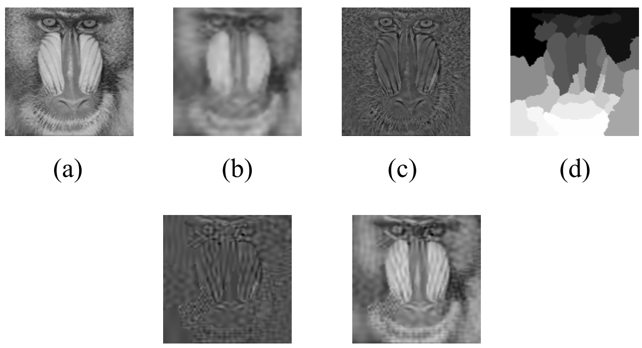


(a) (b)

Fig. 2. Image Baboon 256 x 256 pixels, 256 gray levels: a) original, b) image mask (the segmented image with polygonal approximation applied).

(a) (b) (c)

Fig. 3. Top level processing: a) the subsampled original image (Fig. 2a) by the factor of 4 in each dimension (64 x 64 pixels), b) the subsampled image mask (Fig. 2b) by the factor of 4 in each dimension, c) the decoded subsampled image a) (0.05 bpp of original size).



(a) (b) (c) (d)

(e) (f)

Fig. 4. Second level processing: a) the subsampled original image (Fig. 2a) by the factor of 2 in each dimension (128 x 128 pixels), b) the upsampled decoded image from top level (Fig. 3c) using B-spline interpolation, c) the difference between a) and b), d) the subsampled image mask (Fig. 2b) by the factor of 2 in each dimension, e) the decoded difference c) (0.05 bpp of original size), f) sum of b) and e).

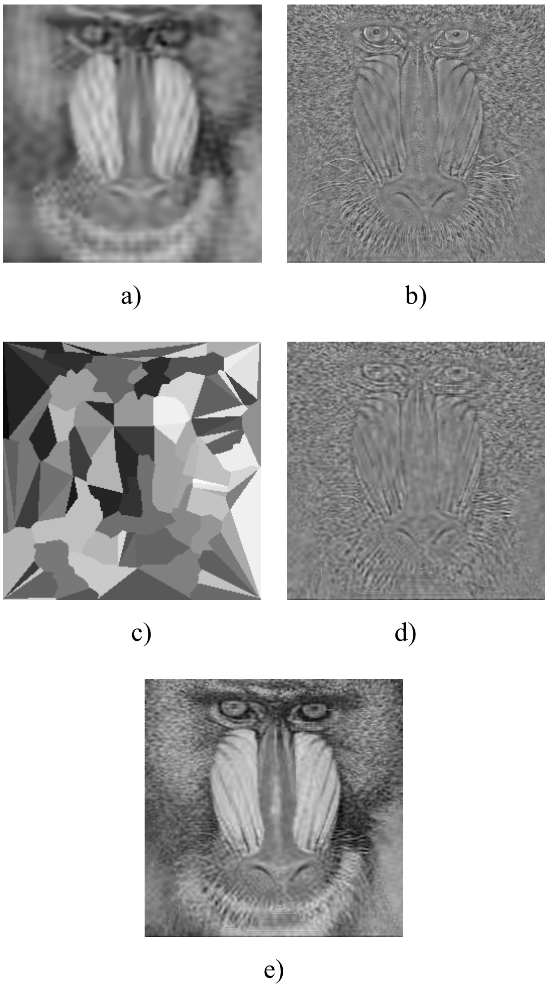


Fig. 5. Base level processing: a) the upsampled decoded image from second level (Fig. 4f) using B-spline interpolation, b) the difference between the original (Fig. 2a) and a), c) the image mask for base level – triangulated Fig. 2b), d) the decoded difference b) (0.35 bpp), e) sum of a) and d) - final product of the decoding process (0.5 bpp; PSNR 23.22 dB).

4. Results

Fig.6 depicts the results of the proposed scheme together with the comparison with sequential segment-oriented shape independent DCT scheme. Our approach is based on [6] and JPEG standard baseline coder [9]. Instead of the default shape independent basis functions any other shape independent basis functions can be used; and there can be different basis functions used on separate levels (until the same basis functions are used in corresponding level decoder). In Fig. 6 a), b), c) DCT-DCT-DCT, DCT-DCT-HT, DCT-DCT-WHT combinations of transforms, respectively, are shown. All of them have very similar PSNR performance even with JPEG but JPEG is affected by disturbing blocking effect. Among basis functions combinations, the DCT-DCT-DCT combination results in the best PSNR of 21.64 dB in the test image. Fig. 7 and 8 portray the PSNR performance of our scheme with DCT-DCT-DCT transform combination, sequential approach and JPEG baseline coder. As expected, very similar results for sequential and hierarchical modes can be observed.

