AN IMPROVED DECODING SCHEME FOR MATCHING PURSUIT STREAMS

Ivan Manić, Rosa M. Figueras i Ventura, Markus Flierl, Pascal Frossard and Pierre Vandergheynst

Signal Processing Institute
Swiss Federal Institute of Technology, Lausanne (EPFL)
{ivan.manic, rosa.figueras, markus.flierl, pascal.frossard, pierre.vandergheynst}@epfl.ch

ABSTRACT

This paper presents an improved coefficient decoding method for Matching Pursuit streams. It builds on the adaptive *a posteriori* quantization of coefficients, and implements an interpolation scheme that enhances the inverse quantization performance at the decoder. A class of interpolation functions is introduced, that capture the behavior of coefficients after conditional scalar quantization. The accuracy of the interpolation scheme is verified experimentally, and the novel decoding algorithm is further evaluated in image coding applications. It can be seen that the proposed method improves the rate-distortion performance by up to 0.5 dB, only by changing the reconstruction strategy at the decoder.

1. INTRODUCTION

Representations with redundant dictionaries are known to be beneficial for efficient signal approximation, especially at low bit rates. This is due to the fact that most of the energy of the signal can be captured by a few atoms. Redundant approximations, though, are not easy to deal with, due to the non-uniqueness of the solution. Matching Pursuit (MP) [1] is a popular greedy algorithm that allows to find a (suboptimal) solution with a limited complexity. The output of the MP algorithm is the sequence of atom indices and coefficients, which normally appear in decreasing order of magnitude.

For coding purposes, these coefficients have to be quantized and entropy encoded. In previous work, a rate distortion (RD) optimized quantizer has been designed [2], that takes into account the specific properties of Matching Pursuit coefficients. In this paper, we use this previous quantization scheme, but with an improved dequantization algorithm. This is inspired by results obtained in Vector Quantization [3], where optimal vector representatives are determined.

There are two approaches to quantize MP coefficients: the *a priori* scheme [4], where the encoder uses the quantized coefficients to update the residual signal and compensate for the quantization errors, and the *a posteriori* scheme, where the quantization does not influence the MP decomposition. The former is used to generate decompositions targeted for specific rates and the later for encoding signals where it is too expensive to run the MP algorithm several times with different quantizers. The MP stream is computed only once and then quantized to meet different rate constraints, with a limited penalty in distortion.

In this work, an improved decoding scheme for MP streams is presented, based on the *a posteriori* MP quantization method proposed in [2]. For high iteration numbers, the reconstruction values of the scalar quantizer are interpolated, giving for the same bit-rate, better image reconstruction quality. This can be very important for out-of-loop quantization, where the PSNR for given number of atoms becomes very close to the reconstruction without quantization, even with a relatively coarse quantization of the most energetic MP coefficients. The performance of the decoder is further evaluated in the context of image coding applications.

The outline of the paper is as follows: In Section 2 we present the novel idea of quantized coefficient interpolation and analyse several interpolation functions. In Section 3 the improvement in RD performance of our system is presented. The paper is concluded in Section 4.

2. RECONSTRUCTION OF MATCHING PURSUIT COEFFICIENTS BY INTERPOLATION

By observing the output of the dequantizer it can be noticed that the output values are relatively coarsely quantized for atoms with high iteration numbers (**Fig. 1**). It is known that coefficient magnitudes are always decreasing (when eventually the decreasing of the coefficients is not strict, a reordering can be done), and the coefficient decay follows a very well defined pattern. It is thus possible to estimate the original values and obtain a better quality for the reconstructed image.

Fig. 1 shows that the behaviour of the coefficient magnitudes is approximately linear in log-log scale starting from a certain iteration number. This linearity has been observed for most of the images studied. Therefore, we approximate the relation between the logarithmic coefficient magnitudes

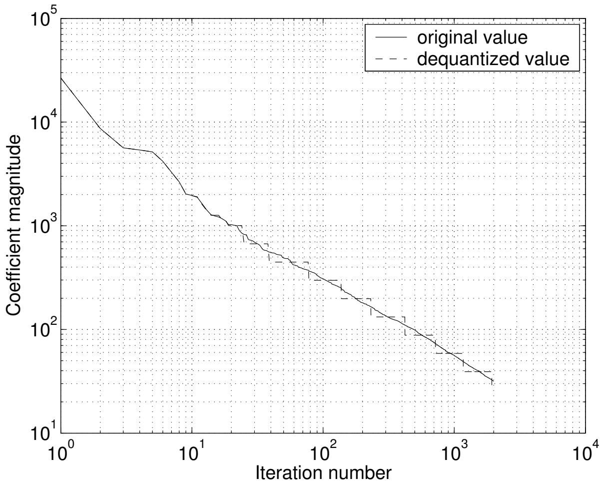

Fig. 1. Image Goldhill (256x256): Behaviour of unquantized coefficient magnitudes in log-log scale compared to dequantized values using 1024 initial quantization levels.

