SCALABLE IMAGE CODING USING GAUSSIAN PYRAMID VECTOR QUANTISATION WITH RESOLUTION-INDEPENDENT BLOCK SIZE

Leszek Cieplinski and Miroslaw Bober

Department of Electronic and Electrical Engineering, University of Surrey
Guildford, Surrey, GU2 5XH, United Kingdom
L.Cieplinski@ee.surrey.ac.uk, M.Bober@ee.surrey.ac.uk

ABSTRACT

We present a new approach to multiresolution vector quantisation. Its main advantage is exploitation of long-range correlations in the image by keeping vector size constant, independent of the image scale. We also developed a variable block-rate version of the algorithm, which allows better utilisation of the available bit budget by refining only those areas of the image which are not efficiently approximated by lower resolutions of the pyramid.

1. INTRODUCTION

Vector quantisation [1] is one of the well established techniques in image coding. It has several advantages as compared to the standard DCT-based methods. Probably the most important one is the possibility of adaptation of the codebook to a specific image type (e.g. faces, medical images). Another one is very low complexity of decoder which can be built as a simple lookup table.

The main drawback of vector quantisation is a very high computational complexity of codebook design and encoding. For the standard technique the complexity increases exponentially with bitrate which practically restricts the block size to 4x4 or 8x8 pixels and which makes the exploitation of long-range correlations in the image virtually impossible. This problem is usually alleviated by applying fast search algorithms to the basic LBG algorithm (see e.g. [1, 2, 3]) or by imposing constraints on the codebook structure (see e.g. [1, 4, 5]). Residual VQ [6, 7] gained popularity in recent years since it reduces complexity without significant decrease of performance. However residual vector quantisers operate on a single image resolution and use constant block size and therefore are not able to benefit from long-range spatial correlations.

In this paper we present a technique with improved performance due to exploitation of correlations on both coarse and fine image resolutions. It has also an important advantage of being scalable. In addition, the multiresolution representation makes it possible to adapt the distortion to the properties of the human visual system. In Section 2 we describe our algorithm, in Section 3 we compare its performance to standard vector quantisation and to the JPEG standard [8].

2. PROPOSED METHOD

Our technique operates in the following manner. We firstly form a Gaussian pyramid by recursively low-pass filtering and subsampling the image. We then perform vector quantisation at the coarsest resolution and calculate the error on the next finer resolution by upsampling of the reconstructed coarse image. The error image is vector quantised and the whole process is recursively repeated up to the original resolution. This process is illustrated in Fig. 1.

Figure 1. An example of Gaussian pyramid

The main difference between our approach and multiresolution VQ is that instead of using multiresolution representation of a single vector, where coarse resolutions have smaller sizes, we keep block size constant across scales. This allows us to exploit correlations between mean values of image subblocks.

The proposed approach has several advantages:

- it does not increase computational complexity and memory requirements
- it constructs a scalable bitstream, where coarse resolution image is a first approximation and successive refinements can be sent on demand
- it is better suited to the human visual system by paying more attention to low frequency components of the image.

To increase efficiency of encoding of the error images at fine resolutions, we introduce a version of this method with variable block rate. The energy of residual vectors is calculated and if it is below a predefined threshold only an escape code is sent to the bitstream. This leaves more bits for encoding of the blocks for which error was higher. To decrease the additional bitrate used by the escape codes we make this decision at the coarser resolution i.e. send single code for four blocks which formed a vector at the previous stage. This means that we need only 0.25 bit per block at the cost of less precise assessment of the error energy. The threshold is constant across resolutions. We made experiments with a threshold adapted to the energy of the error images but it does not seem to improve the performance.

Another variation of the basic scheme is inserting an interpolation step after upsampling of the reconstructed image. We use bilinear interpolation; it is applied both inside vectors and on the boundaries, which reduces blocking effects normally present in vector-quantised images.

3. RESULTS

In this section we present results of various versions of the proposed technique. We use 512x512 versions of "lena", "peppers", and "baboon" image for testing purposes. Codebooks are generated with training data consisting of 23 512x768 images from the Kodak database[1]. To accelerate the codebook design and encoding we use fast nearest neighbour search technique presented in [3], which is 20-50 times faster than the full search LBG.