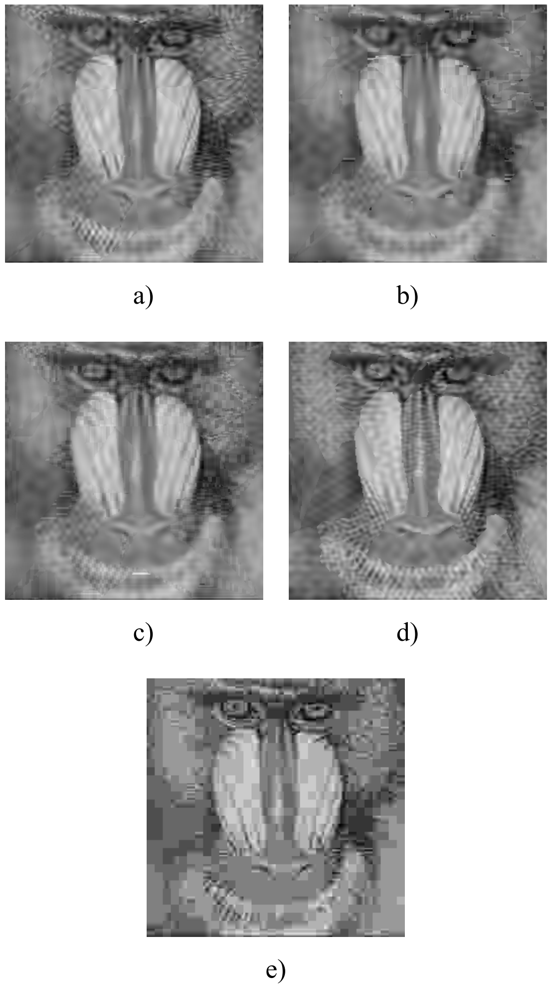


Fig. 6. Different transforms combinations comparison at 0.2 bpp: a) the proposed scheme with DCT-DCT-DCT level transforms (PSNR 21.64 dB), b) the proposed scheme with DCT-DCT-HT level transforms (PSNR 21.32 dB), c) the proposed scheme with DCT-DCT-WHT level transforms (PSNR 21.58 dB), d) the sequential segment-oriented DCT scheme (PSNR 21.3 dB), e) JPEG standard (PSNR 21.2 dB). [Note: DCT – discrete cosine transform, HT – Haar transform, WHT – Walsh-Hadamard transform].

5. Conclusion

The scheme introduces hierarchical coding for the shape independent transform. It offers the additional compression mode for arbitrarily shaped segments coding. Initial segmentation of the test image produced 30 arbitrarily shaped segments. They were processed at different resolutions and overall results were comparable to the sequential mode results. Hence, the functionality of our scheme was demonstrated.

Further research should also evaluate if the extrapolation with the used basis functions can be efficiently applied for the missing data estimation.

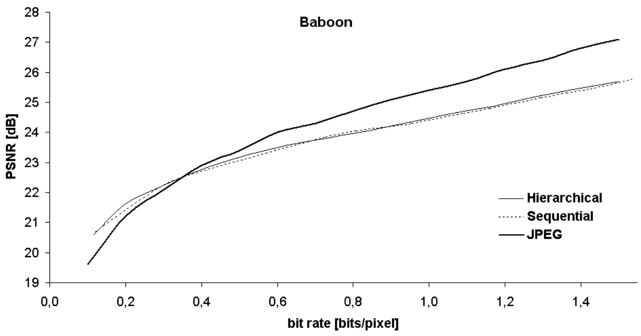


Fig. 7. PSNR performance of hierarchical mode, sequential mode and baseline JPEG coder for the Baboon (256 x 256 pixels) image.

