# GENERALIZED LIFTING FOR SPARSE IMAGE REPRESENTATION AND CODING

*Julio C. Rolón*[†‡] *and Philippe Salembier*[†]

[†]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

## ABSTRACT

This paper investigates the use of generalized lifting to increase the sparseness of wavelet decompositions with application to image representation and coding. As in the bandelet approach, the strategy consists in applying first a separable wavelet decomposition and then in processing the details subbands to further decorrelate the signal representation. For this second step, we use a generalized lifting [13] which allows nonlinear processing of the details subbands. In this paper, the generalized lifting design is based on the pdf of the details coefficients after the separable wavelet decomposition and its goal is to minimize the coefficients energy. Both separable and non separable approaches are investigated. The generalized lifting is shown to reduce significantly the energy and the entropy of the representation. Furthermore, a simple quantification and entropy coding strategy is used to compare the rate-distortion characteristics of wavelet, bandelet and the proposed approach based on generalized lifting. Promising results are demonstrated.

**Index Terms—** Generalized lifting, wavelets, bandelets, image coding, sparse representation, nonlinear lifting

## 1. INTRODUCTION

Wavelet-based image decomposition has been extensively studied in recent years, giving rise among other applications to some of the most advanced image coding algorithms available today. However, one of the known limitations of the wavelet transform when applied in signals with dimensions higher than one, is their inability to deal with higher order singularities, as is the case of contours in images. Contours appear in the details subbands in wavelet domain as the largest magnitude coefficients. From a signal processing perspective, decorrelating the contours is equivalent to an improvement in the sparsity of the whole coefficients set.

Recent image representation methods like Curvelets [1,2] and Contourlets [6], that operate in the frequency domain, are among those methods whose goal is to achieve higher decorrelation than the one that is possible through the use of wavelets. However, the amount of redundancy introduced by them, along with their complex design, prevents their use in image coding applications at present.

Contrary to the previous methods, which belong to the frequency domain approach; Bandelets [8,10] is a locally adaptive decomposition that also decorrelates 1-dimensional singularities but with a line approximation rule, instead of the parabolic rule used by Curvelets and Contourlets. Since the Bandelet method is somehow an extension of wavelets, critical sampling is preserved, which is a major advantage of this method.

The bandelet algorithm performs a 2-D separable wavelet transform, decomposes the image subbands in blocks, and determines the geometric flow inside each block (the presence or absence of a contour) through the use of a 1-dimensional warped wavelet transform, thresholding, and a Lagrangian optimization criterion. The latter criterion is used at a second stage in a bottom-up approach to optimally select those blocks identified as having contours inside. In the end, the algorithm represents the image by a set of coefficients resulting from the second wavelet transform (1- dimensional), a quadtree structure, and a set of orientation values in which each element of the set corresponds to one leaf in the quadtree. Those blocks on which no contours are present, are represented by the coefficients resulting from the original 2-D separable wavelet transform. Bandelets have proved to improve the coding performances of separable wavelet decomposition.

Other possible solutions to the problem of localization and decorrelation of edges in images are those involving adaptive directional lifting schemes. Proposals similar to Bandelets are discussed in [4,5,9,15,16], a common characteristic is that they rely on predictions and interpolations over the pixel grid, be it Cartesian or quincunx. These predictions are made to find the best approximation to the actual contour to be decorrelated. In [3] a bandeletization process made on the resulting coefficients after directional lifting is included. Finally, adaptive update lifting steps are introduced in [7,11,12] for a lossless scenario.