We first present the basic version of our method, without variable-rate and bilinear interpolation modifications. In Fig. 2 we compare performance of the LBG algorithm to our technique. Results were obtained with optimal bit allocation between codebook at coarse and finer resolutions. Block size for both two- and three-level Gaussian pyramid VQ is 4x4. It is seen that for two resolutions performance is almost as good as that of LBG with 8x8 blocks and for three levels it approaches that of LBG with 16x16 blocks.

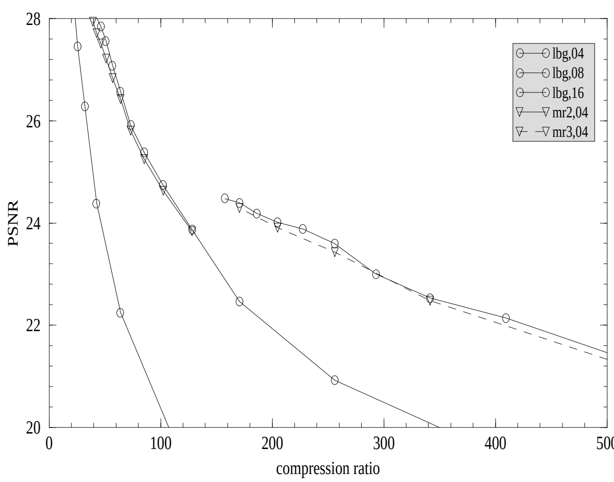

Figure 2. Comparison of LBG (lbg) and Gaussian Pyramid VQ (mr) results

In Fig. 3 we present the visual quality improvement obtained with the multiresolution technique. Both images are compressed 32 times. Codebook size for LBG is 16. For our technique, coarse resolution codebook has 16 vectors, and fine resolution codebook size is 8. This means that computational complexity of the multiresolution technique is about 25% lower. PSNR for LBG is 26.28 dB and for multiresolution 27.00 dB.

Figure 3. Comparison of visual quality between LBG (left) and Gaussian Pyramid Vector Quantisation for the "lena" image

Fig. 4 shows the behaviour of our technique with increasing codebook size at fine resolution. Adding bits at the finer level is seen to be less efficient in this case than increasing codebook at coarse resolution but performance is still better than that of the LBG algorithm. Scalability of the bitstream in this case is an additional gain.

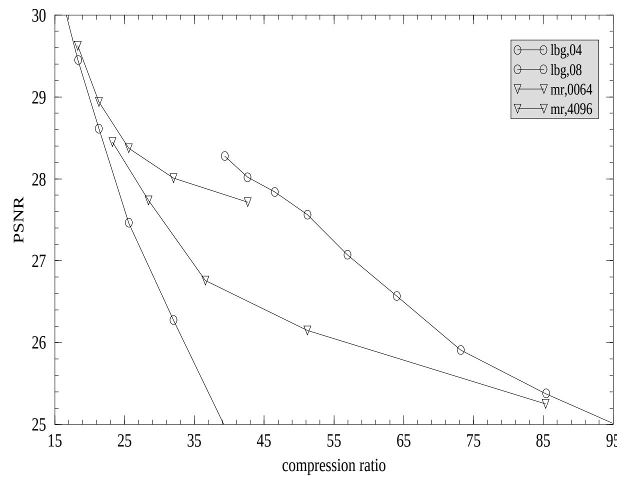

Figure 4. Performance of the Gaussian Pyramid VQ (mr) technique for various fine resolution codebook sizes

The problem of low efficiency of encoding of the error images is alleviated by using thresholding, as explained in the previous section. The performance of constant and variable block rate versions of our technique is compared in Fig. 5 for two coarse resolution codebook sizes. The fine resolution codebook size is varied and a threshold value is set to 81. The variable rate method is seen to outperform the constant rate version by about 0.5 to 1.5 dB. It is also seen that increasing codebook size is more effective in terms of quality improvement.

