OPTIMAL WEIGHTED MODEL-BASED BIT ALLOCATION FOR QUINCUNX SAMPLED IMAGES

Annabelle Gouze, Christophe Parisot, Marc Antonini and Michel Barlaud

Laboratoire I3S, CNRS / UNSA,
2000 route des Lucioles BP121
06903 Sophia Antipolis, France
{gouze,parisot,am,barlaud}@i3s.unice.fr

ABSTRACT

In this paper we address the problem of quincunx sampled images compression. Our objective is to define an efficient bit allocation method adapted to quincunx sampled images. We first estimate the subband optimal weightings for a global distortion estimation. These weightings are derived from the filters used to perform the quincunx wavelet transform. Then, we use our weightings in a model-based bit allocation procedure. Our method uses generalized Gaussians to approximate the probability density function of the wavelet coefficients in each subband. The bit allocation method which is proposed provides both low complexity and high performance.

1. INTRODUCTION

This paper deals with the compression of images directly acquired on a quincunx grid. The compression scheme includes in a first stage, a quincunx wavelet transform [1, 2, 3, 4]. The second stage is the bit allocation procedure. Its purpose is to determine the quantizers in each subband which minimize the total MSE (Mean Squared Error) for a given output bit-rate. The global distortion (MSE) is estimated by a weighted sum of the distortions in each subband of the quincunx wavelet transform. We have:

$$\sigma_x^2 = \sum_{i=1}^{n_{SB}} \pi_i \sigma_{Q_i}^2$$ (1)

Usually, the weightings $\pi_i$ are set to 1 when the synthesis wavelet filters are normalized to 2. But these weightings are correct only for orthogonal filter banks [5]. In this paper, we show how to compute the optimal weightings for the biorthogonal quincunx lifting scheme [6] but the method also applies more generally for quincunx wavelet transforms. The weightings we propose depend on the lifting operators. For our experimental results, we assume that the transmission of compressed data is performed through a rate constrained channel. Thus, rate allocation and control procedures have to be used in order to fit the channel characteristics. In this paper, we adapt the bit allocation method proposed in [7] for quincunx images, using the optimized distortion weights. In section 2, we recall the principle of the quincunx lifting scheme. In section 3, we introduce the computation of the optimal weightings. Section 4 is devoted to the model-based bit allocation procedure and section 5 presents some experimental results.

2. BACKGROUNDS ON THE LIFTING SCHEME

The lifting scheme was introduced by Sweldens [8]. His work is both inspired by the work of Donoho [9] and by the Lounsbery [10] one concerning the transform on meshes. Lifting is related on the filter bank construction of Vetterli and Herley [11]. The lifting scheme is composed of three main steps [8]. First, the original signal is split into two disjoint subsets : $s^{(0)}$ and $d^{(0)}$. Afterwards, a predicting operator is applied in order to obtain the set of wavelet coefficients $d$. Third, an update operator is applied on wavelet coefficients to get the low frequency signal $s$ at a lower resolution. The general scheme can include several lifting steps. The lifting is built by alternating different prediction and update steps. Some methods exist to design lifting operators. This yields to the following equations:

For $l = 1$ to $N_b$ steps do :

$$d^{(l)} = d^{(l-1)} - p^{(l)} * s^{(l-1)}$$

$$s^{(l)} = s^{(l-1)} + u^{(l)} * d^{(l)}$$

where $d = d^{(N_b \text{ steps})}$ and $s = s^{(N_b \text{ steps})}$. The lifting scheme implementation is simple, fast, efficient and easily invertible. In his introduction on lifting [6], Sweldens proposed a method to design 1D lifting schemes. Daubechies showed in [12] how to factorize 1D wavelet filters into lifting scheme. In [13], we showed how to obtain quincunx lifting filters from 1D lifting filters. Kovačević [14] defines a multidimensional two-step lifting scheme based on
the Neville filter construction. In [15], we proposed an optimization method to build quincunx lifting operators. Filters designed by all these methods are not orthogonal but biorthogonal. Thus, the property of energy conservation is not verified. Therefore, we have to compute the coding distortion by taking into account this remark.

