ADAPTIVE CODING OF IMAGES VIA MULTIRESOLUTION ICA

Doru C. Balcan

Computer Science Department
Carnegie Mellon University
Pittsburgh, PA
dbalcan@cs.cmu.edu

Michael S. Lewicki

Electrical Engineering and Computer Science Dept.
Case Western Reserve University
Cleveland, OH
michael.lewicki@case.edu

ABSTRACT

Multiresolution (MR) representations have been very successful in image encoding, due to both their algorithmic performance and coding efficiency. However these transforms are fixed, suggesting that coding efficiency could be further improved if a multiresolution code could be adapted to a specific signal class. Among adaptive coding methods, independent component analysis (ICA) provides the best linear code by finding a linear transform with maximally independent coefficients, given a specific signal distribution. This technique, however, scales poorly with the dimensionality of the data, and has been ill-suited for large-scale image coding. We propose a hybrid method (multi-resolution ICA) which derives an ICA basis for each subband space produced by a given MR transform over the image class. We find that this method produces a significantly more efficient code compared to the MR transform alone. We provide both quantitative and qualitative assessments of coding performance, and illustrate improvement over standard (i.e., non-adaptive) wavelet-based representations such as that used in JPEG2000.

Index Terms— Image coding, Wavelet transforms, Adaptive coding, Unsupervised Learning

1. INTRODUCTION

The problem of efficiently describing visual structure has been of great importance to many research fields, spanning from biology to engineering (see *e.g.*, [1, 2]). Best existing coders (notably JPEG2000 [3]), rely on the flexibility of multiresolution (MR) transforms to capture image structure, by exploiting the intrinsic multiscale character of (natural) images [4, 5]. In spite of their success, wavelets have well known limitations in terms of modeling or detecting the two-dimensional, sharp, arbitrarily oriented (ridge-like) discontinuities. Various types of MR representations emerged in the past decade in computational harmonic analysis, which provably outperform wavelets in approximating particular classes of signals (for an example, see [6]). Because of their great diversity, it is not clear what makes an optimally efficient code for images. Common intuition that optimal image features are smooth surfaces and short straight edges may be accurate for some classes (*e.g.*, natural scenes [4]), but not for others (faces, textures, cartoons, fingerprints, medical images of all sorts).

Separating signal content into different subbands, and concentrating the relevant information into a small set of non-zero coefficients, seems a natural recipe for achieving efficiency. However, a representation is inherently suboptimal unless it can capture the probability density of the data, according to Shannon's source coding theorem. As such, optimal efficiency can only be achieved by adapting the representation to the statistical structure of the target image class. When searching for the "most compact" code, one method to employ is independent component analysis (ICA) [7]. Generally speaking, the goal of ICA is to derive a data dependent linear mapping such that the coefficients in this new representation are maximally independent. Therefore, a suitable mathematical cost to minimize is the mutual information among coefficients. Due to its poor computational scalability with respect to data dimensionality, ICA has been traditionally applied to images (either for analysis, encoding, or denoising) by extracting relatively small image patches to be used as training samples, followed by block-transforming the image. Unfortunately, the arbitrary alignment of the blocks with the image and the insufficient capacity to represent image structure spread across blocks produce artifacts at reconstruction.

In this paper, we propose an ICA-like image representation, which overcomes the artificial block confinement and computational obstacles. Our method consists of a preliminary MR (*e.g.* wavelet) decomposition step, followed by learning an ICA basis for each of the resulting subbands. The purpose of the MR step is to allow easier access to structure at each scale, while reducing the bulk of image information to the coarsest scale; indirectly, this help in eliminating blocking artifacts. Since the learned ICA bases provide the most compact linear code for each subband, we can conclude that this hybrid Multiresolution-ICA procedure (henceforth referred as MrICA) gives an improvement over both types of representations.