A different approach to perform additional decorrelation on wavelet coefficients is presented here. It relies on the use of generalized lifting [13,14]. The generalized lifting framework allows the definition of nonlinear filter banks with perfect reconstruction. Its interest has been demonstrated for lossless compression of images in [14]. In this paper, we extend the work reported in [14] in several directions: First, our approach here deals with lossy compression of images. Then, we investigate the interest of 2-D non-separable generalized lifting. And finally, the generalized lifting is not applied directly on the original signal. We first conduct a 2-D separable wavelet transform as is done in bandelets [8,10]. Once in the wavelet domain, we apply a generalized predict step that minimizes the energy. Interestingly, we are able to increase the decorrelation of the wavelet coefficients as well as to preserve the critical sampling, a desirable condition for image coding. Moreover, initial experimental results show that the entropy and the energy of the subbands are significantly reduced, and that the R-D performance improves when compared to wavelets and bandelets.

In the following section, the generalized lifting is presented, while in section 3, we present the coding strategy, involving first a separable wavelet decomposition and then a generalized lifting. Experimental results are discussed in section 4 and, finally, section 5 reports the main conclusions of this work and ideas for future research.

## 2. GENERALIZED LIFTING

The classical lifting scheme, see fig. 1, is an efficient implementation of the Discrete Wavelet Transform (DWT). It involves a polyphase decomposition (also called Lazy Wavelet Transform, LWT) of the original signal $s[n]$ followed by Predict (P) and Update (U) steps. One of the interesting properties of lifting is that the structure itself guarantees perfect reconstruction.

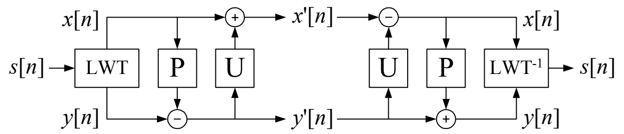


**Fig. 1.** Lifting Scheme

Consider the update step, U, in fig. 1. It takes as input $y'[n]$, computes a value that is added to $x[n]$ to produce $x'[n]$. On the synthesis side, the same operator U takes as input the same $y'[n]$ values and computes the same value as in the analysis side and subtracts, instead of adding it from $x'[n]$. As can be seen, the key of the lifting invertibility is the addition that can be inverted by a subtraction and no assumptions have to be made for the U (or P) operators.

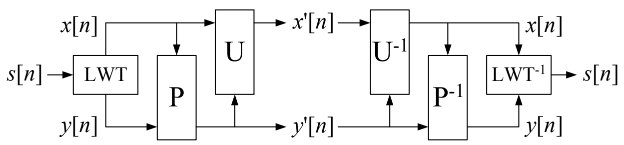


**Fig. 2.** Generalized Lifting Scheme

The generalized lifting (GL) scheme [13] is represented in fig. 2. Let us analyze the generalized prediction P. Instead of considering that the prediction P takes $x[n]$ as input in order to modify $y[n]$, by addition, 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 prediction 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. That is, it is an injective mapping

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

In order to design a 1-D generalized lifting, the context $\{x(n-i)\}_{i\in C}$ should involve pixels that are on the same line (or column) than the pixel being processed. For non separable generalized lifting, the context can be arbitrary and involve pixels surrounding the 2-D space of the pixel being processed.

Apart from the injectivity that is required to get perfect reconstruction, the generalized predict and update may be arbitrary. In [14], a design of the generalized predict minimizing the energy of $y'[n]$ 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.

Depending on the pdf value, this mapping can implement linear as well as highly nonlinear mapping. In this paper, our goal is to investigate the potential of generalized lifting to create sparser representations after an initial separable wavelet decomposition. We will assume that the pdf of the signal to process by the generalized lifting is known. In practice, of course, this pdf will not be known exactly. In [14], several strategies to solve this problem have been discussed: define the pdf of a class of images (here the details subbands of the initial wavelet decomposition) or use an online pdf estimation. In both cases, it has been shown that these strategies are only marginally worse than the optimal results obtained when the exact pdf is known. Anyway, in the sequel we assume that the pdf is known and the results reported here will be an approximation of what can be expected in practice. In the future, this pdf estimation issue will be investigated.