3. RECONSTRUCTION DISTORTION

Most of bit allocation methods minimize the mean squared error. They determine optimal quantization steps to obtain the best bit-rate-distortion ratio. The bit allocation follows a wavelet transform, and the global distortion is estimated by the sum of subband distortions. This method is correct only for orthogonal filter banks, but not for biorthogonal filter banks. The distortion depends on the reconstruction operation and thus of \( h \) and \( g \), the inverse wavelet transform filters. We compute the distortion for a non-separable 2D wavelet transform. Wavelet filters are supposed to be biorthogonal. In the quincunx lifting case, \( h \) et \( g \) are found with the formulas:

\[
\begin{align*}
\tilde{h}^{(0)}[n] &= \tilde{g}^{(0)}[n] = \delta[n_2] \delta[n_3] \\
\tilde{h}^{(l)}[n] &= \tilde{g}^{(l-1)}[n] + \sum_{k \in \mathbb{Z}^2} \tilde{h}^{(l-1)}[n+(1,0)-L(k)] U[l][k] \\
\tilde{h}^{(l)}[n] &= \tilde{h}^{(l-1)}[n] + \sum_{k \in \mathbb{Z}^2} \tilde{g}^{(l)}[n-L(k)-(1,0)] P[l][k]
\end{align*}
\]

with \( 0 < l < L_{1x} \), where \( L_{1x} \) is the number of lifting operators, \( n = (n_2, n_3) \in \mathbb{Z}^2 \) and \( L(x,y) = (x+y,x-y) \). The last lifting step gives filters \( \tilde{h} \) et \( \tilde{g} \). Usevitch [5] gave distortion estimations in the 1D and the separable 2D cases. A wavelet transform, applied on a signal \( s_{1d} \), generates one high frequency \( d_2 \) subband and a low frequency one \( s_2 \). In each subband, the error estimation depends on the quantization step defined by the bit allocation. Let us define \( \varepsilon_s \) et \( \varepsilon_d \), the quantization error images. The global error \( \varepsilon \) is obtained by the application of the inverse wavelet transform to the error images \( \varepsilon_s \) et \( \varepsilon_d \). Thus, \( \forall n = (n_2, n_3) \in I_{SD} \) \( I_{SD} = [1, \ldots, M] \times [1, \ldots, N] \) corresponds to the original image support, and \( I_{S_2} \) to the low frequency image support, \( l \in \mathbb{Z}_l \), with \( M \) and \( N \) being the number of lines and columns of the original image:

\[
\begin{align*}
\varepsilon[n_2,n_3] &= \sum_{k \in \mathbb{Z}^2} \tilde{h}[(n_2,n_3)-L(k)] \varepsilon_s[k] \\
&+ \sum_{k \in \mathbb{Z}^2} \tilde{g}[(n_2,n_3)-L(k)-(1,0)] \varepsilon_d[k]
\end{align*}
\]

The MSE is defined by the formula:

\[
\sigma_e^2 = \frac{1}{MN} \sum_{n \in \mathbb{Z}^2} E[\varepsilon[n]^2]
\]

The MSE is developed as follows:

\[
\sigma_s^2 = \frac{1}{MN} \sum_{n \in \mathbb{Z}^2} E \left[ \left( \sum_{k \in \mathbb{Z}^2} \tilde{h}[n-L(k)] \varepsilon_s[k] \right)^2 \right] + \frac{2}{MN} \sum_{n \in \mathbb{Z}^2} \sum_{l \in \mathbb{Z}_l} \sum_{k \in \mathbb{Z}^2} \tilde{g}[n-L(k)-(1,0)] \varepsilon_d[k] \varepsilon_s[k]
\]