Efficient coding has been a very suitable paradigm in attempting to explain how biological systems cope with processing complex information. The resemblance of the optimally derived linear features learned from natural scenes to the receptive fields of simple cells in primary visual cortex (V1) has led to very interesting hypotheses about the role and function of the brain's sensory systems [1, 8]. A probabilistic modeling approach aimed directly at optimal efficient coding of natural images [9] has revealed that the average entropy improvements of adaptive linear representations over fixed ones (Fourier, DCT, wavelets, Gabor functions) are too important to neglect. However, due to the computational constraints, their representation was derived for relatively small image patches and thus a comparison to multiscale bases was limited. We can mention here two other block-based ICA approaches to image compression [10, 11]. Both compare favorably to JPEG (for faces, or natural images), and the first one even outperforms the WSQ multiscale coder for fingerprint images at low rates; however, they do not exploit the potential of multiresolution. Modeling subband information statistically for image coding has been performed in [12]. Their coder (EPWIC) explicitly exploits statistical relationships between coefficients in different subbands via a parameterized model. Another parametric approach is the adaptive multiscale method presented in

[13], where the objective was to adapt the parameters of a certain wavelet-based transform, to better fit natural images. In contrast, we derive an adaptive non-parametric multiscale image representation.

Structure of the paper. In Section 2, we describe the main constituents of our image encoder. Section 3 describes in detail the experimental results illustrating the encoding performance of the proposed adaptive method, while the last section concludes the paper.

2. ADAPTIVE MULTIRESOLUTION CODING

In this section, we describe the proposed method for MR adaptive image encoding. For background on the main tools we point the reader to [2] (multiresolution transforms and subband coding), [7] (ICA), and [3, 14] (quantization and coding).

We shall start by assuming we have a set of images drawn from a common class. We shall decompose these images by a fixed MR transform and then, for each of the resulting subband spaces we shall learn an ICA basis (using the subband coefficient sets of all the images in the sample, as training data). Finally, we shall use the subband ICA coefficients to design a quantizer. Every new image will then be transformed and quantized, and finally output into a bitstream via an arithmetic coder. In the following, we provide details on each of the modules of our system.

Multiresolution Transforms. The first step of our hybrid method aims at separating image content by projecting images on scale-orientation subbands. The most widely used MR transforms are based on the wavelet decomposition due both to their theoretical and just as importantly, to their computational properties. JPEG2000 itself (Part 1) uses the Cohen-Daubechies-Feauveau 9/7 biorthogonal filters [15], as its only supported "irreversible" wavelet transform [3]. We also chose to employ this wavelet because we wanted test the coding efficiency of our method against that of the most common fixed MR transform. To keep more similarities with existing image coders, we applied this separable decomposition method, using whole-point symmetric edge handling. The implementation we used in our experiments was that of Matlab Wavelet Toolbox 4.2.

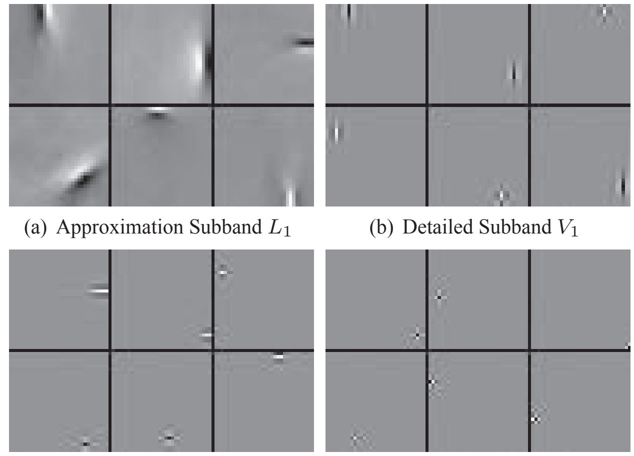
Adaptation. In unsupervised learning, the problem of separating signals into independent linear components can be formulated as follows: given a set of N-dimensional vectors $(\mathbf{y}^{(k)})_{1\leq k\leq K}$, search for a linear transform $\mathbf{A}$ such that the observed vectors are linear mixtures (induced by $\mathbf{A}$) of realizations of an M-dimensional random vector $Z = (z_1, \ldots, z_M)^T$ whose components are as independent as possible. We can express this model compactly as:

