# IMAGE COMPRESSION WITH GENERALIZED LIFTING AND PARTIAL KNOWLEDGE OF THE SIGNAL PDF

*Julio C. Rolón*†‡ *, Philippe Salembier*† *and Xavier Alameda*†

†Technical University of Catalonia (UPC), Dept. of Signal Theory and Communications, Spain
‡National Polytechnic Institute (IPN), CITEDI Research Center, Mexico
{jcrolon, philippe}@gps.tsc.upc.edu, Xavier.Alameda@estudiant.upc.edu

## ABSTRACT

In this paper we deal with the use of Generalized Lifting (GL) for lossy image compression. We have demonstrated in [9] the potential of the method for coding assuming complete knowledge of the pdf of the image to encode. Here, we move towards a realistic scheme that does not assume complete knowledge of the pdf. We show that a multiscale GL produces interesting results even if the pdf of the image to encode is only partially known. We target the compression of a given image class and compute an estimate of the image class pdf. This pdf is available at both encoder and decoder. A decision algorithm minimizes the overhead produced by the difference between the class pdf and the image pdf. This algorithm also removes ambiguities in the decoding process. The encoding strategy is completed using an arithmetic encoder. Results exhibit improvements over the state of the art.

***Index Terms***— Generalized lifting, wavelets, image coding, pdf estimation, nonlinear lifting

## 1. INTRODUCTION

Wavelet-based image coding algorithms are at the core of the most advanced image coders available today. However, one of the known limitations of the wavelet transform when applied to signals with dimensions higher than one, is their inability to deal with higher order singularities, as is the case with contours in images. Contours appear in the detail subbands as correlated coefficients of large magnitude. From a signal processing perspective, decorrelating the contours is equivalent to an improvement in the sparsity of the whole coefficient set.

Two main streams of methods have been followed recently to improve the sparsity of image representations. On one hand, we find methods that operate in the frequency domain like Curvelets [1] and Contourlets [6]; these methods have proven to be efficient in tasks other than coding, as they produce a large amount of redundancy, a pitfall for compression.

The second stream of ideas is conceived to operate in the spatial domain. Among these methods, we find Bandelets [8] and adaptive directional lifting methods [2-5,7,12,13]. A common characteristic of all these methods is that they rely on predictions and interpolations over the pixel grid, be it Cartesian or quincunx. In the case of bandelets, the method may be applied in the wavelet domain, an idea we follow in the present paper, as it allows us to benefit from previous decorrelation of wavelet transform.

A different approach to perform additional decorrelation on wavelet coefficients has been presented in [9]; it relies on the use of generalized lifting [10,11]. The generalized lifting framework allows the definition of nonlinear filter banks with perfect reconstruction. In [9] we presented results on the application of generalized lifting (GL) method to lossy image coding. We followed the idea of bandelets and conducted a wavelet transform over the image before applying the GL; we introduced a scalar quantizer; we used 2-dimensional non-separable sampling pattern for the GL, and made the assumption that the pdf of the image to code was known at the encoder and the decoder ends. This assumption is of course unrealistic in practice but allowed us to study the potential of the approach. The results shown in [9] were very encouraging.

The main contribution of this paper is to move towards a realistic coding scheme where the assumption is that the pdf of the image to code is not known. In this paper, we investigate possible mechanisms to compensate for the lack of knowledge of the pdf, a fundamental element used in the design of the GL. In this paper, we target the compression of images belonging to a given class (as an example we use a class of remote sensing images). We estimate through a training process the image class pdf and use it instead of the pdf of the current image to encode. As a result of the use of a suboptimal pdf, information overhead is produced. We introduce a decision algorithm and a buffer to manage this overhead in an efficient way. Interesting results are obtained with this realistic scheme when compared against JPEG 2000.

The paper is organized as follows. Section 2 presents the generalized lifting method and its multiscale approach. In section 3 the coding strategy is described, as well as details about pdf estimation and managing in coding. Experimental results follow in section 4. Conclusions and ideas for future work are presented at the end.

## 2. GENERALIZED LIFTING

The generalized lifting decomposition shown in fig. 1 was introduced in [10]. The GL decomposition scheme enables the implementation of linear and non-linear operations. The GL involves first a polyphase decomposition or Lazy Wavelet Transform (LWT), followed by generalized predict (P) and update (U) steps.

