Multi-layered Image Compression

F.G. Meyer[1] and A.Z. Averbuch[2] and J.O. Strömberg[3] and R.R. Coifman[4]

[1]Radiology and Computer Science, [4]Mathematics, Yale University, New Haven, CT

[2]School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel

[3]Dept. of Mathematics, Tromsö University, Norway

ABSTRACT

The main contribution of this work is a new paradigm for image compression. We describe a new multi-layered representation technique for images. An image is encoded as the superposition of one main approximation, and a sequence of residuals. The strength of the multi-layered method comes from the fact that we use different bases to encode the main approximation and the residuals. For instance, we can use: – a wavelet basis to encode a coarse main approximation of the image, – wavelet packet bases to encode textured patterns, – brushlet bases to encode localized oriented textured features, etc.

Keywords: Image representation, image compression, best-basis, wavelets, libraries of bases

1. INTRODUCTION

The underlying assumption behind standard compression methods is that the basis (e.g. the DCT basis, or a wavelet basis) used for compression is well adapted to most images. If one wishes to obtain a very sparse representation of an image then it is critical that the functions in the basis bear a strong resemblance to the image being analyzed. If the geometric properties (shape, smoothness, periodicity, etc) of the basis functions match the corresponding characteristics of the data to be analyzed, then the coefficients in such a basis are meaningful, and provide rich information about the data being analyzed. Unfortunately basis functions that compactly code smooth images (such as wavelets with many vanishing moments) are ill suited to represent oscillatory patterns. Reciprocally, oscillatory analyzing functions such as the DCT can efficiently code textures, but they are ill equipped to represent smooth images. To alleviate this issue much larger libraries of waveforms, called wavelet packets,[5] have been developed. In the two dimensional case wavelet packets are patterns that can vary in scale, oscillation, and location. Because each wavelet packet library is overcomplete, it is possible to obtain a very sparse representation of the image by specializing, or "tailoring", the representation. In order to select an orthonormal "best-basis"[6] among all the wavelet packets, these patterns are matched to the image, and a selection of best matches which are sufficient for an efficient reconstruction is made. When coding images that contain a mixture of smooth and textured features, the best basis algorithm is always trying to find a compromise between two conflicting goals: – describe the large scale smooth regions, and describe the local textures. For this reason, the best wavelet packet basis rarely provides the optimal transform to compress an image.

Our approach follows a completely different direction: – instead of forcing all images to adapt to one single basis, we use a collection of libraries of bases to represent a single image. The main contribution of this paper is a new paradigm for image representation and image compression. We describe a new multi-layered representation technique for images. An image is encoded as the superposition of one main approximation, and a sequence of residuals. The strength of the multi-library approach comes from the fact that different sets of basis functions complement each others: –some of the basis functions will give reasonable account of the large trend of the data, while others will catch the local transients, or the oscillatory patterns. By selecting different analyzing functions, we allow different features to be discovered in the image. Furthermore, with each new basis we can encode and compensate for the artifacts created by the quantization of the coefficients of the previous bases.

Send correspondence to François Meyer : E-mail: francois.meyer@yale.edu

2. LIBRARIES OF BASES

During each iteration of the algorithm, we encode the remainder R^i of the previous iteration. This is done by selecting a basis $\mathcal{B} = \{\Phi_\gamma, \gamma \in G\}$ adapted to the image among a very large and highly redundant *library* of waveforms: $\mathcal{L} = \{\Phi_\gamma(t),\ \gamma \in \Gamma\}$:

$$R^i(x,y) = \sum_{\gamma \in G} \alpha_\gamma \Phi_\gamma(x,y) \tag{1}$$

where γ is a parameter indexing the waveforms. The library is *overcomplete* because, the number of basis functions Φ_γ is much greater than the effective dimension of the input space. Because the library is highly redundant, the representation is not unique. This redundancy allows us to choose among many possible representations the one that is best suited to our purpose. We review here the concept of libraries of orthonormal bases.

