PREDICTION DESIGN FOR DISCRETE GENERALIZED LIFTING

Joel Solé and Philippe Salembier

Technical University of Catalonia (UPC)
Dept. of Signal Theory and Communications, Campus Nord, Barcelona, SPAIN
{joel, philippe}@gps.tsc.upc.es

ABSTRACT

Lifting scheme is a useful tool to create different types of wavelets, including adaptive and nonlinear decompositions. Generalized lifting scheme is more flexible and can improve lifting results, but the design of generalized prediction and update is a difficult task. This paper proposes a design strategy to optimize the prediction step, in the sense of minimizing detail signal energy, and shows some possible applications. Promising results are reported for certain classes of images.

1. INTRODUCTION

Wavelet multi-resolution decomposition of images have shown their efficiency in many image processing areas, and specifically in compression. However, due to the non-stationarity of the image signal, efforts were naturally focused on the creation of adaptive schemes that modify the transform according to local signal characteristics [1]. The main goal is to adaptively choose a good wavelet basis for each sample of the image so that local statistics are taken into account. As a result, a compact representation is obtained for all kind of areas because wavelet transform is adapted to every signal point.

Lifting scheme is a suitable framework for developing time-varying/nonlinear wavelet filter banks, mainly because the lifting structure itself assures reversibility. In [2], a generalization of the lifting scheme was proposed in order to include information given by the same branch to be filtered and to allow nonlinear combinations of the samples between the two branches. Moreover, new conditions to assure reversibility were required and established. The generalized lifting scheme may potentially improve the transform properties, but the framework is essentially devoted to nonlinear processing and it implies two fundamental drawbacks. First, the wavelet multi-resolution framework of nested subspaces for signal description is lost. Second, in the context of compression, filter design for an embedded lossy-to-lossless code becomes very difficult, since frequency-band notion disappears and inter-band relations are less obvious. In addition, the output may not be a continuous function of the input samples and so, quantification errors may be magnified.

In [2], we also proposed a generalized discrete prediction step and discussed its interest for lossless compression. In this paper we formulate the design of the prediction step as an optimization problem that depends on the probability density function of the signal. Moreover, we apply the resulting lifting for biomedical images (mammography) and remote sensing images (sea surface temperature) and evaluate the gains provided by the proposed prediction.

The organization of this paper is as follows: In Section 2, the classical, the adaptive and the generalized lifting are briefly reviewed. Section 3 focuses on the optimization of the prediction step. Experiments are reported in section 4 and, finally, conclusions are established in section 5.

2. LIFTING SCHEME, ADAPTIVE LIFTING AND GENERALIZED LIFTING

The lifting scheme (Figure 1) introduced by Sweldens [3] is a well-known method to create bi-orthogonal wavelet filters from other ones. Usually, a polyphase decomposition (LWT) of the input signal $\underline{x}_o$ is initially done, obtaining an approximation signal $\underline{x}$ and a detail signal $\underline{y}$. Then, lifting steps are performed by predicting the detail signal from the $\underline{x}$ samples (1) and updating the approximation signal with the $\underline{y}$ samples (2). The so-called prediction and update lifting steps improve the initial wavelet properties

$$y'[n] = y[n] - P[\underline{x}] \qquad (1)$$

$$x'[n] = x[n] + U[\underline{y}'] \qquad (2)$$

Several lifting steps may be concatenated in order to reach the desired properties for the wavelet basis. These prediction and update operators may be a linear combination of $\underline{x}$ and $\underline{y}$, respectively, or any nonlinear operation, since by construction the lifting scheme is always reversible.