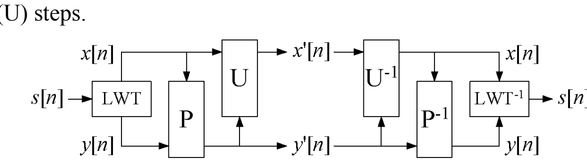



**Fig. 1.** Generalized Lifting Scheme

Let us analyze the generalized predict operator P. In the classical lifting case, P takes $x[n]$ as input in order to predict $y[n]$ and the details signal $y'[n]$ is seen as a prediction error. In the GL case, P is viewed as a mapping between $y[n]$ and $y'[n]$ that takes into account a context represented by samples from $x[n-i]$ for $i \in C$, $C$ being the set of sample positions that constitutes the context. Formally, the generalized predict operation can be written as

$$y[n]\Big|_{with\ context\ \{x[n-i]\}_{i\in C}} \xrightarrow{P} y'[n] .$$

Assuming discrete signals, the mapping itself is discrete. To get perfect reconstruction, the mapping P should be invertible, that is, it should be an injective mapping. If the number of possible values for $y[n]$ and $y'[n]$ is the same, then the mapping should be bijective. The same reasoning can be done for the generalized update operation. That is, it is an injective mapping

$$x[n]\Big|_{with\ context\ \{y'[n-i]\}_{i\in C}} \xrightarrow{U} x'[n] .$$

Apart from the injectivity that is required to get perfect reconstruction, the generalized predict (P) and update (U) operators may be arbitrary. In [11], a design of P that minimizes the energy of the details coefficients is proposed. The design can be intuitively described as follows: assume that the pdf of $y[n]$ conditioned on the context $\{x(n-i)\}_{i\in C}$ is known. For any given context $\{x(n-i)=x_i\}_{i\in C}$, in order to minimize the energy of $y'[n]$, the most probable value of $y[n]$ should be mapped to $y'[n]=0$. The next most probable value of $y[n]$ should be mapped to $y'[n]=1$; the next one to $y'[n]=-1$, etc. As can be seen, values of $y[n]$ of decreasing probability are successively assigned to values of $y'[n]$ of increasing energy. Note that this design of the generalized predict does not rely on the exact pdf values but on the order of the pdf values. So two different pdfs with the same order of their pdf value will produce the same mapping.

We have introduced the use of the generalized lifting operator for lossy image coding in [9]. The GL was based on a unique generalized predict P (in particular, no generalized update was used). The mapping criterion of P in [9] was the minimization of the energy of the wavelet coefficients. A 2-dimensional non-separable context that involved 4 samples was used. The assumption in [9] was that the pdf of the image to encode was known at the coder and the decoder. So the generalized predict was designed on the basis of this pdf. In this paper, we assume that the pdf of the image is not known; our approach here is to use a reference pdf which remains unchanged during the coding process. Our goal is to demonstrate that GL-based image coding is possible even when an approximate pdf is used instead of the exact pdf of the image.

## 3. CODING SCHEME

In this paper, we propose a lossy coding algorithm that does not assume knowledge of the pdf of the image to be coded. Instead, we assume the image to encode belongs to a given class of images and that we have some estimate of the image class pdf. The GL encoder is illustrated in fig. 2. The GL operation is performed in the wavelet domain. We use a well known linear-phase biorthogonal wavelet (9/7 or CDF 9/7 after Cohen-Daubechies-Feauveau [4]) to perform a 5-scale discrete wavelet transform (DWT), this wavelet filter is also used by JPEG 2000 lossy compression mode.

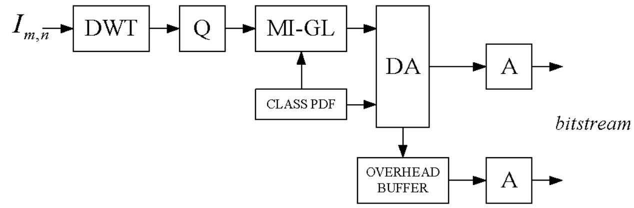



**Fig. 2.** Coding Scheme

As an initial and simple strategy, a uniform scalar quantizer is applied before the GL operator to preserve the bijectivity of the GL decomposition. A 2-dimensional non-separable quincunx sampling pattern is used. The four closest neighboring pixels compose the context while the central sample is the value to be mapped. To fully exploit the energy minimization potential of the mapping, we iterate the GL decomposition over the approximation signal to generate a multiscale GL decomposition. We reach the condition where, for a given wavelet subband, only one wavelet coefficient of this subband is left as the GL approximation coefficient. We define this condition as maximally-iterated (MI) GL decomposition.