To reduce blocking artifacts in the image we performed experiments with postfiltering of the error images by bilinear interpolation. Fig. 6 shows performance of the modified algorithm compared to the standard one. A three level

[1]All the images used were obtained from the ftp site `ftp.ipl.rpi.edu`.

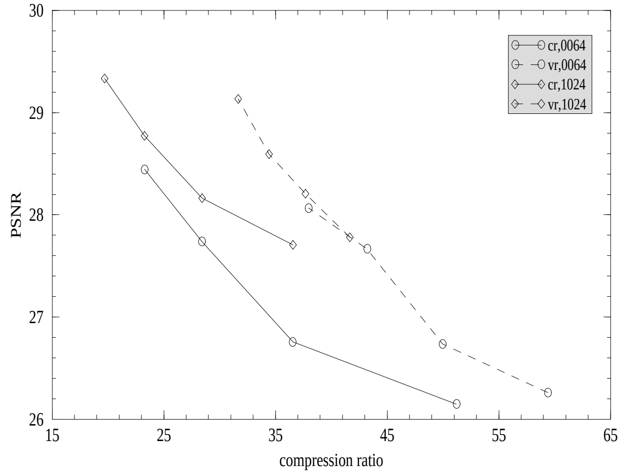

Figure 5. Comparison of constant (cr) and variable (vr) block rate versions of the algorithm

pyramid was used, codebook sizes for two coarse resolutions were 256, and codebook size for the fine resolution varied from 0 to 1024. The results without interpolation are denoted by "copy", results with bilinear interpolation by "bi". It is seen that additional quality improvement of about 1.5 dB, almost independent of the compression ratio is obtained for this case.

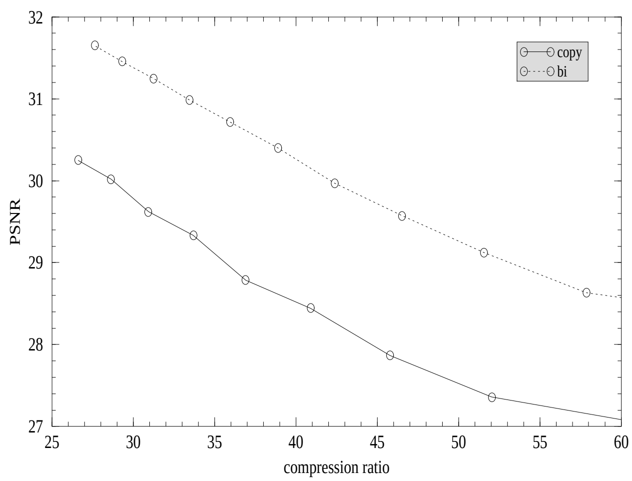

Figure 6. Performance of the bilinear interpolation version

Finally, in order to assess the usefulness of our technique in a more objective way, we compare it to the JPEG standard. We use the popular software implementation of JPEG available from the Independent JPEG Group[2]. It is seen that our technique performs significantly better than JPEG for high compression ratios.

Also visual quality obtained with our technique is much better than for JPEG, as illustrated in Figs. 8,9,10 for "lena", "peppers", and "baboon" images. In all cases a

[2]The current version is available by ftp from `ftp.uu.net:/graphics/jpeg/jpegsrc.v6a.tar.gz`.