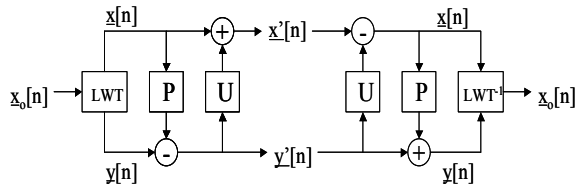

Figure 1. *Lifting Scheme*

The adaptive lifting scheme [4] is a modification of the classical lifting. Figure 2 illustrates an example of adaptive update step followed by a fixed prediction step. At each sample n, an update operator is chosen according to a decision function that depends on $\underline{y}$, but that might also depend on the sample $x[n] \in \underline{x}$ being updated. In this case, a problem arises because the decoder does not know the sample $x[n]$ used at the coder for the decision. Instead, the decoder has access to $x'[n]$, the updated version of $x[n]$ through an unknown update filter. Therefore, the challenge is to find a decision function and a set of filters which allow reproducing the decision at the decoder, thus obtaining a reversible decomposition scheme.

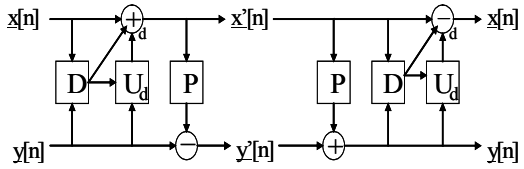

Figure 2. *Adaptive Update Lifting Scheme*

A generalization of the lifting scheme was proposed in [2]. As can be seen in Figure 3, the generalized prediction and update steps combine the filtering stage as well as the addition of classical lifting. This leads to a more general framework, allowing more complex, possibly adaptive or nonlinear operations.

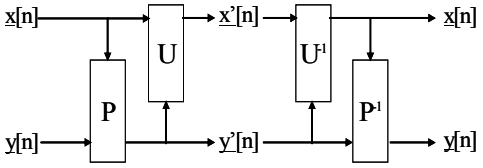

Figure 3. *Generalized Lifting Scheme*

To establish a formal definition of the generalized prediction and update, let A be the set of functions a from $\Re \times \Re^k$ to itself:

$a \in A \Leftrightarrow a : \Re \times \Re^k \rightarrow \Re \times \Re^k$, such that:

$\{z_1'[n], z_2'[n-n_1], \ldots, z_2'[n-n_k]\} =$

$a\ \{z_1[n], z_2[n-n_1], \ldots, z_2[n-n_k]\}$

We denote the samples by z in order to maintain the same definition for both the prediction and the update steps. For the prediction (update) step, it is assumed that $z_1[n]=y[n]$ and $z_2[n]=x[n]$ ($z_1[n]=x[n]$ and $z_2[n]=y[n]$). Let A_0 be the subset of A containing all functions that do not modify $z_2[n]$, that is, for which the restriction to $\Re^k$ is the identity: $A_o = \left\{a \subseteq A \mid a_{|\Re^k \rightarrow \Re^k} = Id\right\}$. Then, a generalized lifting step is any function belonging to A_0.

In order to have a reversible scheme, the generalized prediction and update cannot be chosen arbitrarily. To attain reversibility the generalized steps must be bijective functions of A_0.

As presented, the scheme assumes that the values taken by $\underline{x}$ and $\underline{y}$ are real numbers. However, the proper design of the resulting generalized prediction and update steps is very difficult to handle. A possible solution is to consider the discrete version of the generalized lifting scheme. To this goal, we assume that the values taken by $\underline{x}$ and $\underline{y}$ are integers, and that the generalized steps outputs are also integers. We consider the framework for discrete gray-scale images where each pixel is represented by 8 bits. Without loss of generality, we assume that sample values range from -128 to 127. Let us call Z_{255} the set of integers that belong to the interval $[-128,127]$. The discrete generalized update and prediction are now functions from the $Z_{255} \times Z_{255}{}^k$ space to itself that can only modify the first component. Note that the output values are also restricted to the interval [-128,127]. The statements made for the real case are still valid. In particular, reversibility is obtained if the following a mappings are bijective:

$\{z_1'[n], z_2[n-n_1],\ldots, z_2[n-n_k]\} = a\{z_1[n], z_2[n-n_1],\ldots, z_2[n-n_k]\}$

For $z_2[n-n_1], \ldots, z_2[n-n_k]$ fixed, the set of all possible values of $z_1[n]$ describes a column in the $Z_{255} \times Z_{255}{}^2$ space. Let $C_{i \in Z_{255}{}^k}$ denote such a column:

$$C_{i \in Z_{255}{}^k} = \left\{z_1[n], z_2[n-n_1] = i_1, \ldots, z_2[n-n_k] = i_k\right\}$$

As the generalized update and prediction can only modify the first component $z_1[n]$, they map the column $C_{i \in Z_{255}{}^k}$ to itself. In order to have a reversible scheme, the mapping of $C_{i \in Z_{255}{}^k}$ to itself should be bijective for all columns. Figure 4 illustrates the case where $k=2$.