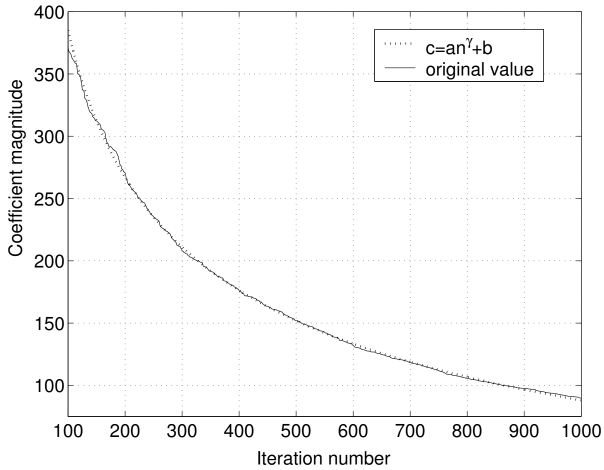

Fig. 2. Image Barbara (256x256): Approximation of coefficient magnitudes with (1); parameters a, b and γ are calculated by using a linear regression on original coefficient values for iteration numbers between 120 and 1000.

and the logarithmic iteration number by:

$$Y = AX + B,$$

where $Y = \log(\hat{c} - b), X = \log(n)$; $\hat{c}$ is the estimated coefficient magnitude and n is the iteration number. By solving the above equation, we obtain:

$$\hat{c} = e^B n^A + b = an^\gamma + b. \tag{1}$$

Parameters A, B and b can be computed by using a linear regression procedure [5], while iterating the value b, for which the squared error $e = (c - \hat{c})^2$ is minimized. The variable c denotes the actual value of the coefficient magnitude for a given iteration number. Parameter b is introduced to model slight changes in the linearity for different iteration numbers. **Fig. 2** shows how close is the approximation of original values with (1) for iteration numbers larger than 120 for the image Barbara. This function is calculated by linear regression for the iteration numbers in the range 120 to 1000.

In the quantization scheme proposed in [2], the number of quantization levels as a function of the iteration number decreases very rapidly, due to the quick decreasing of the coefficient magnitudes. The dequantized values are often constant on an interval, while the original values are decreasing slowly. Therefore, these constant regions could be dequantized with lower error if this decreasing of the values is taken into account.

For these constant regions, the most precise information about the coefficient magnitude values can be obtained from edge points (where the dequantized values change): the left edge point, $A\ (n_A, c_A)$ and the right one, $B\ (n_B, c_B)$. The only exception is the last region (where the stream is cut), for which no precise value of B can be found. To compute the coordinates of the edge points, the decoder performs the dequantization and notes two iteration numbers n_1 and n_2, when the dequantized value c_{deq} changes. The points A and B are then calculated as:

$$n_A = n_1 - 1/2, \quad c_A = (c_{deq}(n_1 - 1) + c_{deq}(n_1))/2$$
$$n_B = n_2 - 1/2, \quad c_B = (c_{deq}(n_2 - 1) + c_{deq}(n_2))/2$$

These values represent a very good approximation to the actual values of the coefficient magnitudes (see **Fig. 3**; the edge points are represented by asterisks). The decoder knows those edge points and can estimate the values for other iterations by using an interpolation function, better than a constant function. The interpolation of the dequantized values will only be done in the constant regions, while leaving the other values unchanged.

Starting from (1) we can derive several simple, but very useful interpolation functions, by fixing one of the three degrees of freedom, so that the other two can be estimated using the two edge points A and B. For each constant region with edge points A and B, the decoder can perform interpolation using the function $c = a/n + b$, which goes through these points (Eq. (1) with $\gamma = -1$) . The parameters a and b can be estimated as:

$$a = \frac{c_B - c_A}{\frac{1}{n_B} - \frac{1}{n_A}}, \quad b = c_A - \frac{a}{n_A}.$$

For each constant region the calculation of parameters is independent from neighbouring regions, except for the last segment (where the stream is cut), which requires special attention. As the decoder knows only point A, but not point