Such libraries are composed of functions, or atoms, with different time frequency localization. The atoms provide an overcomplete representation: - there is not a unique decomposition of each image over the library. Among all possible decompositions of a given image, one would like to pick up the most compact decomposition, which yields a sparse representation. Examples of such libraries are:

- **local trigonometric functions**.[1] Local trigonometric transforms provide an adaptive tiling of the spatial domain in terms of oscillating patterns. An image can be broken into blocks of different sizes within which a local Fourier expansion, or a DCT, is performed. Instead of abruptly cutting blocks in the image, we use a new family of *smooth orthogonal projectors*.[1,4,15]
- **wavelet packets**.[5] Loosely speaking, wavelet packets make it possible to adaptively tile the frequency domain into different bands of arbitrary size. Wavelet packets have been used to characterize textures in images.[2,7,9,?] However a wavelet packet is always associated with two peaks in frequency that does not allow to selectively localize a unique frequency.
- **brushlets**.[10] Brushlets are new families of steerable wavelet packets that adaptively segment the Fourier plane to obtain the most concise and precise representation of the image in terms of oriented textures with all possible directions, frequencies, and locations.

Because each library is overcomplete, it is possible to obtain a very sparse representation of an image by specializing, or "tailoring", its representation. It is critical that the basis functions in a library bear a strong resemblance to the signal being analyzed with this library. If the geometric properties (shape, smoothness, periodicity, etc) of the basis functions match the corresponding characteristics of the data to be analyzed, then the coefficients in such a basis are meaningful, and provide rich information about the data being analyzed.

Even though one could work with other overcomplete representations that do not necessarily contain orthogonal bases,[3,8] the libraries of orthonormal bases offers many advantages:

- they provide very large sub-collection of orthogonal bases,
- in an orthogonal basis the decomposition, and the reconstruction can be performed using very fast algorithm, which are numerically exact and stable,[5]
- there exist some fast algorithms that can be applied in real time, to select the optimal decomposition over the library.[6]

2.1. Best basis algorithm

The heart of the matter is then the selection of a subset of basis functions $\mathcal{B} = \{\Phi_\gamma, \gamma \in G\}$ of $\mathcal{L}$ such that the expansion (1) is optimal for a criterion Q. Such a subset is called the *best adapted basis*.[6,15] The best adapted basis can be expressed as the solution of the following optimization problem:

$$\max_{\mathcal{B} \subset \mathcal{L}} Q(\mathcal{B} = \{\Phi_\gamma,\ \gamma \in G\}) \quad \text{subject to} \sum_{\gamma \in G} \alpha_\gamma \Phi_\gamma = R^i \tag{2}$$

At first sight, the numerical solution of (2) appears to be non trivial: – a constrained nonlinear minimization problem, with possibly many local minima. However, a beautiful property of the wavelet packet dictionaries is that they can be organized in a hierarchical fashion, and the problem (2) can then be solved with fast algorithms.[6] The "best basis" paradigm permits a rapid (order $N\log(N)$, where N is the number of pixels in the image) search among the large collection of orthogonal bases to find that basis which permits the best approximation according to a given cost function (e.g. coding efficiency). The best-basis which minimizes this criterion is searched in this binary tree using a "divide and conquer" algorithm: – at each node, the cost is compared with the cost of the union of its two children's nodes and if the node's cost is smaller than the children's costs, the node is retained; otherwise, the children nodes are retained instead of the node itself. This process is recursively applied from the bottom to the top of the tree. Ramchandran and Vetterli[?] have proposed to select the best basis according to the rate distortion criterion. Each node of the wavelet packet tree is associated with the best scalar quantizer for that node. Then the best basis and the best set of quantizers are obtained using a pruning algorithm. The pruning procedure needs to be iterated several times to find the optimal slope on the rate distortion curve, at which all the quantizers of the best basis operate. We note that the results published in[?] correspond to hypothetical compression rates, since the first order entropy was chosen to measure the rate. In principle the approach of[?] is optimal for scalar quantizers. In practice, their approach is computationally intensive since it requires to search for the best basis several times in order to find the optimal slope on the rate distortion curve.

Our approach is much faster, and requires only one single pruning of the wavelet packet tree. Firstly, we de-noise the image: we threshold the coefficients in the wavelet packet tree to remove those coefficients whose magnitude are below a given threshold. The threshold is defined as the amplitude of the smallest non-zero coefficients that can be reconstructed after inverse quantization. Discarding the small coefficients permits to choose the best basis from the set of coefficients that will really contribute to the reconstruction of the image. Secondly, we measure the compactness of a basis with the first-order entropy. For each node of the tree, we calculate an histogram of the coefficients. This provides us with an approximation $\{p_i\}$ of the probability density function of the coefficients. The cost of the node is defined as: $-\sum_i \ p_i \ \log(p_i)$. We have tried several other cost measures. After de-noising, the first order entropy provides a very good measure of the overall budget required to encode the coefficients. Another excellent cost measure, which is faster to calculate, is defined as follows: – let λ be a given threshold, of the same order of magnitude as the quantization step, the cost of the node is the number of coefficients larger than λ.