In our proposal, we use the pdf of the image class instead of the pdf of the image to encode. The class pdf is obtained through a training process, and once trained; the class pdf remains unchanged and known for both the coder and the decoder. During the coding process, a decision algorithm (DA) handles the overhead produced by the disparity between the class pdf and the actual pdf of the image to encode. The GL mapped coefficients as well as the overhead information are encoded with an arithmetic encoder.

### 3.1 Training of the class pdf

In the ideal case [9], both the encoder and the decoder knew exactly the pdf of the image to encode. The P mapping under this condition is ideal since the energy minimization is optimal. In this paper, we remove this assumption but assume that the image to encode belongs to a given image class. Through a training process, we estimate the class pdf. For this pdf estimation, the simplest possible strategy is used: we define a set of training images that should represent the image class and compute the histogram of $y[n]$ samples conditioned by the context $C$. The class pdf is defined as the normalized (so that the sum of probabilities is equal to one) histogram of the class. As mentioned in section 2, two pdfs having different values but the same order of their pdf values will give the same mapping. So here we hope that the order of the class pdf values will be statistically close to the order of the pdf values of the image to encode.

In the proposed coding algorithm we have the combination of two decompositions: wavelet and GL. The statistical behavior depends on the combination wavelet/subband-scale and GL-scale. For this reason, the training process is conducted locally at every wavelet subband and at every GL decomposition scale for every wavelet subband.

### 3.2 Decision algorithm

The operation of the GL mapping when the exact pdf of the image is not known is more complex. The main issue is to know how to handle the cases of $y[n]$ values conditioned by $\{x(n-i)\}_{i\in C}$ for which the class pdf is equal to zero. In practice, it means that these values of $y[n]$ values conditioned by $\{x(n-i)\}_{i\in C}$ have never been seen in the training set. As a result, the mapping of these values has not been defined. In practice, we distinguish between three possible cases:

*1.* The class pdf value for $y[n] \| \{x(n-i)\}_{i\in C}$ is different from zero. In this case, a mapping has been defined:

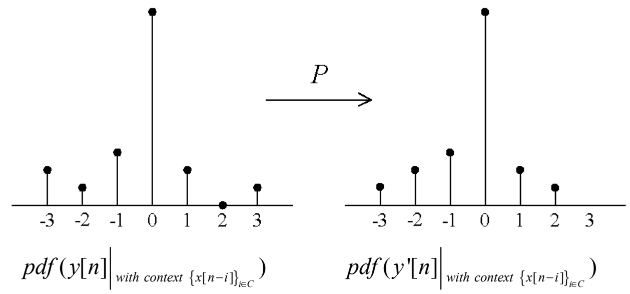
$$y[n]\Big|_{with\ context\ \{x[n-i]\}_{i\in C}} \xrightarrow{P} y'[n]\ .$$

As a result, this mapping is performed and $y'[n]$ is generated as output coefficient.

*2.* The class pdf value for $y[n] \| \{x(n-i)\}_{i\in C}$ is equal to zero and the particular context $\{x(n-i)\}_{i\in C}$ has never been seen in the training set. In this case, the original $y[n]$ is used as output value. During the decoding process, the decoder will see that the current context has never been seen in the training set and therefore, will interpret the coefficient value as the value of the unmapped $y[n]$.

*3.* The class pdf value for $y[n] \| \{x(n-i)\}_{i\in C}$ is equal to zero but the particular context $\{x(n-i)\}_{i\in C}$ has been seen in the training set (with other values of $y[n]$). In this case, no mapping of the $y[n]$ has been defined but we cannot just use the unmapped $y[n]$ value as output coefficient because the decoder will not know whether the received value is a mapped $y'[n]$ or an unmapped $y[n]$ value. To solve this problem, we use as output coefficient an *escape* code to signal the receiver that we are in this third case and store the actual $y[n]$ value in an overhead buffer that is encoded separately by an arithmetic encoder (see figure 2). The value of the escape code is context dependent. For a given context, the escape code value is defined as the output value of lowest energy that is never used for that context. To illustrate this case, consider the histogram plots of the class pdf, for a given context, shown in fig. 3. On the left (a), a plot of the instances of $y[n] \| \{x(n-i)\}_{i\in C}$ found during the training is shown. On the right (b), the mapped values that correspond to those instances are shown. Now, suppose during the encoding process we find $y[n]=2$ for this context. We see from fig. 3(a) that during the training process this case was never found, then we have $p(y[n]=2 \big| \{x(n-i)\}_{i\in C})=0$ . An *escape* code has to be sent. The decision algorithm (DA) of fig. 2 searches in the mapped domain (fig. 3b) for the first available value, in this case 3. The output coefficient value will then be 3, and $y[n]=2$ will be fed to the buffer.