\[
\sigma_d^2 = \frac{1}{MN} \sum_{n \in \mathbb{Z}^2} E \left[ \left( \sum_{k \in \mathbb{Z}^2} \tilde{g}[n-L(k)-(1,0)] \varepsilon_d[k] \right)^2 \right]
\]

\[
\sigma_s^2 = \frac{1}{MN} \sum_{n \in \mathbb{Z}^2} E \left[ \left( \sum_{k \in \mathbb{Z}^2} \tilde{h}[n-L(k)] \varepsilon_s[k] \right)^2 \right] + \frac{2}{MN} \sum_{n \in \mathbb{Z}^2} \sum_{l \in \mathbb{Z}_l} \sum_{k \in \mathbb{Z}^2} \tilde{h}[n-L(k)] \varepsilon_s[k] \varepsilon_s[k]
\]

Finally, we have:

\[
\sigma_s^2 = \frac{1}{MN} \left( \sum_{k \in \mathbb{Z}^2} \sum_{l \in \mathbb{Z}_l} \sum_{n \in \mathbb{Z}^2} \tilde{h}[n-L(k)] \tilde{h}[n-L(l)] \right) + \frac{2}{MN} \left( \sum_{k \in \mathbb{Z}^2} \sum_{l \in \mathbb{Z}_l} \sum_{n \in \mathbb{Z}^2} \tilde{g}[n-L(k)] \varepsilon_s[k] \varepsilon_s[l] \right)
\]

To simplify computations, we make two assumptions on error images \( \varepsilon_s \) et \( \varepsilon_d \).

1. \( \varepsilon_s \) and \( \varepsilon_d \) are mutually decorrelated:

\[
E[\varepsilon_s \varepsilon_d] = 0
\]

2. \( \varepsilon_s \) and \( \varepsilon_d \) correspond to a white noise and verify

\[
E[\varepsilon_s[k] \varepsilon_s[l]] = \sigma_s^2 \quad \forall k \neq l, k \in I_{S_2}
\]

\[
E[\varepsilon_s[k] \varepsilon_d[l]] = 0 \quad \forall k \neq l, k \in I_{S_2}
\]

\[
E[\varepsilon_d[k] \varepsilon_d[l]] = \sigma_d^2 \quad \forall k \neq l, k \in I_{D_2}
\]

\[
E[\varepsilon_d[k] \varepsilon_d[l]] = 0 \quad \forall k \neq l, k \in I_{D_2}
\]

Gersho et Gray [16] define limitations lied on these assumptions (3) et (4). By considering the two assumptions, the simplification of the MSE definition (2) gives:

\[
\sigma_e^2 = \frac{1}{2} \sum_{n \in \mathbb{Z}^2} \tilde{h}^2[n] \sigma_s^2 + \frac{1}{2} \sum_{n \in \mathbb{Z}^2} \tilde{g}^2[n] \sigma_d^2
\]
The supports $I_{S_{2}}$ and $I_{S_{4}}$ of $\varepsilon_{a}$ and $\varepsilon_{d}$ are twice more little than $I_{S_{0}}$ of size $MN$. The size ratio, between the considered subband and the original image, weights the subband variance. The distortion for one level of decomposition is:

$$\sigma_{\varepsilon}^{2} = \frac{\|\hat{y}\|^{2}}{2} \sigma_{\varepsilon}^{2} + \frac{\|\tilde{y}\|^{2}}{2} \sigma_{\varepsilon}^{2}$$

The distortion for $L$ decompositions is defined by weighted sum of subband distortions:

$$\sigma_{\varepsilon}^{2}(x, x) = \frac{\|\hat{y}\|^{2L}}{2L} \sigma_{\varepsilon}^{2} (s_{K + \frac{1}{4}, s_{K + \frac{1}{4}}}) + \frac{\|\tilde{y}\|^{2(l-1)}}{2^{l}} \sigma_{\varepsilon}^{2} (d_{K + \frac{1}{4}, d_{K + \frac{1}{4}}})$$