3. MULTI-LAYERED IMAGE COMPRESSION

3.1. Cascade of compressions: the "multi-layered" paradigm

A block diagram of the multi-layered algorithm is shown in Fig. 1. The multi-layered compression algorithm consists in a cascade of compressions applied successively to the image itself and to the residuals that resulted from the previous compression. An initial main approximation is obtained by compressing the input image with a wavelet basis. We then reconstruct the compressed part, and we calculate the error between the original and compressed data. This compression error defines the first remainder, or *residual*. Residuals are compressed with wavelet packets or local trigonometric bases. Once the first residual is compressed, one defines the second residual as the compression error of the first residual. The algorithm keeps on compressing the successive residuals until we reach a residual that contains no more structure. We describe now in details the different stages of the algorithm. We consider a sequence of libraries of orthonormal bases:

$$(\mathcal{L}_i)_{i\in Z} \tag{3}$$

where each $\mathcal{L}_i$ is a library of bases. In this work $\mathcal{L}_0$ contains only one single wavelet basis, and $\mathcal{L}_1$ can be the library of wavelet packets associated with some different QMFs, or a library of local cosine transforms. The choice of the libraries, and their arrangement are parameters of the algorithm, and need to adjusted only once for large classes of images.

3.1.1. Initialization

Let I be an image. We first compress I over the library $\mathcal{L}_0$, using the budget b_0: – the approximation is performed under a budget constraint: the result of the approximation should be described with at most b_0 bits. Let $\hat{R}_0$ be the result of the compression, after decompression. $\hat{R}_0$ is an approximation of the original image I, and we have

$$I = \hat{R}^0 + R^1 \tag{4}$$

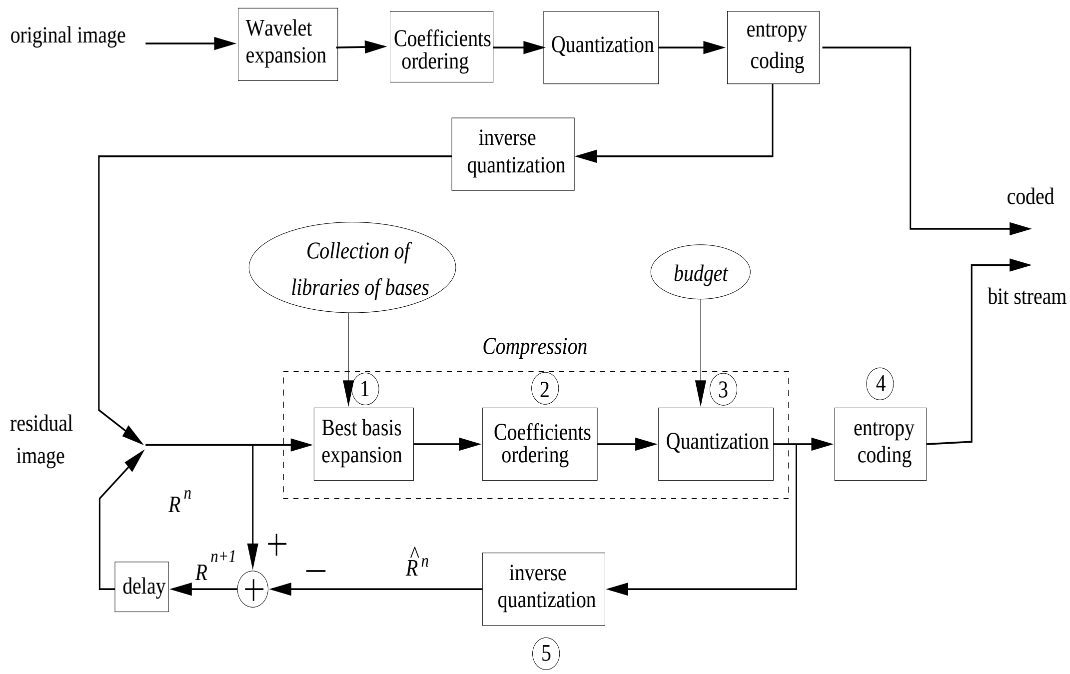

Figure 1. Block diagram of the multi-layered compression algorithm. In the first pass of the algorithm, the switch is turned towards the original image (i.e the image that we want to compress). In the subsequent refinement passes the switch is turned towards the residual image. The block in the dotted line compresses either the original image, or the residual. This single pass compression consists of three parts: (1) best basis selection, and calculation of the coefficients of the image using the best basis, (2) ordering of the coefficients, and (3) quantization of the stream of coefficients. Finally, in (4) the quantized coefficients are entropy coded. The residual error is calculated, and is fed back to the compression algorithm.