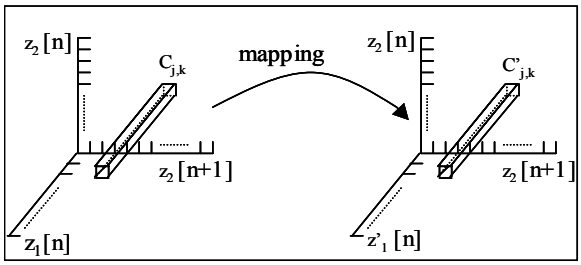

Figure 4. *Discrete mapping from* $Z_{255} \times Z_{255}{}^2$ *to itself. The lifting step is reversible if all mappings from every column* $C_{j,k}$ ($C_{i \in Z_{255}{}^k}$) *to itself are bijective.*

3. PREDICTION DESIGN

Once the previous definitions have been stated, the critical problem is to design useful generalized lifting steps for specific applications. In this paper, we consider lossless compression. This section focuses on the design of a prediction step, in order to apply an appropriate transform to sample $y[n]$ knowing its neighbors, $\underline{x}[n]$. In this case, we define a column as:

$$C_{i\in Z_{255}{}^k} = \{y[n], x[n-n_1]=i_1,\ldots,x[n-n_k]=i_k\}$$

The filter design problem amounts to find a mapping for every column $C_{i\in Z_{255}{}^k}$ of the $Z_{255}\times Z_{255}{}^k$ space to the transformed column, noted $C'_{i\in Z_{255}{}^k}$.

$$C'_{i\in Z_{255}{}^k} = \{y'[n], x[n-n_1]=i_1,\ldots,x[n-n_k]=i_k\}$$

For the transform to be reversible, every column mapping should be bijective. In addition, columns form a partition of the $Z_{255}\times Z_{255}{}^k$ space, so the prediction (P) mappings are independent Accordingly, every mapping (P_i) can be designed independently from each other.

$$P(y[n], \underline{x}[n]) = \bigcup_{\forall\, i\in Z_{255}^k} P(y[n], \underline{x}[n])|_{\underline{x}[n]=i} = \ldots$$

$$\ldots = \bigcup_{\forall\, i\in Z_{255}^k} P_i(y[n])|_{\underline{x}[n]=i}$$

Given $i\in Z_{255}{}^k$, the transform relates every input value $y[n]\in Z_{255}$ one-to-one to every output value $y'[n]\in Z_{255}$. So the output values for each i are related to the input values simply through a permutation matrix.

$$C'_i = \underline{P_i} C_i$$

The prediction step can be seen as the union of $card(Z_{255}{}^k)$ permutation matrix, $\underline{P_i}$. Consequently, the complexity associated to this formulation grows rapidly with k. In practice, one has to use a low value of k (i.e., a reduced number of context values $x[n]$) or to take advantage of the similarities between permutation matrices that may arise.

State-of-the-art entropy coders profit from several characteristics of wavelet coefficients. Specifically, they tend to increase their performance when the coefficients energy is minimized. As a result, we aim at designing a mapping that minimizes the expected energy of the detail signal, $y'[n]$.

$$P_{opt} = \arg\min_p E\{y'^2\} = \bigcup_{i\in Z_{255}{}^k} \arg\min_{p_i} E\{y'^2 \mid \underline{x}=i\}$$

Second equality is due to the independency of columns.

$$E\{y'^2\} = \sum_{i\in Z_{255}^k} E\{y'^2 \mid \underline{x}[n]=i\} P(\underline{x}[n]=i)$$