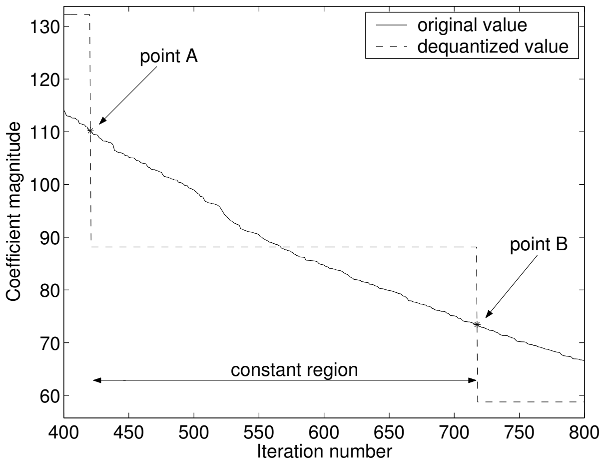

Fig. 3. Selection of end-points A and B for interpolating from dequantized values in a given constant region; end points are denoted by asterisks.

B, the interpolation will be done using the parameters computed from the previous segment,

An example of interpolation using the rational function is given in **Fig. 4**, which shows the PSNR as function of the iteration number for the image Cameraman. The number of the quantization levels for the first atom is 128, which is *very* small for 1000 iterations. It is important to know which cases should be interpolated. If not done properly, the interpolation process can actually decrease the quality of the reconstructed image. Inappropriate interpolation can take place if the interpolation is performed on a constant region of dequantized values that is not caused by relatively coarse quantization but by coefficient behaviour. This can happen only for low iteration numbers. The reduction in PSNR for one such case is shown in **Fig. 4**. Since the decoder knows the number of quantization levels it can start the interpolation when it is needed. It is much more difficult to predict the values of the coefficient magnitudes for low iteration numbers, and if the number of quantization levels is sufficiently high, the coefficient values should not be modified.

The interpolation scheme does not perform equally well for all images and all numbers of quantization levels. The image Barbara with 128 initial levels is such an example. The stream is cut after 1000 atoms. The obtained PSNR for rational interpolation using 128 initial levels is better than using 64 initial levels for most of the iteration numbers, but it deteriorates in the end. The reason for this deterioration is the interpolation for the last segment. For the case of 128 initial levels the final segment starts at iteration number 405 and continues till the end of the stream. The interpolation is based on parameters calculated for the previous segment. Since this last segment is very long, the proposed method does not perform very well. For the case of 64 levels, the last segment starts at iteration number 827 and is much shorter. This last constant segment can be held relatively short by using sufficiently high number of initial quantization levels.

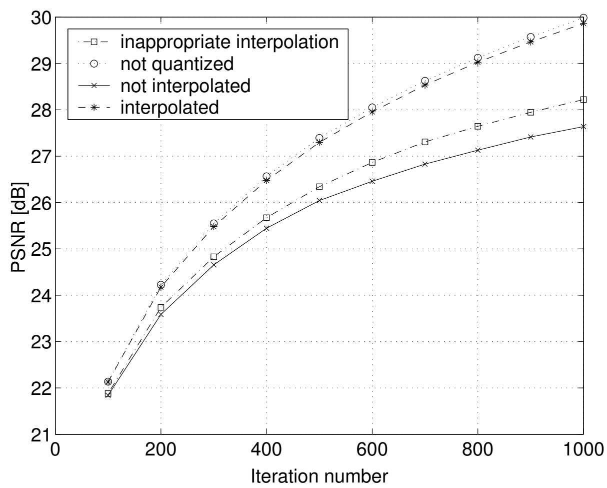

Fig. 4. Image Cameraman (256x256): interpolation with the rational function and 128 initial levels, compared to non-quantized values. Inappropriate interpolation (done before the number of quantization levels becomes low enough) is also shown. The experiments have been conducted with MP decompositions over Anisotropic Refinement Atoms and Gaussian Atoms subdictionaries, as described in [6].

Instead of the rational function $c = a/n + b$, we can use function $c = an^{\gamma}$, that also goes through two points A and B as discussed in the previous part. This is the special case of (1) with $b = 0$. The parameters can be calculated as:

$$\gamma = \frac{\log \frac{c_B}{c_A}}{\log \frac{n_B}{n_A}}, \qquad a = \frac{c_A}{n_A^{\gamma}}$$

We get very similar results with these two functions. The only noticeable difference is for the last segment, where the latter function performs better for certain images. Using sufficiently high initial number of levels the interpolated coefficient magnitudes are very close to the original values for both functions. However, the best approximation can be obtained if the parameters are calculated at the encoder, using (1), and sent to the decoder. If this is done, the quantized coefficient values for this region do not have to be transmitted at all.