where R^1 is the approximation error. Let $\left\{\Phi_j^0,\ j \in E_0\right\}$ be the basis from $\mathcal{L}_0$ that was selected to compress I. We have

$$\hat{R}^0 = \sum_{j \in E_0} q_j^0 \Phi_j^0 \tag{5}$$

where $\left\{q_j^0,\right\}_{j \in E_0}$ are the quantized coefficients, and E_0 is the set of indices of the basis functions that constitute the best basis.

At this point we can refine the approximation of I by calculating an approximation of the residual R^1. This is achieved by compressing the residual. In order to discover different features in the image, we use a different library to compress R^1. We use a budget of b_1 bits to compress R^1, and we select the best basis $\left\{\Phi_j^1,\ j \in E_1\right\}$ from the library $\mathcal{L}_1$. We obtain:

$$R^1 = \hat{R}^1 + R^2 \tag{6}$$

with

$$\hat{R}^1 = \sum_{j \in E_1} q_j^1 \Phi_j^1 \tag{7}$$

We note that we can reconstruct an approximation of I:

$$\hat{I}^1 = \hat{R}^0 + \hat{R}^1 \tag{8}$$

where $\hat{I}^1$ is an image that can be compressed with the budget of $b_0 + b_1$ bits.

3.1.2. Main loop of the algorithm

Figure 1 shows the main loop of the algorithm. Let us assume that we have carried the approximation of I up to step $n-1$, and let us explain how to calculate the term n of the approximation. Let R^n be the residual of the approximation at step $n-1$. We select the best basis $\{\Phi_k\}_{k\in E_n}$ from the library $\mathcal{L}_n$, where E_n is the set of indices of the basis functions that constitute the best basis. We expand R^n into the best basis $\{\Phi_k\}_{k\in E_n}$, and we compress R^n using the budget b_n:

$$R^n = \hat{R}^n + R^{n+1} \tag{9}$$

where $\hat{R}^n$ is the compressed residual (after decompression):

$$\hat{R}^n = \sum_{j\in E_n} q_j^n \Phi_j^n. \tag{10}$$

Finally we can reconstruct an approximation of I using the $n+1$ compressed residual images:

$$\hat{I}^n = \sum_{k=0}^{n} \sum_{j\in E_k} q_j^k \Phi_j^k \tag{11}$$

where $\hat{I}^n$ is an image that can be compressed with the budget of $\sum_{k=0}^{n} b_k$ bits.

Note that the coefficients of the non-linear approximation are not the inner products with the basis functions. They are the quantized versions of the coefficients of the non-linear approximation.

3.2. Completion of the algorithm

We can stop the iterations when the residual is noisy and does not contain any apparent structure. This noisy part constitutes the incoherent noise that was present in the image. At this point no more structure that can be pulled out of the residual image.This can be made explicit by looking at the statistical distribution of the coefficients of the expansion.

3.3. Lossless compression

After the final compression iteration, one might still want to be able to save the final residual. This can happen for legal reasons, or simply in order to be able to come back to the raw data. This functionality is fully available in the multi-layer algorithm. The final residual can be entropy coded. Because during each iteration the amplitude of the residual decreases, the amplitude of the residual is usually small, and can be efficiently coded using a DPCM algorithm, followed by an entropy coder.

3.4. Toy example

Figures 2,3 and 4 illustrate the algorithm. Fig. 2 shows the original image *Barbara* on the left, and the final result after a compression of 32 with the multi-layer algorithm. The first two iterations of the algorithm are displayed in the following images. We show in Fig. 3 the result of the first iteration: – on the left, the smooth regions of the image have been described with a few wavelet coefficients (compression ratio= 254), – on the right is the residual: what is left when we subtract the smooth regions $\hat{R}^0$ out of the original image. The residual R^1 contains all the edges, as well as the textured regions. The wavelet compressed image $\hat{R}^0$ constitutes the first part of the compressed image. During the second pass of the algorithm, we compress the first residual R^1 with a wavelet packet basis (compression factor = 36.61), with the best basis algorithm. The result of the compression, $\hat{R}^1$, is shown in Fig. 4 on the left, and the error, R^2, is shown on the right. When the compressed residual $\hat{R}^1$ is added to the second residual R^2, we obtain the first residual, R^1, back. $\hat{R}^1$ constitutes the second layer. We note that most of the strong textures, as well as the edges have been removed from the image, and are encoded into this second layer.