$$\mathbf{Y} = \mathbf{A}\mathbf{Z} \qquad (1)$$

where $\mathbf{Y} \in \mathbb{R}^{N\times K}$, $\mathbf{A} \in \mathbb{R}^{N\times M}$, and $\mathbf{Z} \in \mathbb{R}^{M\times K}$. The ICA objective is then to find $\mathbf{A}$, such that the mutual information among the coefficients z_i is minimized. To simplify the description, we will assume that $\mathbf{A}$ is square and invertible (that is, $M = N$) and if we denote its inverse by $\mathbf{W}$, the problem is reduced to minimizing:

$$I(z_1, \ldots, z_N) = \sum_{j=1}^{M} H(z_j) - H(Y) - \log|\det \mathbf{W}| \qquad (2)$$

Since the entropy of the observed mixture Y is constant, imposing that $|\det \mathbf{W}| = 1$ results in the quantity we seek to minimize to be the marginal entropy sum; that is, ICA searches for the transformation giving the (potentially) most compact linear code of the data. By interpreting the ICA objective as maximizing the (log-)likelihood of the data under the linear model, or as minimizing the Kullback-Leibler divergence between the joint probability and the product of marginals, has produced several families of algorithms for ICA (see [7, 16]). For the results in this paper we preferred the Relative Trust-Region algorithm [17], due to its high flexibility and robustness.



(a) Approximation Subband L_1 (b) Detailed Subband V_1

(c) Detailed Subband H_1 (d) Detailed Subband D_1

Fig. 1. Basis functions computed for MrICA (32x32, L=1) log-scale. For each subband, a random set of basis functions are shown.

Quantization and Coding. Next, we shall describe the subband coding procedure employed to transform the coefficients into bitstreams, for both the wavelets and MrICA. For a group of images from the training set, we group the MR coefficients belonging to the same subband and from the whole group, we estimate a scalar quantizer. Note that scalar quantization is justified in the case of MrICA, since coefficients within each subband are as independent as possible. To design the subband quantizers, individual bit rates are allocated according to the relative energy within each subband. Since we are interested in the potential improvement of the adaptive representation, and less so in the great many practical issues of image coding, we will compute the "optimal" entropy-constrained scalar quantization [14] for each subband. This should provide a reliable *upper bound* for the performance of each representation. After quantizing the coefficients, we use Matlab Communication Toolbox's arithmetic coder to construct the bistreams and record the total bitstream length and the reconstruction SNR for each test image. Then, we take the average over the whole test set to estimate the coding efficiency of the distribution. We repeat this procedure for various target rates, and by interpolation we construct the rate-distortion curve.

3. EXPERIMENTAL RESULTS

We shall illustrate the encoding performance of MrICA by plotting the average rate-distortion curves generated by coefficient quantization at various levels of precision, for both the fixed and the adaptive MR transforms. First of all however, let us comment on the features learned by our method, when applied to natural images.

We applied MrICA to encoding natural scenes randomly cropped from van Hateren's database of natural stimuli [18]. We tested our method on images of two sizes, 32×32 and respectively, 64×64 pixels[1]. Instead of working with the pixel intensities, we took the

[1] As they are similar to JPEG2000 standard code blocks sizes, we consid-

logarithm of these intensities before any further processing; as explained in [18], the reasons for this operation are to incorporate contrast invariance of natural scenes, get better first-order statistics of the natural image data, and better mimic the operations performed by the first stages of visual systems. On each of these logarithmically transformed images, we applied the discrete wavelet decomposition and learned the subband ICA matrices, as described in the previous section. Figure 1 displays a random set of such ICA basis functions learned from the 32×32 data set with one decomposition level. MrICA basis functions of the approximation subband retain the aspect of classic image ICA basis functions (relatively low spatial frequency, all orientations) [18, 8], not a surprise considering that the approximation subband contains a low-resolution version of the original image. The detailed subbands basis functions look like localized features, maintaining the dominant orientation of the subband. Besides the quantitative (mutual information) difference between the MrICA detailed bases and corresponding wavelets, we note that the adaptive features are also more diverse in shape.