At this moment, given that the class pdf is unchanged, we repeat the *escape* code as many times as necessary during the coding process even if the particular instance of $y[n] \| \{x(n-i)\}_{i\in C}$ has already been sent. In a future version of the algorithm, an adaptive pdf strategy will prevent repetitions of the *escape* code. The decision algorithm would eventually provide proper feedback to the adaptation mechanism.




(a) (b)

**Fig. 3.** Mapping operation in the class pdf for a given context

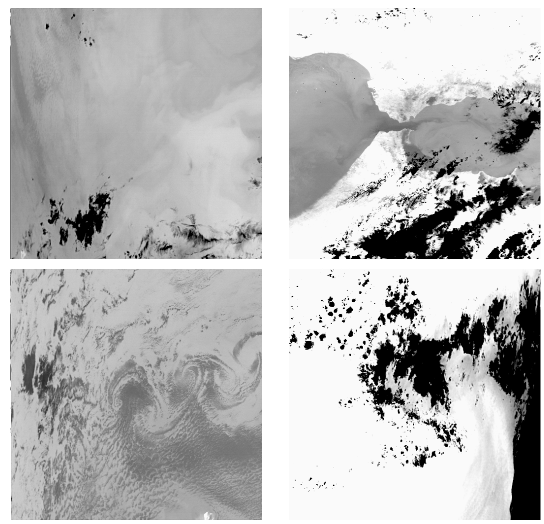


**Fig. 4**. Examples of SST images used to train the class pdf

## 4. EXPERIMENTAL RESULTS

In this section, we evaluate the R-D performance of the coding algorithm for a class of images, namely, remote sensing images of sea surface temperature (SST). To compute the class pdf, we use 100 images of 512 × 512 pixels. A CDF 9/7 wavelet was used to transform the set of images in 5 scales. A MI-GL decomposition was then applied to each wavelet subband. Fig. 5 shows the image *SST* to be coded; this image does not belong to the training set so that the pdf is not biased.

A comparison between JPEG 2000, the R-D characteristic of the *ideal* case in which the pdf of the image to encode is completely known, and the proposed codec in which only the pdf of the image class is known is shown in fig. 6. As expected, when we have complete knowledge of the pdf of the image (ideal pdf curve in fig. 6) we reach the upper performance limit of the GL algorithm. Fig. 6 shows that even when the pdf of the image is not known (class pdf curve), the use of a class pdf allows the GL coder to outperform the R-D characteristic of JPEG2000 by about 0.4 dB in the interval from 0.2 to 0.5 bpp. The performance of GL algorithm in fig.6 is typical for images of the class that do not belong to the training set. A mean coding gain of 0.3 dB at 0.25 bpp, and 0.5 dB at 0.5 bpp was obtained for a set of these images. Overhead produced by buffer ranges from 3 to 6 % of the total bitrate in the best case, to a range of 5.7 to 10 % in the worst case.