## 3. CODING SCHEME

The coding strategy is illustrated in fig. 3. It is composed of a separable 2-D Wavelet decomposition of 5 scales using a CDF 9/7 filter. All subbands except the low pass are quantized with a uniform scalar quantizer in order to obtain the R-D curve. Quantization is applied before the generalized lifting to prevent high distortion effects generated by the nonlinear mapping. Quantized subbands are processed independently with a generalized lifting scheme that contains only the prediction (P) operator. Note, that in this experiment, the generalized lifting scheme is applied just once and not iterated.

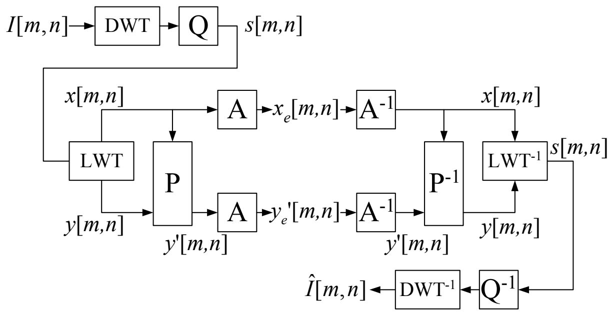


**Fig. 3.** Coding Scheme

A 1-dimension and a 2-dimension contexts are defined to investigate their performance. For the 1-dimensional case, the role of the LWT operator is to separate consecutive samples of $s[m,n]$ into even and odd sets of samples, generating $y[m,n]$ and $x[m,n]$ respectively. This operation is performed independently in two ways: row-wise and column-wise.

The energy is computed for the whole subband using the sets

$$\left\{x[m,n]\bigcup y'[m,n]\right\}\Big|_{row-wise},\ \left\{x[m,n]\bigcup y'[m,n]\right\}\Big|_{col-wise}.$$

The lowest valued is selected. One bit is used to indicate the selected orientation properly, this introduces a negligible overhead in this 1-dimensional case only. For the row-wise operation, the context of $y[m,n]$ is $\{x[m,n], x[m,n+1]\}$. For the column-wise operation, the context is $\{x[m,n], x[m+1,n]\}$.

The mapping is $y[m,n]\Big|_{with\ context\ \{x[m,n],x[\cdot]\}} \xrightarrow{P} y'[m,n]$,

where $x[\cdot]$ is used accordingly.

In the 2-dimensional case, the role of the LWT operator is similar to the 1-D case, but here a quincunx sampling pattern is created. So the separation in even and odd samples of $s[m,n]$ is done in an alternating row-wise; i.e., row $s[k,n]$ starts with an even sample, corresponding to $y[m,n]$ following in $s[k+1,n]$ with an odd sample first, corresponding to $x[m,n]$. The context of $y[m,n]$ is defined as $\{x[m,n-1], x[m-1,n], x[m,n], x[m+1,n]\}$ for the even rows, while for the odd rows is

$$\{x[m,n], x[m-1,n], x[m,n+1], x[m+1,n]\}.$$

The GL mapping is computed over

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

An arithmetic encoder is applied to both the approximation $x[m,n]$ and the details $y'[m,n]$ separately. The energy minimization and the higher decorrelation achieved through the generalized prediction allows us to obtain better encoding performance by using $y'[m,n]$ than what would be obtained coding $y[m,n]$ directly. Since the mapping (P) is invertible, distortion is introduced only by the quantizer. This single scale decomposition using GL scheme would be further improved by iterating the GL on the approximation to create a multiscale representation, which would moreover enable spatial scalability.

## 4. EXPERIMENTAL RESULTS

Experiments towards demonstrating the ability of GL to increase the sparsity of wavelet subband coefficients, as well as to evaluate the coding gain this additional sparsity may convey were conducted.

A set of five images with different characteristics were evaluated, namely *lena*, *barbara*, *mandrill*, *dintel*, and *bike*, being this latter of size 256 by 256 pixels, the former being all 512 by 512 pixels. All images have 8 bits per pixel (i.e. 256 gray levels).