Figure 2. Original image (left), and compressed image by a factor of 32 (right)

Figure 3. The first layer $\hat{R}^0$ (left), and the first residual part R^1 (right). The first layer is a very sparse representation of the smooth trends in the image. This first layer was obtained by compressing the original image in a wavelet basis, by a factor of 254.

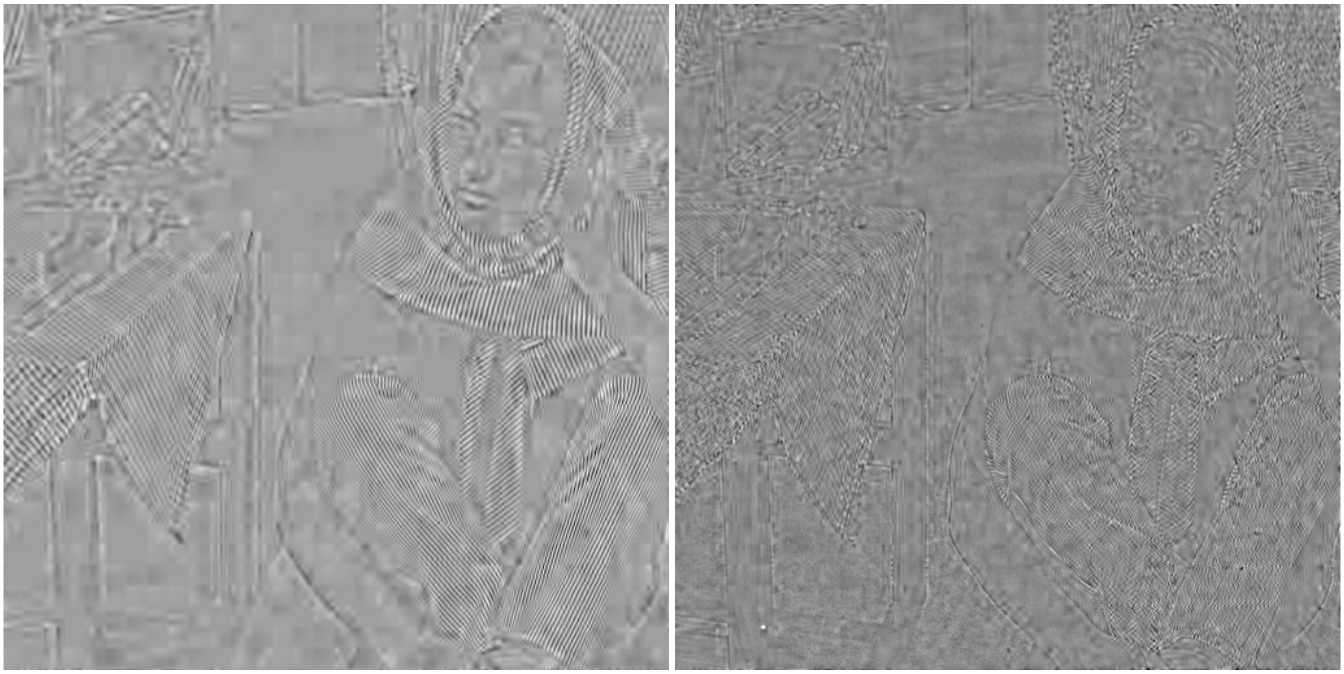

Figure 4. The second layer $\hat{R}^1$ (left), and the second residual part R^2 (right). The second layer is a sparse representation of the texture in the image. This second layer was obtained by compressing the first residual image using an adapted local cosine basis by a factor of 36.61.

Note that the final residual R^2 is visually almost random noise, i.e. we do not perceive any structure inside the image. When $\hat{R}^0$, $\hat{R}^1$, and $\hat{R}^3$ are added together, we obtain an image which is compressed by a factor of 32.

4. FREQUENCY ORDERING OF THE COEFFICIENTS

We explain here how to organize, and quantize the coefficients of the residual R^i, for each iteration i of the algorithm. We use the same ordering, and the same quantization methods for wavelet libraries, and local trigonometric libraries.