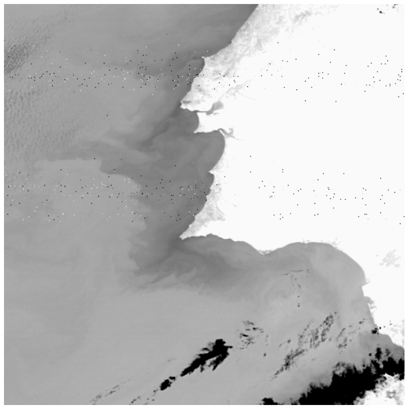


**Fig. 5.** *SST* test image

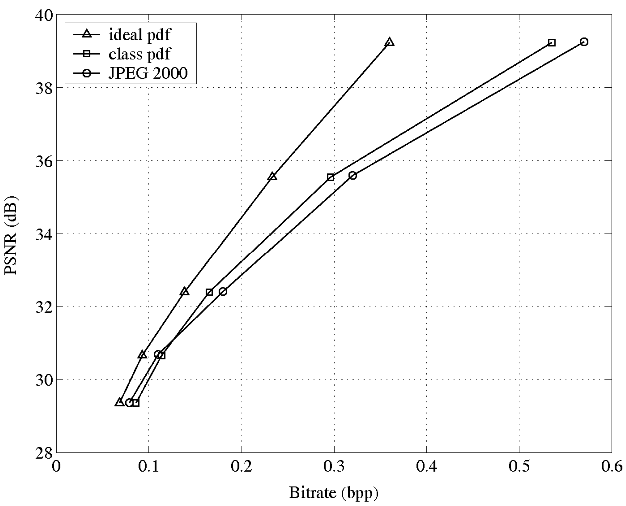



**Fig. 6.** Rate-distortion plot for *SST*

## 5. CONCLUSIONS

The main contribution of this paper is the demonstration that for a class of images, the GL method improves the R-D performance even when the pdf of the signal to encode is not known. We have implemented a multiscale, maximally-iterated GL mapping that is applied independently to all wavelet subbands; a class pdf that has been used to enable the GL operator to perform the mapping; and a decision algorithm minimizes the energy of the overhead and prevents ambiguity in the decoder. The results are very encouraging as a typical improvement over JPEG 2000 of about 0.4 dB in the interval from 0.2 to 0.5 bpp has been demonstrated, while overhead produced by buffer does not exceed 10 percent in worst case.

Of course, the scheme described in this paper has to be considered as an initial scheme on which many improvements can be studied and tested. Important issues to be tackled in the future involve 1) the study of pdf estimation strategies that do not assume that the image belongs to a given class (the pdf should ideally model the statistics of the wavelet coefficients after the initial DWT); 2) the inclusion of the quantization process within the GL; 3) the design criterion of the generalized predict P; 4) the optimization of the Decision Algorithm (DA); and 5) finally the design of an efficient entropy coder specifically devoted to coefficients resulting from a Generalized Lifting.

## 6. ACKNOWLEDGMENT

This work has been partially supported by COFAA-IPN of Mexico, and by the TEC2007-66858/TCM PROVEC project of the Spanish government.

## 7. REFERENCES


[1] Candès, E.J., *et al.*, Fast Discrete Curvelet Transforms, *SIAM Multiscale Model. Simul.*, vol. 5, no. 3, pp. 861-899, 2006.

[2] Chang, C-L., Girod, B., Direction-Adaptive Discrete Wavelet Transform Via Directional Lifting and Bandeletization, *IEEE Intl. Conf. on Image Processing, ICIP'06*, pp.1149-1152, 2006.

[3] Chappelier, V., Guillemot, C., Oriented Wavelet Transform for image Compression and Denoising, *IEEE Transactions on Image Processing,* vol. 15, no. 10, pp. 2892-2903, 2006.

[4] Cohen, A., Daubechies, I., Feauveau, J.-C., Biorthogonal Bases of Compactly Supported Wavelets, *Comm. Pure Appl. Math.* Vol. XLV, pp. 485-560, 1992.

[5] Ding, W.P., *et al.*, Adaptive Directional Lifting-Based Wavelet Transform for Image Coding, *IEEE Transactions on Image Processing,* vol. 16, no. 2, pp. 416-427, 2007.

[6] Do, M.N., Vetterli, M., The Contourlet Transform: An Efficient Directional Multiresolution Image Representation, *IEEE Transactions on Image Processing,* vol. 14, no. 12, pp. 2091-2106, 2005.

[7] Mehrseresht, N., Taubman, D., Spatially Continuous Orientation Adaptive Discrete Packet Wavelet Decomposition for Image Compression, *IEEE Intl. Conf. on Image Processing, ICIP'06*, pp.1593-1596, 2006.

[8] Peyré, G., Mallat, S., Discrete Bandelets with Geometric Orthogonal filters, *Proc. of the ICIP,* vol.1, pp. 65-68, 2005.

[9] Rolón, J.C., Salembier, P., Generalized Lifting for Sparse Image Representation and Coding, *Picture Coding Symposium, PCS 2007*, pp. 1-4, 7-9 Nov 2007, Lisbon, Portugal.

[10] Solé, J., Salembier, P., Adaptive discrete generalized lifting for lossless compression, in proc. *IEEE ICASSP 2004*, 17-21 May 2004 pp.iii - 57-60 vol.3, Montreal, Canada.

[11] Solé, J., Salembier, P., Generalized Lifting Prediction Optimization Applied to Lossless Image Compression, *IEEE Signal Processing Letters*, vol. 14, no. 10, pp. 695-698, Oct. 2007.

[12] Wang, D., *et al.*, Curved Wavelet Transform for Image Coding, *IEEE Transactions on Image Processing,* vol. 15, no. 8, pp. 2413-2421, 2006.

[13] Zhang, N., *et al.*, Directional Lifting-Based Wavelet Transform for Multiple Description Image Coding with Quincunx Segmentation, *Advances in Multimedia Information Processing,* LNCS Vol. 3768, pp. 629-640, Springer-Verlag, 2005.