ADAPTIVE MULTIRESOLUTION IMAGE CODING WITH MATCHING AND BASIS PURSUITS

Hamid R. Rabiee¹, R.L. Kashyap¹, and S.R. Safavian²

¹School of Electrical & Computer Engineering
Purdue University
West Lafayette, IN 47907-1285

²LCC Inc., Cellular Institute
Arlington, VA 22201

ABSTRACT

In recent years there has been a growing interest in representation and compression of signals by using dictionaries of basis functions other than the traditional dictionary of sinusoids. These new set of dictionaries include cosine packets, chirplets, Gabor functions, wavelets, and wavelet packets. In this paper Matching Pursuit and Basis Pursuit with finite dictionaries of convolutional splines are used for adaptive multiresolution image compression. At the cost of computational complexity these algorithms outperform the DCT based JPEG both in terms of PSNR and subjective image quality at lower bit rates.

1. INTRODUCTION

Signal expansion is one of the major component of many image and video compression algorithms in digital signal processing. The popular block-based discrete cosine transform, which uses a fixed dictionary of sinusoids as basis functions for signal expansion, has been adopted by many of the emerging standards such as JPEG, H.261, and MPEG for image and video compression [1]. More recently, there has been a growing interest in representation and compression of signals by using dictionaries of basis functions other than the traditional dictionary of sinusoids [2-9]. These new set of dictionaries include cosine packets, Gabor functions, chirplets, wavelets, and wavelet packets.

Most of these dictionaries are overcomplete and lead to nonunique expansion of signals because some of the bases could be represented by linear combination of other bases in the dictionary. This non-uniqueness can be exploited by using efficient adaptive algorithms to achieve signal expansions that are sparse, high resolution and robust. These properties are central to image compression at low bit rates because sparsity leads to representation of signals with a few number of significant coefficients, high resolution results in better subjective quality, and robustness guarantees that small perturbations would not seriously degrade the quality of the representation. Matching pursuit [4] and basis pursuit [5] are two effective iterative algorithms for adaptive expansion of signals over finite dictionaries of basis functions. In this paper matching pursuit (MP) and basis pursuit (BP) algorithms are used for multi-resolution representation and compression of images over hierarchical data structures. The organization of this paper is as follows. Section 2 provides an overview of adaptive representation of signals with MP and BP. Section 3 presents the compression algorithm and explains the quantization and coding of the expansion coefficients. Finally, the experimental results and conclusions are provided in section 4.

2. ADAPTIVE REPRESENTATION OF SIGNALS

Let \( x = \{ x[n] : n = 0, 1, \ldots, N - 1 \} \) be a discrete-time signal in a finite dimensional Hilbert space \( H = \mathbb{R}^N \), with the inner product of \( x, y \in H \) defined as \( \langle x, y \rangle = \sum_n x[n] \cdot y[n] \), and the norm as \( \|x\| = \langle x, x \rangle^{1/2} \). Given a dictionary of basis functions \( D = \{ \phi_\gamma : \gamma \in \Gamma \} \) in \( H \), with \( \|\phi_\gamma\| = 1 \) and \( \text{span}(D) = H \), the goal is to obtain a representation of \( x \) with linear combinations of a small number of basis functions such that

\[
x = \sum_{\gamma \in \Gamma} a_\gamma \phi_\gamma \tag{2.1}
\]

or in an approximate decomposition

\[
x = \sum_{\gamma \in \Gamma} a_\gamma \phi_\gamma + R \tag{2.2}
\]

where \( \phi_\gamma \)'s represent the basis functions in the dictionary, \( a_\gamma \)'s are the coefficients of decomposition, and \( R \) is the residual. Typically, the dimension of \( D \) is larger than \( N \) and thus we have a redundant or overcomplete set of basis functions in \( H \), for signal decomposition.
Alternatively, if we write out all the vectors in the dictionary as columns of a matrix \( \Phi \), and all the coefficients as a column vector \( a \), then the above decomposition problem is that of finding an exact or approximate solution for the linear equation \( \Phi a = x \). The goal of these decompositions is to achieve sparsity, high resolution, robustness, and speed. The popular techniques for finding solutions to the above problem are the method of frames [2], best orthogonal basis [3], matching pursuit (MP) [4], and basis pursuit (BP) [5]. The MP and BP algorithms are explained below.