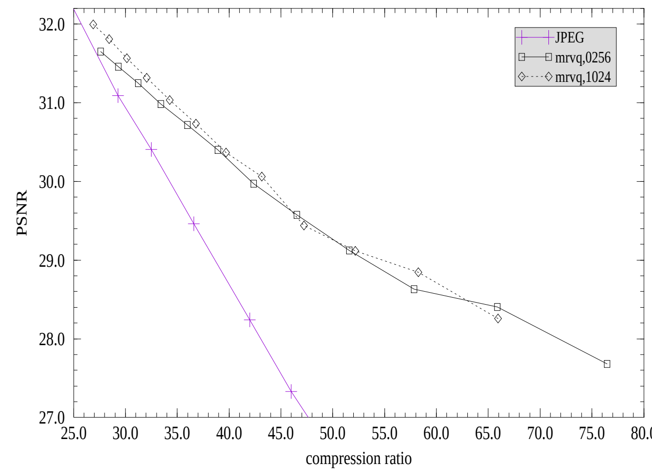

Figure 7. Comparison of Gaussian Pyramid Vector Quantisation to JPEG standard

three-level Gaussian pyramid was used, and the threshold was set to 100. Compression ratios are 62.25 and 65.87 for "lena", 64.08 and 68.57 for "peppers", 55.29 and 57.64 for "baboon" for JPEG and our technique, respectively. PSNR is in all three cases higher for vector quantisation than for JPEG.

Figure 8. Visual quality comparison between JPEG (left) and Gaussian Pyramid VQ (right) for the "lena" image

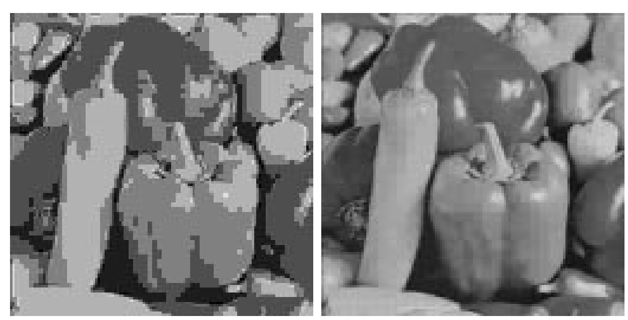

Figure 9. Visual quality comparison between JPEG (left) and Gaussian Pyramid VQ (right) for the "peppers" image

Results presented above were obtained with independent codebook design for different resolutions. It has been

Figure 10. Visual quality comparison between JPEG (left) and Gaussian Pyramid VQ (right) for the "baboon" image

shown [7] that for residual VQ with more than two stages independent codebook generation is suboptimal. Several approaches to joint codebook optimisation were developed (see [7]). Work is currently in progress on application of joint codebook optimisation techniques to our multiresolution algorithm.

4. CONCLUSIONS

We have shown that our approach gives substantial improvement of image quality for a given bitrate and computational complexity of the algorithm. Image quality measured by PSNR is from 0.5 to 3 dB higher than for the standard LBG with the same block size and only slightly lower than for two times larger blocks.

This work has been partially supported by European Community ACTS AC077 grant "Scalar" and by the Polish KBN grant 8 T11E 035 10.

REFERENCES

[1] A. Gersho and R.M. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic Publishers, Boston/Dordrecht/London, 1993.

[2] M.T. Orchard, "A Fast Nearest-Neighbor Search Algorithm", in *Proc. International Conference on Acoustics, Speech, and Signal Processing*, Toronto, Canada, April 1991, vol. 4, pp. 2297–3000.

[3] L. Cieplinski and C. Jedrzejek, "An Acceleration of Nearest Neighbour Search for Vector Quantization", in *Proc. IEEE Information Theory Workshop ITW'95*, June 1995, p. R11.

[4] R.L. Baker and R.M. Gray, "Differential Vector Quantization of Achromatic Imagery", in *Proc. International Picture Coding Symposium*, march 1983.

[5] T. Lookabaugh, E.A. Riskin, P.A. Chou, and R.m. Gray, "Variable Rate Vector Quantization for Speech, Image, and Video Compression", *IEEE Transactions on Communications*, vol. 41, no. 1, pp. 186–199, January 1993.

[6] F. Kossentini, M.J. Smith, and C.F. Barnes, "Image Coding Using Entropy-Constrained Residual Vector Quantization", *IEEE Transactions on Image Processing*, vol. 4, no. 10, pp. 1349–1357, October 1995.

[7] C.F. Barnes, S.A. Rizvi, and N.M. Nasrabadi, "Advances in Residual Vector Quantization: A Review", *IEEE Transactions on Image Processing*, vol. 5, no. 2, pp. 226–262, February 1996.

[8] W.B. Pennebaker and J.L. Mitchell, *JPEG Still Image Data Compression Standard*, Van Nostrand Reinhold, New York, 1993.