So, the design of the prediction function reduces to the definition of the optimal column mapping P_i (or permutation matrix $\underline{P_i}$) for all columns:

$$E\{y'^2 \mid \underline{x}=i\} = \sum_{n=-128}^{127} n^2 P(y'=n \mid \underline{x}=i)$$

$$= \sum_{n=-128}^{127} n^2 P(P_i(y)=n) = \sum_{n=-128}^{127} n^2 P\left(y=P_i^{-1}(n)\right)$$

This equation, because of P_i is an isomorphism, can be expressed as

$$E\{y'^2 \mid \underline{x}=i\} = \sum_{n=-128}^{127} P(y=n \mid \underline{x}=i)\left(P_i^{-1}(n)\right)^2$$

$$= \underline{v}_i \left(P_i^{-1}(-128)^2 \ldots P_i^{-1}(127)^2\right)^T$$

Where $\underline{v}_i = (P(y=-128 \mid \underline{x}=i)\ldots P(y=127 \mid \underline{x}=i))$.

Introducing the permutation matrix, we obtain

$$E\{y'^2 \mid \underline{x}=i\} = \left((-128)^2 \ldots (127)^2\right)\underline{P}_i \underline{v}_i{}^T$$

which is minimized when permutation relates input values of high probabilities with small output values.

Then, assuming that the pdf is known, a column map is created by constructing a vector with input values sorted by their probability in descending order. The first element of this vector, which is the more probable input sample for the given context, is assigned (mapped) to a *0* output value (the minimum energy output). Following, *1* is assigned to the vector second element, *–1* to the third, *2* to the fourth and so on. In conclusion, a prediction step is performed by column mappings which are look-up-tables that re-order input values according to their probabilities. These look-up-tables are different and more practical representations of permutation matrices.

4. EXPERIMENTS AND RESULTS

In this section, we apply the previous design strategy to three kind of images, natural, biomedical images (mammography) and remote sensing images (sea surface temperature, SST). The last two image classes have been chosen because of their pdf, which considerably differs from that of natural images. We restrict ourselves to *k=2*, that is, the probability distribution function of a sample used for the prediction design is estimated conditioned to its two neighbors.

The decomposition is followed by an entropy coder. Two are used for the experiments: SPIHT [5] and the EBCOT from the JPEG2000 standard [6]. Although frequency-band interpretation is not valid for this nonlinear scheme, we consider here a transform band as a subset of samples from which we expect to share the same statistics, and these bands are coded as if they were obtained from the usual time-frequency wavelet decomposition. This assumption is surely not optimal, but it does not seem to worsen performance significantly.

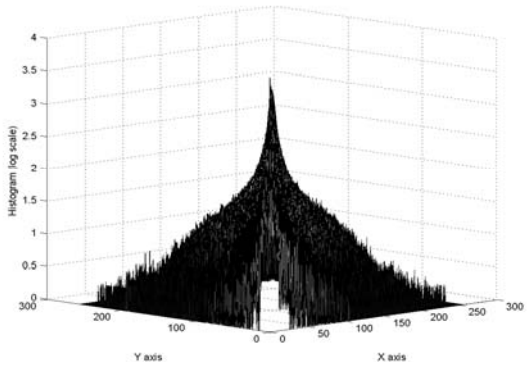

Figure 5. *Pdf (histogram) of y[n] (x axis) conditioned to the mean value of its two vertical neighbors (y axis) for the set of natural images.*

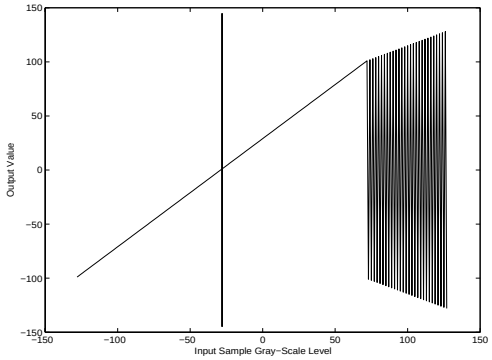

Figure 6. *Mapping for natural images. Vertical line at -28 marks the value of neighbor samples. A linear part on the left hand side (input values between -128 and 72) and a nonlinear part on the right hand side (input values between 73 and 127) can be observed.*