Next, we illustrate the improvement in coding efficiency afforded by adaptiveness. In addition to comparing the adaptive and non-adaptive MR methods described in the previous section, we also compare both of them against JPEG2000; for this purpose, we used Jasper[19], a software package implementing JPEG2000. The coding cost of the adaptive and non-adaptive representations does not include the basis functions (respectively, the wavelets), as these can rightfully be considered part of the coder, and not of the code; also, in case of Jasper, we report only the codestream (*i.e.*, not including the metadata). In the case of MrICA, the images included for the evaluation are taken from the testing set, that is, they belong to the same signal class as those in the training set, but have not been used during learning. Let us point out that the JPEG2000 performs quantization for each individual image, and not over a whole sample set, unlike our method. In this respect, our quantizers take advantage of more information. On the other hand, JPEG2000 performs surprisingly well considering that we used an optimal ECSQ, and not a uniform one. The rate-distortion trade-off obtained for the 32×32 test images, with one MR level, for the three encoding methods, are presented in Figure 2. The top plot shows the relative coding gain, while the bottom plot shows the relative rate difference of the three methods, taking the non-adaptive wavelet representation as reference. The better coding efficiency of MrICA (more apparent at low bit rates) has two important consequences, namely the same distortion (or SNR) can be achieved by the adaptive method for a significantly lower rate (*i.e.*, bit cost), and reciprocally, for the same bit rate we can get a significantly better improvement in fidelity.

Since SNR is not a relevant measure of perceptual distortion, evaluating the representational power of MrICA should involve assessing presence of reconstruction artifacts. For this purpose, we chose to display several test images from the two datasets, their encoded version via the adaptive and non-adaptive MR transforms, and the residual errors[2]. Figure 3 illustrates the encoding results of five examples from the 32×32 dataset. Each image has been encoded to a quality of 25dB; the coding gain of MrICA over the non-adaptive wavelet method for these images is of 0.62 bpp, 2.91 bpp, 2.78 bpp, 3.39 bpp, and 2.69 bpp. Figure 4 shows three examples of 64×64 images encoded at 20dB. The coding gain of the adaptive method was in this case of 1.48 bpp, 0.33 bpp, and 0.23 bpp. (For both figures, the colormaps were maximally stretched to enhance visibility.) As a general conclusion, MrICA obtains a better coding rate than the fixed wavelet representation, with less reconstruction artifacts.

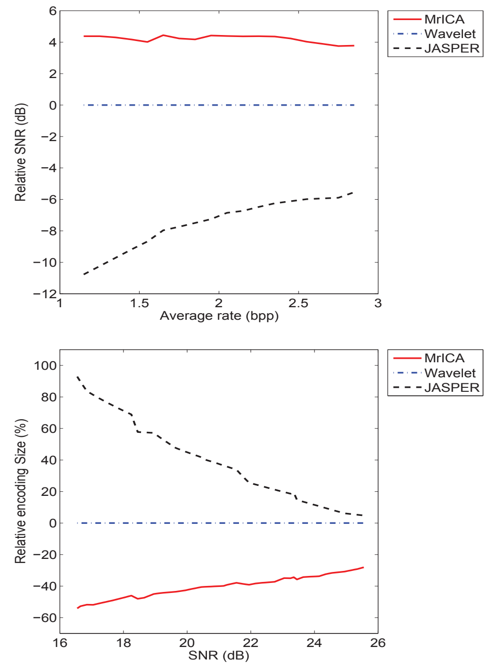


Fig. 2. Relative rate-distortion performance of three methods (MrICA, wavelet, JASPER) computed for the 32×32 test images.