Once the best basis has been selected, the residual is represented by a set of coefficients. Before quantization we organize the coefficients in such a way that we can take advantage of the expected decay of the coefficients. After quantization, this organization results in a very sparse description of the non zero coefficients. Our scanning method exploits the fact that if an image is smooth, then the amplitude of the coefficients decreases as the frequency of the basis function increases. This result is certainly true if we use local Fourier or cosine bases. In the case of wavelet packet we have a similar result[11]:

LEMMA 4.1.[11] *If I is a C^r regular function, then $\exists C > 0$ such that the wavelet packet coefficients of I, $w_{n,\,j,\,l}$, satisfy*

$$\forall q \geq 0, \quad \forall n = 2^q, \ldots, 2^{q+1} - 1 \qquad |w_{n,\,j,\,l}| \leq C\, 2^{-q/2}\, 2^{-r(q+j)} \tag{12}$$

This bound on the coefficient tells us two things. On the one hand, we notice that as the scale j increases, the wavelet packet coefficients have a geometric decay 2^{-rj}. We note that this bound is similar to the bounds obtained in the standard wavelet basis. On the other hand, as the frequency n increases, the wavelet packet coefficients decay faster than n^{-r}, if the original image if C^r. The decay of the wavelet packet coefficient is thus similar to the decay of the Fourier coefficients. It also tells us that, even though we cannot exploit a multiscale structure among the wavelet packet coefficients, we should organize those coefficients by increasing frequency. More detail about the ordering can be found in.[9]

5. QUANTIZATION

Once the coefficients are ordered, we quantize them with a uniform scalar quantizer. The quantizer step is the same for all coefficients. Once the coefficients are quantized, a stream of bits is generated by bitplane encoding the absolute value of the quantized coefficients. If a coefficient is not quantized to zero, then its sign is also encoded. The bit stream contains long stretches of zeros and thus we use a zero-runlength coder to encode the stream. The parameters of the quantization are optimized in order to reach an exact budget. The uniform quantizer is characterized by two parameters: Δ the cell size, and $\Delta\, 2^{-l}$ the radius of the dead-zone around zero. The bitplane encoding consists in transmitting the digits 0 and 1 of the binary representation of each quantized coefficient. The bitplane encoding is characterized by two parameters: l the index of the smallest bit encoded, and m the index of the largest bit encoded. A coefficient x is quantized into $q(x)$ using the following rule:

- if $|x| < \Delta\ 2^l$ then $q(x) = 0$
- if $|x| \geq \Delta\ 2^l$ then we code the sign of x, and we define the quantized coefficient $q(x)$ as follows:

$$|x| = \Delta\ q(x) + r(x) \quad \text{with} \quad 0 \leq r(x) < \Delta\ 2^l \tag{13}$$

In order to reach the targeted budget, we optimize the value of Δ, and l, the two parameters of the quantization. The parameter m is the minimum number of planes needed to bitplane encode the absolute value, M, of the largest coefficients. We have

$$2^{m-1} \leq q(M) < 2^m\ -\ 1 \tag{14}$$

where $q(M)$ is the integer defined as follows

$$M = q(M)\ \Delta + r(M) \quad \text{with} \quad 0 \leq r(M) < \Delta\ 2^l \tag{15}$$

We need to be aware that Δ and l are dependent, since we obtain the following constraint from (13) and (15):

$$\frac{M}{2^l\ (2^{n-1}+1)} < \Delta \leq \frac{M}{2^l\ 2^{n-2}} \tag{16}$$

We want to find the smallest l, and Δ such that the number of bits needed to encode all the quantized coefficients and their signs matches the budget. The optimization proceeds as follows. First, we find m such that (14) and (15) are true. Secondly, we find the smallest value of l such that we can encode all the $q(x)$, as defined in (13), for all the coefficients x, and with $\Delta = M/2^l$. Finally, we apply a dichotomic search to find the optimal value of $\Delta \in [M/2^{l-1}, M/2^l]$. The sequence of 0 and 1 generated by the bitplane encoding is then encoded using a runlength coding technique. The best basis geometry is described by a quadtree. We code the quadtree, with an adaptive arithmetic coder.[16] The number of bytes required to encode the geometry is always is small fraction of the total budget.