4.1. Natural images

The probability distribution function of a sample $y[n]$ conditioned to the value of its two vertical neighbors, $x[n]$ and $x[n+1]$, have been extracted from a set of seven natural images (those of table 1). Figure 5 partly represents this pdf: it is a histogram depicting the frequency of apparition of a sample value in function of the mean of its two neighbors values. A complete representation would have 4 dimensions, since the histogram depends of both neighbors values, not only its mean. A common pattern has been observed for all contexts in this pdf. Concretely, it has a maximum at the mean of the two neighbors, $(x[n]+x[n+1])/2$, an decreases monotonically and symmetrically on both sides. This observation allows us to avoid the implementation of the prediction step by means of the look-up table design described in the previous section.

Rate (bytes)	2 resolution levels		3 resolution levels	
Image / Filter	LeGall	Opt. Pred.	LeGall	Opt. Pred.
Lenna	203402	190142	159958	157030
Peppers	212171	215093	172529	172552
Girl	195707	197849	149062	150920
Baboon	235983	224736	210328	211192
Barbara	212913	199905	178565	172672
Cameraman	55100	48273	42524	40892
Goldhill	52731	53262	48545	48390
Mean	**166858**	**161323**	**137359**	**136235**

Table 1. *Natural Image Lossless compression with SPIHT coder.*

For testing purposes, the seven images have been compressed with this one dimension 2-taps prediction and a SPIHT coder. Note that no update stage is used. The image is first filtered vertically and the approximation signal is filtered horizontally, resulting in a three-band decomposition. The same prediction is used vertically and horizontally for all resolution levels.

For comparison, results with the LeGall wavelet used in JPEG2000 standard for lossy-to-lossless compression are also shown. It is worth to note the support of these decompositions. While LeGall low-low band output samples have a support of 5x5 input samples, and 5x3, 3x5 and 3x3 samples the three other bands, the proposed scheme is a decimation for the approximation band and the other two bands have a support of only 3 samples. Despite of this, our prediction (table 1) performs better for 2 resolution levels and marginally better for 3 resolution levels.

However, with EBCOT coder very similar results were obtained for both decompositions. This suggests that the design strategy in the case of natural images does not provide a prediction scheme that is significantly different from the LeGall linear case. In order to clarify this point, let us analyze the prediction resulting from the optimization strategy.

The kind of prediction mapping that arises from the pdf of natural images has two differentiated parts, a linear and a nonlinear part. Figure 6 shows the prediction mapping when the context $x[n]=x[n+1]=-28$. The context value is indicated by a vertical line at -28. Input values between -128 and 72 are linearly mapped to output values between -100 and 100. This mapping is equivalent to the linear combination:

$$y'[n] = y[n] - (x[n]+x[n+1])/2$$

This linear part of the mapping is due to the pdf that has a maximum for $(x[n]+x[n+1])/2$ and that decreases monotonically and symmetrically on both sides. For input values above 72, the mapping is highly nonlinear but it arises from our choice to work with output values between -128 and 127. As a result of this analysis, we can see that for most probable input values, the mapping is the same as the LeGall filter. Therefore powerful coder, as EBCOT, returns practically the same results for both decompositions.

Rate (bytes)	2 resolution levels		3 resolution levels		4 resolution levels	
Image / Filter	LeGall (jpg2k)	Opt. Pred.	LeGall (jpg2k)	Opt. Pred.	LeGall (jpg2k)	Opt. Pred.
Mam-7	400109	388646	376297	364153	371006	358588
Mam-8	446962	428159	418873	399129	412455	392471
Mam-9	409648	401539	382768	374069	376228	367613
Mam-10	240535	217647	215511	192471	209235	186559
Mam-11	255827	230639	228213	202218	220197	194732
Mean	**350616**	**333326**	**324332**	**306408**	**317824**	**299993**