### 2.1 Matching Pursuit

The method of MP [4] uses a greedy algorithm that adaptively refines the signal approximation with an iterative procedure. Let \( D = \{ \phi_i ; i \in \Gamma \} \) be a dictionary of unit vectors in \( H = \mathbb{R}^n \), and \( x \in H \) be the input signal. The MP starts by searching \( D \) for some \( \phi_{i_0} \) such that

\[
|\langle \phi_{i_0}, x \rangle| \geq \alpha |\langle \phi_i, x \rangle|, \quad i \neq i_0
\]

where the parameter \( 0 < \alpha \leq 1 \), is an optimality factor which is typically close or equal to 1. At the first iteration of MP the signal \( x \) can be approximated as the sum of its projection onto \( \phi_{i_0} \) and the residue \( R_0x \)

\[
x = \langle \phi_{i_0}, x \rangle \phi_{i_0} + R_0x
\]

The algorithm then chooses the next basis in \( D \) to match \( R_0x \) and proceeds iteratively on the residues until some measure of error or convergence criterion such as \( l^2 \) norm is met. For example, the algorithm could be terminated at iteration \( p \) if \( \|R_px\| < \varepsilon \|x\| \) for some \( \varepsilon > 0 \). Therefore if we let \( R_0x = x \), the signal \( x \) can be expressed as

\[
x = \sum_{i = 0}^{p-1} \langle \phi_{i'}, R_0x \rangle \phi_{i'} + R_px
\]

Although in the above expansion the residue \( R_0x \) is orthogonal to \( \phi_{i_0} \), it may not be orthogonal to the other bases in the expansion (2.5). Therefore, even in a finite dimensional space the MP algorithm may converge slowly. To avoid this problem the MP algorithm can be orthogonalized to insure that at each level of iteration the residue is orthogonal to all the previous terms in the expansion (2.5). In [10] after each iteration the chosen basis from \( D \) is orthogonalized with respect to the previously selected bases before the residue at that level is calculated. In [7, 11] an orthogonal MP based on Gram-Schmidt procedure is driven that simply keeps an orthogonal set of best matches by successive projection of the remaining bases onto the orthogonal complement with respect to the current orthogonal set of chosen bases. The orthogonalized MP will converge in \( N \) steps [10, 11] due to the following theorem.

**Theorem 2.1** - Let \( H \) be an \( N \)-dimensional Hilbert space and let \( x \in H \) \((N \text{ may be infinite})\). Then an orthogonal pursuit converges in less than or equal to \( N \) iterations.

It is important to note that for \( P \) iterations of the MP algorithm the non-orthogonal MP is \( P \) times faster to compute than the orthogonal MP. Therefore only for small \( P \) the orthogonal MP is advantageous because it converges faster. Moreover, the problem of finding an optimal expansion for a signal with MP over redundant dictionaries is a NP-hard problem [11]. However, it is possible to find sub-optimal approximations over finite dictionaries which are capable of recovering the underlying sparse structure of the signals.

### 2.2 Basis Pursuit

The method of BP [5] uses a convex optimization procedure that adaptively refines the signal approximation over a redundant dictionary of basis functions. Given \( \Phi a = x \), as defined above, and an overcomplete dictionary \( D \) whose elements belong to a finite dimensional Hilbert space \( H = \mathbb{R}^n \), the method of BP tries to find a representation of the signal whose coefficients have minimal \( l^1 \) norm. That is, one have to solve an optimization problem of the form

\[
\min \|a\|_1 \text{ subject to } \Phi a = x \tag{2.6}
\]

From one point of view, the method of BP is very similar to the method of frames [2], because it simply replaces the \( l^2 \) norm in the method of frames with the \( l^1 \) norm. However, this minor change has a major impact on the outcome of the optimization problem of (2.6). While the method of frames solves a quadratic optimization problem, the method of BP should solve a convex and nonquadratic optimization problem. Although the method of BP involves nonlinear optimization, it is possible to reformulate the equation (2.6) into a linear optimization problem with the method of slack variables [12]. Moreover, for a data set at noise level \( \sigma > 0 \), it is possible to obtain an approximate solution as in (2.2) by solving the following optimization problem

\[
\min \|\Phi a - x\|_2^2 + \lambda_n \|a\|_1 \tag{2.7}
\]

where \( \lambda_n = \sigma \sqrt{2 \log (\text{card}(D))} \), and \( \text{card}(D) \) corresponds to the number of distinct basis functions in the dictionary.