4. CONCLUDING REMARKS

We proposed MrICA, a hybrid adaptive multiresolution method that combines the advantages of the two families of representations. We illustrated the significant coding efficiency gain of MrICA over the wavelet transform when applied to natural images, which is explained by the ability of the new method to adaptively describe image structure at all scales. This suggests that a image coder devoted a given class of signals should use not only multiresolution, but also adaptivity to optimize encoding performance.

5. REFERENCES

[1] B. A. Olshausen and D. J. Field, "Sparse coding with an overcomplete basis set: A strategy employed by V1?," *Vision Res.*, vol. 37, pp. 3311–3325, 1997.

[2] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*, Prentice Hall, 1995.

[3] M. Marcellin D. Taubman, Ed., *JPEG2000: Image Compression Fundamentals, Standards and Practice*, Springer, 2001.

[4] D. J. Field, "Relations between the statistics of natural images and the response profiles of cortical cells," *J. Opt. Soc. Am. A*, vol. 4, pp. 2379–2394, 1987.

[5] S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans. PAMI*, vol. 11, no. 7, pp. 674–693, July 1989.

[6] E. J. Candès and D. L. Donoho, "New tight frames of curvelets and optimal representations of objects with piecewise $\mathcal{C}^2$ singularities," *Comm. Pure Appl. Math.*, vol. 57, no. 2, 2004.

ered these image sizes relevant to use for comparison purposes.

[2] To evaluate this appropriately, we recommend zooming in on the pdf document, rather than on the printed version. Also, due to lack of space we limit the number of examples included in the paper; for more, we invite the reader to `http://www.cs.cmu.edu/~dbalcan/mrica/`.

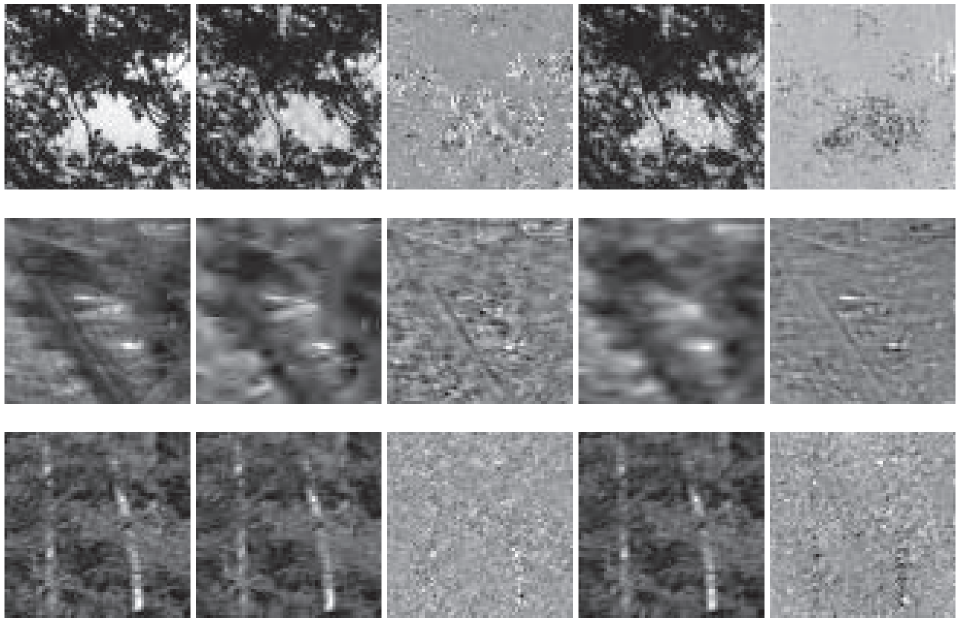


Fig. 4. Examples of 64×64 images reconstructed at 20dB. Column 1.Original images; 2.Image encoded by non-adaptive method; 3.Error; 4.Image encoded by MrICA ; 5.Error