3. RD PERFORMANCE

Increasing the initial number of quantization levels normally gives an increase in decoded image quality. However, this

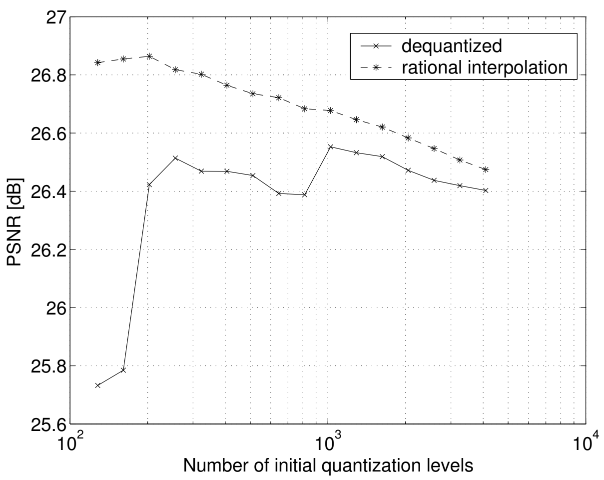

Fig. 5. Image Cameraman (256x256): Optimal number of levels for the first atom, for the given bit rate of 0.2bpp. The number of levels for the first atom ranges between 128 and 4096 with a resolution of 3 values per octave.

also gives an increase in bit rate, so there is an inflection point for which the obtained RD performance is best, and further increase in number of quantization levels produces much higher rates.

The best results in RD sense, when using interpolation, are obtained with a lower number of quantization levels than for the case without interpolation. Of course, the optimal number of levels depends on the rate. To determine the optimal number of levels, we fix the bit rate. In **Fig. 5**, the rate is 0.2bpp for the image Cameraman and the optimal number of levels without interpolation is 1024. By using interpolation and the same number of levels, we get an improvement in PSNR by 0.12dB. But by decreasing the number of levels, we get an additional increase by 0.19dB. The maximum RD performance point for the image Cameraman for 0.2 bpp is 1024 levels without interpolation, and 203 levels with interpolation (see **Fig. 6**).

4. CONCLUSION

The proposed method for implementing interpolation as an additional step after dequantization for MP coded images improves the image quality at the decoder, due to lower quantization noise. There are two possible approaches for doing this: for the first, the encoder stays the same and only the decoder is modified. In this case, there is backwards compatibility with the adaptive atom quantization of successive coefficients. For the second case, the encoder can calculate the parameters of the interpolation function (1), and, after some iterations, the interpolation parameters can be sent instead of the quantized values. This scheme may have better RD results at the possible price of a restrained flexibility. The design of a joint encoder quantization, and decoder interpolation scheme, is currently under study.

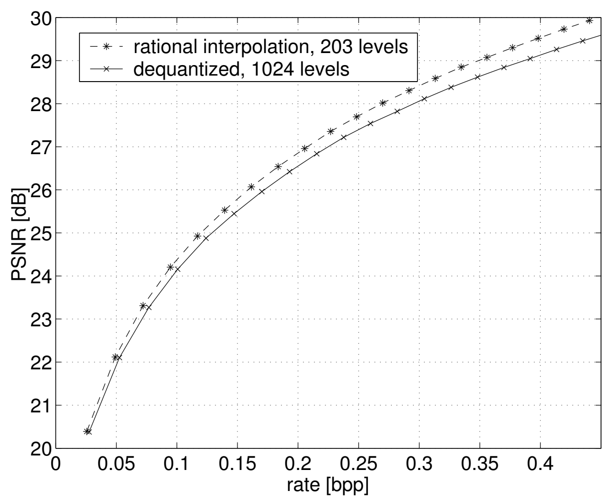

Fig. 6. Image Cameraman (256x256): The RD performance obtained with the optimal initial number of levels from the **Fig. 5** (203 with interpolation, 1024 without interpolation).

5. REFERENCES

[1] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.

[2] P. Frossard, P. Vandergheynst, R. M. Figueras i Ventura, and M. Kunt, "A posteriori quantization of progressive matching pursuit streams," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 525–535, Feb. 2004.

[3] P. A. Chou, T. Lookabaugh, and R. M. Gray, "Entropy-constrained vector quantization," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 1, pp. 31–42, Jan. 1989.

[4] C. de Vleeschouwer and A. Zakhor, "In-loop atom modulus quantization for matching pursuit and its application to video coding," *IEEE Transactions on Image Processing*, vol. 12, no. 10, pp. 1226–1242, Oct. 2003.

[5] N. R. Draper and H. Smith, *Applied regression analysis*, John Wiley & Sons, 3rd edition, 1998.

[6] R. M. Figueras i Ventura, P. Vandergheynst, and P. Frossard, "Low rate and scalable image coding with redundant representations," *IEEE Transactions on Image Processing*, July 2004, Submitted.