### 4.1. Energy and entropy

The entropy and the energy of each detail subband of the separable decomposition were measured before and after the GL. The energy gain was computed as follows:

$$E_{\text{gain}} = 100\left(1 - \frac{Energy\{x, y'\}}{Energy\{s\}}\right)$$

For the entropy, a similar gain was computed:

$$H_{\text{gain}} = 100\left(1 - \frac{Entropy\{x, y'\}}{Entropy\{s\}}\right)$$

**Table I.** Entropy mean-gain values (%)

| scale | $H_{gain}$ (1-D context) | | | $H_{gain}$ (2-D context) | | |
|---|---|---|---|---|---|---|
| | HL | LH | HH | HL | LH | HH |
| 1 | -12.38 | -12.72 | -8.86 | -26.03 | -27.93 | -23.52 |
| 2 | -18.48 | -18.30 | -17.22 | -30.53 | -31.03 | -29.37 |
| 3 | -24.72 | -24.89 | -22.92 | -34.04 | -34.13 | -32.88 |
| 4 | -30.23 | -32.00 | -28.96 | -36.83 | -36.48 | -35.86 |
| 5 | -34.85 | -35.94 | -33.96 | -38.15 | -37.70 | -37.73 |

Results reported in Tables I and II, are the arithmetic mean gain values for each of the subbands obtained for the set of images. This gain was computed for each scale of the separable wavelet decomposition (from scale 1 to 5). The negative sign indicates that a decrease in entropy and energy is achieved.

**Table II.** Energy mean-gain values (%)

| scale | $E_{gain}$ (1-D context) | | | $E_{gain}$ (2-Dcontext) | | |
|---|---|---|---|---|---|---|
| | HL | LH | HH | HL | LH | HH |
| 1 | -47.69 | -44.89 | -35.89 | -49.03 | -49.35 | -47.12 |
| 2 | -50.69 | -48.33 | -47.92 | -49.60 | -50.02 | -49.80 |
| 3 | -49.87 | -50.31 | -50.06 | -50.60 | -50.82 | -52.00 |
| 4 | -48.87 | -47.71 | -50.02 | -48.46 | -50.13 | -49.39 |
| 5 | -48.51 | -46.81 | -47.21 | -48.60 | -55.54 | -46.98 |

The high variation in entropy values between 1-D and 2-D contexts that is reported in table I, where values of the 2-D context almost double those of its counterpart, confirms the fact that using 2-D information allows us to achieve higher decorrelation. It also allows us to infer that contour coefficients are implicitly having an impact in the construction of the mapping (through the context and the pdf), thereby establishing a link between the GL approach and other approaches like bandelets or adaptive directional lifting methods.

While the values of entropy gain vary significantly from the 1-D context to the 2-D context; table II shows that the energy variation is not so large when comparing both contexts. Energy gain is around 50% in both cases. In a single scale GL decomposition, this is the expected energy gain. From the original set of coefficients, only 50% become details ($y'[m,n]$), while the other 50% retain the energy ($x[m,n]$). In a multiscale decomposition, the size of the set of approximation coefficients becomes $100(2^{-J})\%$, where $J$ is the number of scales. The more the iterations, the lower the energy in the low-pass. Only under this latter condition it may be possible to evaluate the energy difference produced by the two contexts.

### 4.2. Rate-Distortion

To evaluate the coding gain, an R-D curve was obtained with the simple scheme of fig. 3. Shown in fig. 4 is the one corresponding to the 2-D context. The 2-D context was used as the previous section clearly showed that it is significantly more efficient than the 1-D context.