The methods of BP and MP are both greedy algorithms,
but unlike the MP that builds up a new model in a greedy fashion with an iterative algorithm, BP starts with an initial model and iteratively improves the model by swapping the appropriate basis functions. In addition, based on the theory of linear programing it is possible for the linearized BP to converge to a global optimum, whereas the global optimality of MP is not guaranteed [12].

3. IMAGE COMPRESSION WITH MP AND BP

Because of sparsity, high resolution, and robustness of signal representation with the methods of MP and BP, it is possible to use these representations to achieve compression at lower bit rates. The method of MP has been successfully used for encoding of image and video at low bit rates [7-9]. In [9] we used the method of MP to generate multiresolution representation of images over quadtree (QT) and BSP tree hierarchical data structures [14]. In this paper we use both basis pursuit and matching pursuit over a finite dictionary of convolutional splines [13] to achieve low bit rate image compression. The MP and BP are used as a criterion for adaptive generation of quadtree hierarchical data structures to fully exploit the trade off between the complexity of the representation and the complexity of the tree. The BPQT (MPQT) compression algorithm can be summarized as follows. Given an uncompressed image, the method of basis pursuit (matching pursuit) is used to approximate the data at the finest level. The algorithm is then precedes to recursively break the data into dyadic intervals. The approximation continues until the desired peak signal-to-noise ratio (PSNR) or the desired measure of error is satisfied. The trees and the corresponding coefficients are then separately quantized and entropy coded. Since the success of our compression algorithm is highly correlated with the quantization of the coefficients of MP and BP representations, a consistency-based quantizer similar to the one explained in [16] were designed for optimal quantization of BP and MP coefficients. In our implementation the coefficients were quantized prior to the next step of the iteration to avoid propagation of quantization errors to subsequent iterations. This resulted in a PSNR gain of one to three dB over the Lloyd-Max quantizers that was used in [9] for quantization of the MP coefficients. The tree information and the quantized coefficients were separately entropy coded and combined to form the compressed bit stream. The experimental results are provided in the following section.

4. RESULTS AND CONCLUSIONS

The proposed algorithm was tested on the Lenna test image (512x512x8) which is shown in Fig. 4.1(a). The decoded Lenna at 0.125 bit/pixel and PSNR of 30.49 dB which has been encoded with the adaptive MPQT algorithm of section 3 is shown in Fig. 4.1(b). The decoded Lenna at 0.125 bit/pixel and PSNR of 30.58 dB which has been encoded with the adaptive BPQT algorithm of section 3 is shown in Fig. 4.1(c). For comparison, the JPEG reconstructed Lenna at 0.125 bit/pixel and PSNR of 22.36 dB is shown in Fig. 4.1 (d), and the PSNR curves for MPQT, BPQT, JPEG, and EZW [17] algorithms at rates between 0.0635 to 1 bit/pixel (compression ratios 128:1 to 8:1) are shown in Fig. 4.2. The quantization technique in [16] was used to quantize the MP and BP coefficients. The resulting quantized values and the tree informations were then separately encoded by using an adaptive arithmetic encoder [15]. In general, our segmentation based algorithms perform better than JPEG at rates below 1 bit/pixel based on subjective image quality (as shown in Fig. 4.1) and the PSNR measure of quality (as shown in Fig. 4.2). Moreover, due to the multiresolution structure and segmentation properties of the hierarchical data structures, the compressed bit streams of MPQT and BPQT algorithms can be used for progressive image transmission and browsing applications. The performance of our algorithms is also comparable to EZW in terms of subjective quality at the given rates, although in general EZW decoded images have a slightly higher PSNR at rates above 0.15 bit/pixel. The performance of our algorithm can be greatly improved by designing new optimal quantizers for MP and BP algorithms and enriching the dictionary of basis functions. The main disadvantage of MP and BP compression techniques is their computational complexity compared to block-based DCT and EZW algorithms. In our experiments the convergence of the BP algorithm was slower than the MP algorithm for a given measure of error. Finally, these algorithms introduce a variable size square block structure into the encoded images which may cause in some blocking effects at very low bit rates.

ACKNOWLEDGMENTS

The work of the first two authors was partially supported by the Innovative Science and Technology (IST) program of the BMDO monitored by the Office of Naval Research and a grant from National Science Foundation. WaveLab 7.0 was used to run some of the simulations.

REFERENCES