Table 2. *Mammography. Lossless compression with EBCOT.*

Rate (bytes)	2 resolution levels		3 resolution levels		4 resolution levels	
Image / Filter	LeGall (jpg2k)	Opt. Pred.	LeGall (jpg2k)	Opt. Pred.	LeGall (jpg2k)	Opt. Pred.
SST AfricaNW 4	2033824	1682236	2036641	1611577	2036443	1592365
SST AfricaNW 5	1590859	1310141	1590796	1256018	1590782	1242411

Table 3. *SST images lossless compression with EBCOT.*

Therefore, there is potentially more compression to be gained for those images belonging to a class whose pdf differs significantly from that of natural images. Next subsections illustrate some results for two classes of medical and satellite images.

4.2. Medical images

To realize the experiment for biomedical images we have selected a set of 11 mammography images from our database. Figure 7 shows an example. Their size is about 1 Mbyte without compression. Six images are used to estimate the pdf for this type of images. The resulting pdf does not exhibit a regular pattern as in the case of natural images.

As figure 8 shows, mammography images pdf is not as structured as natural images pdf. It is not symmetrical neither decreases monotonically. Usually, several maxima appear and also, the darker values are rather probable in most of the contexts. Figure 9 depicts the mapping when $x[n]=x[n+1]=-88$. As can be seen, in this case the mapping of the most probable input values (values around -88) is already nonlinear.

The decomposition is performed for the other 5 images and compressed with the EBCOT. For comparison, the same process is realized using the LeGall transform. Results (Table 2) are 5 to 6% better for the generalized prediction for all resolution levels.

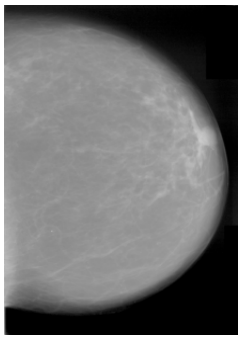

Figure 7. *Example of a mammography image.*

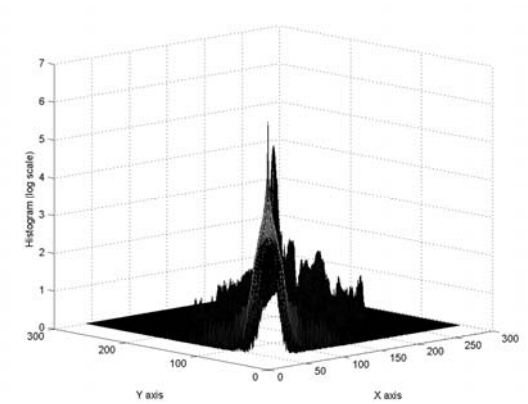

Figure 8. *Pdf (histogram) of y[n] (x axis) conditioned to the mean value of its two vertical neighbors (y axis) for the set of mammography.*

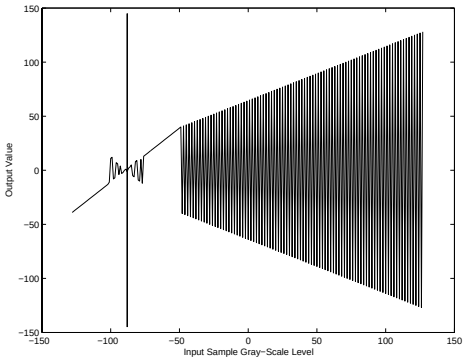

Figure 9. *Example of mapping for mammography. Vertical line marks the value of the two neighbor samples (which is the same).*

4.3. SST images

The last experiment applies the optimized prediction to a set of images of the sea surface temperature (SST) obtained with the AVHRR/2&3 sensors from the NOAA satellite series. An example is shown in figure 10. Images size range from 5 to 7 Mbytes. This specific set

is devoted to the African northwest coast and forms a huge image corpus, so modeling the pdf is worth compared to the gains in compression.

Three SST images are used to estimate the pdf. The resulting mapping is stored in memory and so, prediction is performed using look-up tables. The conditional pdf (Figure 11) of this kind of images significantly differs from that of natural images. The most light and dark values are highly probable for all context. Earth regions are almost binary. In these circumstances LeGall prediction performs poorly, so there is much to be gained escaping from the linear processing. Figure 12 shows and example of mapping when $x[n]$=-1 and $x[n+1]$=12. As can be seen, the optimized prediction mappings for these images are quite nonlinear.