6. EXPERIMENTS

We have implemented the coder and decoder, and an actual bit stream is generated by the coder. Note that for all experiments we generated a compressed file with a size equal to the targeted budget. We present the results of the multi-layer compression algorithm, using the following four test images: 512×512 "Barbara", and 512×512 "Houses". These images are difficult to compress because they contain a mixture of large smooth regions, and long oscillatory patterns. In order to evaluate the performance of our algorithm, we compared it to one of the best wavelet coder that was available to us: the SPIHT wavelet coder of Amir Said and William A. Pearlman.[12] A comparison with other wavelet coders (e.g.[13,14,17]) would result in different but comparable results. The performance of the algorithm is summarized in Tables 1, and 2. We work with 8 bit images, and we define the Peak Signal to Noise Ratio (PSNR) of the compressed image I_c as $PSNR = 10\ \log_{10} \frac{255^2}{\frac{1}{N^2}\sum_{i,j=0}^{N-1} |I(i,j) - I_c(i,j)|^2}$.

Figure 5. Barbara, compression ratio: 32. Left: SPIHT (wavelet) . Right: Multi-layer.

Figure 6. Detail on the right leg of Barbara, compression ratio: 32. Left: SPIHT (wavelet) . Right: Multi-layer.

Figure 7. Houses, compression ratio: 32. Left: SPIHT (wavelet) . Right: Multi-layer.

Barbara

Fig.2 (left) shows the original image Barbara. Fig.5 (left) shows the result of a compression of 32 using SPIHT, and on the right is the result of the multi-layer coder for the same compression of 32. In this example we used two layers: – a first compression with wavelets, and a second compression with an adapted local cosine basis. To better evaluate the visual quality of the compression, we have magnified a detail of the image: the right leg of Barbara. This detailed view is shown in Fig. 6. We note that the texture on the pants of Barbara is very well preserved with the multi-layered representation. Furthermore, the multi-layer coder does not introduce unpleasant ringing artifacts. A quantitative comparison of SPIHT and the multi-layer coder is provided in table 1. The multi-layer coder clearly outperforms SPIHT on the image Barbara, both in terms of PSNR and visual quality.

Houses

Fig. 7 (left) shows the result of a compression of 24, using SPIHT, and on the right is the result of the multi-layer coder at the same compression rate. In this example we used two layers: – a first compression with wavelets, and a second compression with an adapted wavelet packet basis.[9] We notice in Fig. 7 that the multi-layer coder has kept all the details on the shutters, that have been erased by SPIHT. A quantitative comparison of SPIHT and the multi-layer coder is provided in table 3. In this case, the multi-layer outperforms SPIHT on the image Houses in terms of visual quality.

7. DISCUSSION AND CONCLUSION

One can interpret this new coding algorithm in terms of a multi-layered representation of the image: Each residual $\hat{R}^i$ constitutes one set – or layer – of features that share the same structural properties. A layer is defined with respect to the choice of a particular basis. For instance

- the wavelet basis defines the layer of smooth regions,
- the wavelet packet defines the layer of textured patterns,
- the local trigonometric transforms define the localized textured features, etc.

Barbara				Houses			
Rate (bpp)	Compression	SPIHT	Hybrid	Rate (bpp)	Compression	SPIHT	Hybrid
1	8	36.41	36.58	1	8	30.84	30.44
0.67	12	33.40	33.76	0.67	12	28.07	27.88
0.5	16	31.39	31.97	0.5	16	26.15	26.30
0.4	20	30.10	30.70	0.4	20	25.06	25.27
0.333	24	29.13	29.73	0.333	24	24.33	24.43
0.286	28	28.27	28.94	0.286	28	23.75	23.75
0.25	32	27.58	28.27	0.25	32	23.17	23.24
0.20	40	26.65	27.25	0.20	40	22.33	22.46
0.154	52	25.79	26.19	0.16	50	21.65	21.76
0.125	64	24.86	25.37	0.125	64	20.98	20.95
0.10	88	24.25	24.61	0.10	88	20.37	20.33

The algorithm encodes iteratively all those features that have a strong correlation with the basis elements during this iteration. Because each layer has a sparse representation in the associated basis, the superposition of layers achieves a very compact coding.

7.1. Relation to other work

It is possible to draw some connections between our algorithm and other related ideas. There are several different methods that are related to the multi-layered coding technique.

Best Orthogonal Basis. This is the original best basis algorithm developed by Coifman and Wickerhauser in.[6] If the signal is composed of highly non orthogonal components, then the method may not yield a sparse representation. As explained previously, if the image is composed of a mixture of libraries, then the best-basis will not provide a sparse representation.