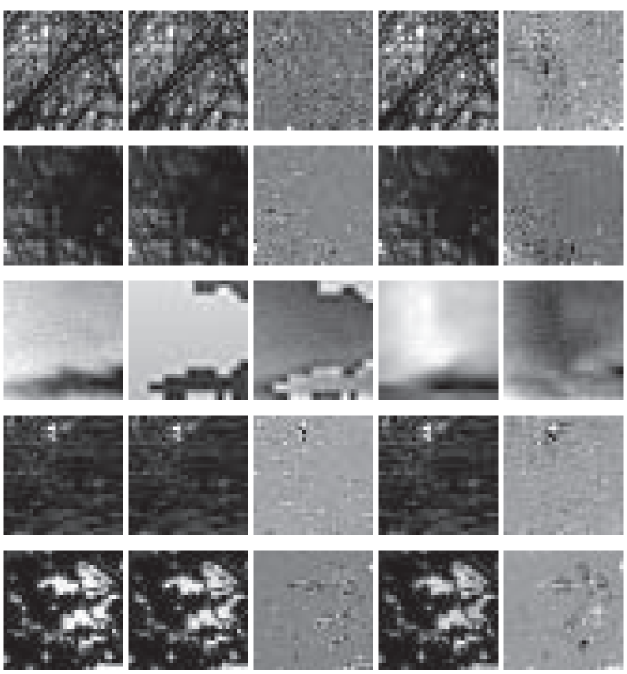


Fig. 3. Examples of 32×32 images reconstructed at 25dB. Column 1.Original images; 2.Image encoded by non-adaptive method; 3.Error; 4.Image encoded by MrICA ; 5.Error

[7] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley & Sons, 2001.

[8] A. J. Bell and T. J. Sejnowski, "The independent components of natural scenes are edge filters," *Vision Res.*, vol. 37, pp. 3327–3338, 1997.

[9] M. S. Lewicki and B. A. Olshausen, "A probabilistic framework for the adaptation and comparison of image codes," *JOSA A*, vol. 16, no. 7, pp. 1587–1601, 1999.

[10] A. J. Ferreira and M. A. T. Figueiredo, "Class-adapted image compression using independent component analysis," in *Proc. IEEE ICIP*, 2003, pp. 625–628.

[11] M. Narozny and M. Barret, "Ica-based algorithms applied to image coding," in *Proc. IEEE ICASSP*, 2007, pp. I–1033 – I–1036.

[12] R. W. Buccigrossi and E. P. Simoncelli, "Image compression via joint statistical characterization in the wavelet domain," *IEEE Trans. Image Proc.*, vol. 8, no. 12, pp. 1688–1701, 1999.

[13] B. A. Olshausen, P. Sallee, and M. S. Lewicki, "Learning sparse images codes using a wavelet pyramid architecture," in *Adv. NIPS*, 2000.

[14] N. Farvardin and J. Modestino, "Optimal quantizer performance for a class of non-gaussian memoryless sources," *IEEE Trans. Inform. Th.*, vol. 30, no. 3, pp. 485–497, 1984.

[15] I. Daubechies, *Ten Lectures on Wavelets*, Number 61 in CBMS/NSF Series in Applied Math. SIAM, 1992.

[16] A. Cichocki and S. i. Amari, *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*, J. Wiley&Sons, 2002.

[17] H. Choi and S. Choi, "A relative trust-region algorithm for independent component analysis," *Neurocomputing*, vol. 70, no. 7–9, pp. 1502–1510, 2007.

[18] J. H. van Hateren and A. van der Schaaf, "Independent component filters of natural images compared with simple cells in primary visual cortex," *Proc. R. Soc. Lond. B*, vol. 265, pp. 359–366, 1998, (data available online at `http://hlab.phys.rug.nl/imlib/`).

[19] R. K. Ward M. D. Adams, "JASPER: A portable flexible open source software tool kit for image coding/processing," in *Proc. ICASSP*, 2004, `http://www.ece.uvic.ca/~mdadams/jasper/`.