Two other SST images where compressed by these means followed by EBCOT. A gain of 20% is obtained compared to lossless JPEG2000 (Table 3). For the SPIHT coder gains are even larger in terms of bit saving. For instance, the image SST North Africa 5 is compressed to 3034846 bytes with optimal prediction, and only to 3451307 bytes with the LeGall decomposition.

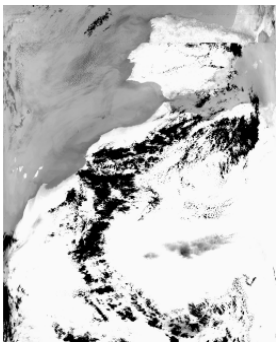

Figure 10. *Example of a Sea Surface Temperature (SST) image.*

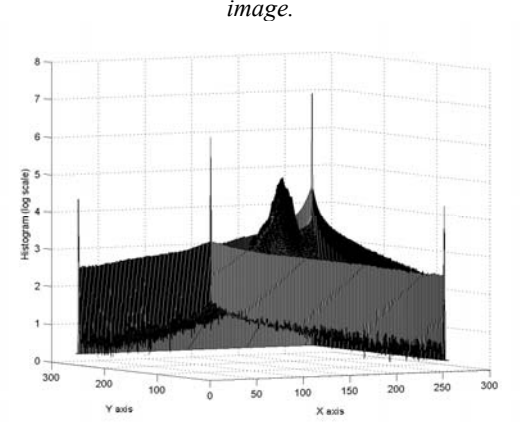

Figure 11. *Probability distribution function for the set of SST images.*

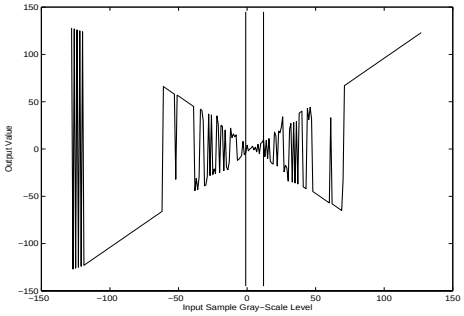

Figure 12. *Example of mapping for SST images. The two vertical lines mark the value of both neighbors*

5. CONCLUSIONS

The generalized lifting scheme framework is used in this paper to derive an optimized prediction step that minimizes the detail signal energy. It is used for the decomposition of different types of images. Results are promising, and even very good for the SST images. This is because they have a pdf that results in a prediction step differing considerably from LeGall filter. Even if the filter support is smaller, compression gains are up to 20% better. Anyway, a major difficulty appears at this point, since larger supports seem to be difficult to handle in practice if the pdf does not show any structure. An other solution to improve the results would be to design an optimized update step following a strategy similar to the one described in this paper for the prediction. We will study these points in the future.

6. REFERENCES

[1] O.N. Gerek and A. E. Cetin, "Adaptive polyphase subband decomposition structures for image compression", *Trans. on Image Processing*, vol. 9, num. 10, pp. 1649 –1660, Oct. 2000.

[2] J. Solé and P. Salembier, "Adaptive Discrete Generalized Lifting for Lossless Compression", ", *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing 2004*, Montreal, May 2004.

[3] W. Sweldens, "The lifting scheme: A custom-design construction of biorthogonal wavelets", *Appl. Comput. Harmon. Anal.*, vol. 3, num. 2, pp. 186-200, 1996.

[4] B. Pesquet-Popescu, G. Piella and H. Heijmans, "Adaptive update lifting with gradient criteria modeling high-order differences", *IEEE Int. Conf. on A.S. and S. Processing 2002*, vol. 2, pp. 1417 –1420, 2002.

[5] A. Said and W.A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees", *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 6, num. 3, 243-250, June 1996.

[6] ISO/IEC 15444-1: JPEG 2000 image coding system, 2000.