Matching pursuit. This technique was developed by Mallat and Zhang in order to provide an adaptive representation of signals. The matching pursuit algorithm is a greedy algorithm that selects at each iteration the waveform that best correlates with a large “library” of waveforms. Matching pursuit has a myopic view, and therefore cannot select a set of features all at once. As opposed to the best basis algorithm, the library need not be composed of orthogonal “atoms”, and thus the final signal representation is not constructed with orthonormal waveforms. With regards to our goal: compression, the algorithm may therefore yield a representation that is redundant.

Basis Pursuit. This technique was developed by Chen and Donoho to provide a representation of a signal with the minimum l^1 norm of the coefficients. Basis pursuit has shown to be very useful to obtain very sparse representations of signals. Unfortunately, unlike the best basis algorithm, Basis Pursuit cannot be applied to real time signal processing.[3]

7.2. Future work

The question of how many residuals, and how much budget (how many bits b_i) should be allocated to each residual R^i remains open. At the moment our approach consist in using two layers: wavelets, and wavelet packets or local trigonometric transforms. We use an exhaustive search to find the optimal allocation of the budget between the two layers. In the future we intend to address this problem with a comprehensive methodology, based on a variational approach, that will provide a clear understanding on how to adjust these parameters.

7.3. Conclusion

We have addressed the problem of efficiently coding images that contain a mixture of smooth and textured features. We have shown that a new solution to the image coding problem can be provided by “multi-layered” representations. Any image can be broken into a sequence of layers:

- the layer of smooth regions, described with wavelet bases,
- the layer of textured patterns, described with wavelet packets,
- the layer of oscillatory patterns, described with local trigonometric transforms

The evaluation of the algorithm indicates that this new coder outperforms the best wavelet coding algorithms,[12,13] both visually and in term of the quadratic error.

REFERENCES

1. P. Auscher, G. Weiss, and M.V. Wickerhauser, *Local sine and cosine bases of Coifman and Meyer*, Wavelets-A Tutorial, Academic Press, 1992, pp. 237–56.
2. T. Chang and C.C.J. Kuo, *Texture analysis and classification with tree-structured wavelet transform*, IEEE Trans. Image Processing **2,(4)** (1993), 429–441.
3. S.S. Chen, *Basis pursuit*, Ph.D. thesis, Stanford University, Dept. of Statistics, November 1995.
4. R.R. Coifman and Y. Meyer, *Remarques sur l'analyse de Fourier à fenêtre*, C.R. Acad. Sci. Paris I (1991), 259–61.
5. ______, *Size properties of wavelet packets*, Wavelets and their Applications (Ruskai et al, ed.), Jones and Bartlett, 1992, pp. 125–150.
6. R.R. Coifman and M.V. Wickerhauser, *Entropy-based algorithms for best basis selection*, IEEE Trans. Information Theory **38** (1992), no. 2, 713–718.
7. J. Li, P.Y. Cheng, and C.C.J. Kuo, *An embedded wavelet packet transform technique for texture compression*, SPIE Vol 2569, 1995, pp. 602–613.
8. S. Mallat and Z. Zhang, *Matching pursuits with time-frequency dictionaries*, IEEE Trans. on Signal Processing **41(12)** (1993), 3397–3415.
9. F.G. Meyer, A.Z. Averbuch, and J-O. Strömberg, *Fast adaptive wavelet packet image compression*, IEEE Transactions on Image Processing (2000), 792–800.
10. F.G. Meyer and R.R. Coifman, *Brushlets: a tool for directional image analysis and image compression*, Applied and Computational Harmonic Analysis **4** (1997), 147–187.
11. Yves Meyer, Personnal communication, 1997.
12. A. Said and W.A. Pearlman, *A new fast and efficient image codec based on set partioning in hierarchical trees*, IEEE Trans. Circuits. Syst. Video Tech. **6** (1996), 243–50.
13. J.M. Shapiro, *Embedded image coding using zerotrees of wavelet coefficients*, IEEE Trans. on Signal Processing **41** (1993), no. 12, 3445–3462.
14. P. Sriram and M.W. Marcellin, *Image coding using wavelet transforms and entropy-constrained treillis quantization*, IEEE Trans. Image Processing **4** (1995), 725–733.
15. M.V. Wickerhauser, *Adapted wavelet analysis from theory to software*, A.K. Peters, 1995.
16. I.H. Witten, R.M. Neal, and J.G. Cleary, *Arithmetic coding for data compression*, Communications of the ACM **30,6** (1987), 520–540.
17. Z. Xiong, K. Ramchandran, and M.T. Orchard, *Space-frequency quantization for wavelet image coding*, IEEE Trans. Image Process. **6** (1997), no. 5, 677–693.