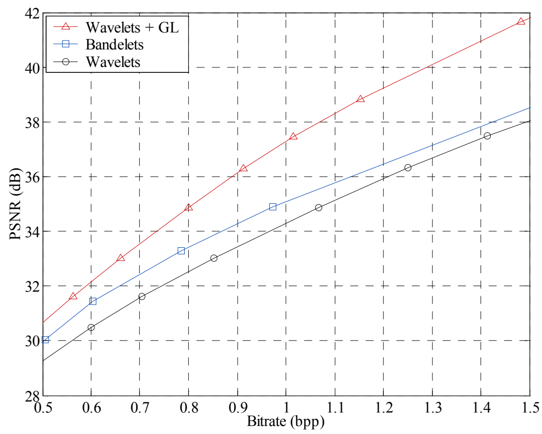


**Fig. 4.** Rate-distortion plot for *barbara*, 2-dimensional context

The encoding scheme has been kept simple to be able to compare the three approaches; wavelets only, i.e. $x[m,n]$ and $y[m,n]$ values; wavelets *and* generalized lifting, i.e. $x[m,n]$ and $y'[m,n]$; and bandelets. Our focus is on the comparison of the transforms themselves at this moment, rather than on other issues like specific entropy coding and scalability that would be analyzed in the future. The proposed setting, a uniform scalar quantizer (Q) followed by an arithmetic encoder (A) is also used in [8] to produce the R-D curve.

Interestingly, through the use of a 2-D context, it is possible to improve over bandelets from around 1dB at 0.7 bpp, to a little more than 2dB at 1bpp. As mentioned in section 2, these results can only be considered as preliminary since the generalized predict has been designed assuming that the pdf of the signal to decompose is known. In practice, this is not a reasonable assumption and strategies as those proposed in [14] will be investigated in the future.

## 5. CONCLUSIONS

In this paper, a generalized lifting (GL) design has been applied to wavelet coefficients after an initial separable wavelet decomposition in order to increase the sparsity of the representation. It has been demonstrated that GL gives very promising results at reducing the entropy and the energy of the coefficients, while improving the R-D performance, when compared against wavelets and bandelets in an image coding setting.

While the drop in the entropy values produced by GL is greatly influenced by the dimension of the context, the energy reduction may be directly related to the number of iterations of the GL scheme.

Some of the issues that will be further investigated in the near future are: *pdf estimation*, where the ideas in [13,14] will be investigated. *Multiscale approach,* the single GL decomposition reported in this paper will be extended to multiscale through iterations of the GL. This will enable spatial scalability, higher energy minimization, and hopefully increased sparsity. *Efficient entropy coding* mechanisms like SPIHT, EBCOT, or the like will be explored in the context of GL, in search for SNR scalability and greater compression.

## 6. ACKNOWLEDGMENT

This work has been partially supported by COFAA-IPN of Mexico, and by the TEC2004-01914 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] Candès, E. J., Donoho, D. L., Curvelets – A Surprisingly Effective Nonadaptive Representation for Objects with Edges, *in Curves and Surface Fitting,* Schumacher, L.L., Cohen, A., Rabut, C., eds., Vanderbilt Univ. Press, 2000.

[3] 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.

[4] 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.

[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] Heijmans, H.J.A.M., Pesquet-Popescu, B., Piella, G., Building Nonredundant Adaptive Wavelets by Update Lifting, *Applied Computational Harmonic Analysis*, no. 18, pp. 252-281, 2005.

[8] Le Pennec, E., Mallat, S., Sparse Geometric Image Representations with Bandelets. *IEEE Transactions on Image Processing,* vol. 14, no. 4, pp. 423-438, 2005.

[9] 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.

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

[11] Piella, G., *et al*, Combining Seminorms in Adaptive Lifting Schemes and Applications to Image Analysis and Compression, *Journal of Mathematical Imaging and Vision,* vol. 25, no. 2, pp. 203-226, 2006.

[12] Piella, G., Heijmans, H.J.A.M., Adaptive Lifting Schemes With Perfect Reconstruction, *IEEE Transactions on Signal Processing,* vol. 50, no. 7, pp. 1620-1630, 2002

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

[14] Solé, J., Salembier, P., Generalized Lifting Prediction Optimization Applied to Lossless Image Compression, *IEEE Signal Processing Letters*, Vol. 14, no. 11, Nov. 2007.

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

[16] 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.