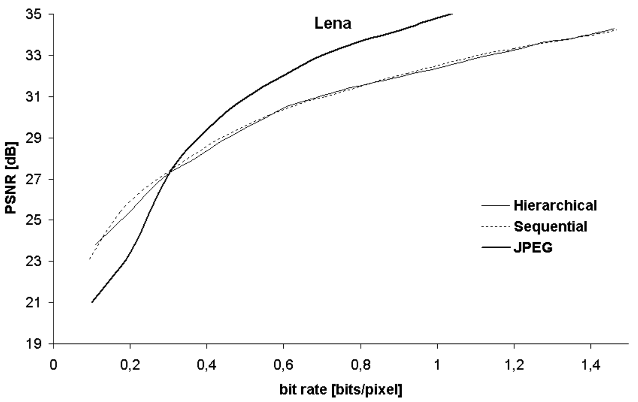


Fig. 8. PSNR performance of hierarchical mode, sequential mode and baseline JPEG coder for the Lena (256 x 256 pixels) image.

Acknowledgments

The research described in the paper was financially supported by the Slovak Research Grant Agency (VEGA) under grant No. 1/3094/06.

References

[1] GHANBARI, M. *Video Coding: An Introduction to Standard Codecs.* IEE, London, 1999, p. 60-62.

[2] PAL, N. R., PAL, S. K. A review on image segmentation techniques. *Pattern Recognition*, 1993, vol. 26, no. 9, p. 1277 - 1294.

[3] DENG, Y., MANJUNATH, B. S. Unsupervised segmentation method for colour-texture regions in images and video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, August 2001,http://csdl.computer.org/dl/trans/tp/2001/08/i0800.pdf.

[4] GONZALES, R. C., WINTZ, P. *Digital Image Processing.* 2nd ed. Reading. Addison-Wesley Publishing Company, Tokyo, 1987.

[5] KOLINGEROVA, I. On triangulations. In *Computational Science and Its Applications – ICCSA 2004: International Conference, Proceedings, Part II.* Assisi (Italy), May 14-17, 2004, p. 544 – 553.

[6] KAUP, A., AACH, T. Coding of segmented images using shape-independent basis functions. *IEEE Transaction on Image Processing*, 1998, vol.7, no.7, p. 937-947.

[7] POLEC, J., PAVLOVICOVA, J., VANYA, R., KARBULIKOVA, T., MASLIK, L., PARTYK, M. New scheme for region approximation and coding with shape independent transform. In *Proceedings IAPRS, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.* Graz (Austria), 2002, vol. XXXIV, part 3A/B, p. B-214-217.

[8] VARGIC, R., PROCHASKA, J. An adaptation of shape adaptive wavelet transform for image coding. In *EURASIP2005.* Smolenice (Slovakia), June 29 – July 2, 2005.

[9] BIEMOND, J., et al. *Image and Video Compression Learning Tool VcDemo,* http://www-it.et.tudelft.nl/~inald/vcdemo/, 2004.

[10] GILGE, M., ENGELHARDT, T., MEHLAN, R. Coding of arbitrarily shaped image segments based on a generalized orthogonal transform. *Signal Processing: Image Communication.,* Oct. 1989, vol. 1, p. 153–180.

[11] POLEC, J., PAVLOVICOVA, J., KARBULIKOVA, T. Application of shape-independent orthogonal transforms for image interpolation. *Radioengineering,* 2002, vol. 11, no. 1, p. 27-30.

[12] PARTYK, M., POLEC, J., KOLINGEROVA, I. Hybrid scheme with triangulations for transform coding. *Radioengineering,* 2003, vol. 12, no. 3, p. 42-45.

[13] KARBULIKOVA, T., POLEC, J. Extrapolation of incomplete image data with Discrete Orthogonal Transforms, *Radioengineering*, 2004, vol. 13, no 4, p. 35-37.

[14] WU, X., FANG, Y. Segmentation-based progressive image coding. In *Data Compression Conference*, 1993, p. 262-271.

About Authors...

Martin MOJZEŠ was born in 1982 in Bratislava, Slovak Republic. In 2006 he graduated (MSc.) from the Dept. of Telecommunications, Faculty of Electrical Engineering and Information Technology, Slovak University of Technology, Bratislava. Since then he has been a PhD. student at the Department of Telecommunications. His scientific research is focused on image and signal processing, image compression and reconstruction.

Jaroslav POLEC was born on 1964 in Trstená, Slovak Republic. He received his MSc. and PhD. degrees in telecommunication engineering from the Faculty of Electrical Engineering and Information Technology, Slovak University of Technology in 1987 and 1994, respectively. Since 1997 he has been associate professor at the Dept. of Telecommunications of the Faculty of Electrical Engineering and Information Technology, Slovak University of Technology, and since 1998 at the Dept. of Computer Graphics and Image Processing, Faculty of Mathematics, Physics and Informatics of the Comenius University. His research interests include Automatic-Repeat-Request (ARQ), channel modeling, image coding, interpolation and filtering.