The factors $2^{l}$ for $l \in \{1, \ldots, L\}$ represent the ratio between the considered subband size and the original image size. In the following, we note $\pi_{i}$ the weighting of the subband $i$, and $\sigma_{Q_{i}}^{2}$ its distortion. Thus,

$$\pi_{i} = \frac{\|\hat{y}_{i}\|^{2(l-1)}}{2^{l}} \|\tilde{y}\|^{2} \quad \forall i \in \{1, \ldots, L\}$$

$$\pi_{L+1} = \pi^{#SB} = \frac{\|\hat{y}_{i}\|^{2L}}{2L}$$

In the case of hexagonal sampled data, Payan and Antonini compute weightings in a similar way [17].

4. MODEL-BASED BIT ALLOCATION

The output bit-rate can be expressed as the following weighted sum:

$$R_{output} = \sum_{i=1}^{#SB} a_{i} R_{i}$$

with $R_{i}$ the output bit-rate for subband $i$ and $a_{i}$ the weight of subband $i$ in the total bit-rate ($a_{i}$ is the ratio of the size of subband $i$ divided by the size of the image).

The solution of the entropy constrained bit allocation can be obtained thanks to Lagrangian operators by minimizing the following criterion:

$$J_{\lambda}(\{q_{i}\}) = \sum_{i=1}^{#SB} \pi_{i} \sigma_{Q_{i}}^{2} (q_{i}) + \lambda \left( \sum_{i=1}^{#SB} a_{i} R_{i} (q_{i}) - R_{T} \right)$$

where $R_{T}$ denotes the target output bit-rate and both $R_{i}$ and the distortion $\sigma_{Q_{i}}^{2}$ depend on the quantization step we are looking for.

An efficient way to minimize (7) without pre-quantizing subbands is to model the distribution $p(x)$ of the wavelet coefficients and use theoretical models for both distortion and bit-rate. In each subband, the probability density function of the wavelet coefficients can be approximated with a generalized Gaussian [18]. Therefore, we have

$$p(x) = ae^{-|bx|^e}$$

with $b = \frac{1}{\pi} \sqrt{\frac{\Gamma(1)}{\Gamma(1/e)}}$ and $a = \frac{\Gamma(1/e)}{\pi^{1/e}}$. We also assume that wavelet coefficients are independent and identically distributed (i.i.d.) in each subband.

For generalized Gaussian distributions, we have shown in [7, 19] that (7) becomes

$$J_{\lambda}(\{q_{i}\}) = \sum_{i=1}^{#SB} \pi_{i} \sigma_{Q_{i}}^{2} (q_{i}) + \lambda (\sum_{i=1}^{#SB} a_{i} R_{i} (q_{i}) - R_{T})$$

with $\tilde{q}_{i} = \frac{q_{i}}{a_{i}}$. The normalized distortion function $D$ depends only on the shape parameter of subband $i$.

To find the minimum of $J$, we differentiate it with respect to $q_{i}$ and $\lambda$. This provides a set of $\#SB + 1$ equations and $\#SB + 1$ unknowns which can be solved using the algorithms proposed in [7] or [19].

5. APPLICATION

![Quincunx Image: Nice.](image)

In this section, we show some results obtained with the proposed method for image Nice shown on figure Fig. 1 (SPOT5 acquisition simulation of the city of Nice provided by CNES, the French Space Agency of Toulouse). This image is coded with 10 bpp [20]. We provide the results of the proposed method for the (4,2) and (6,2) quincunx lifting...
filters. Fig. 2 shows that the weighted version of the model-based bit allocation procedure provides SNR improvements belonging to the range 0.1 dB to 0.2 dB. These results show a performance improvement when one compute the real distortion weights. Thus, there is a real qualitative interest to compute distortions with good weightings. Furthermore, the computation of the optimal weightings has to be done only once for a given set of filters. It does not increase the complexity of the encoder nor the decoder.

6. CONCLUSION

In this paper, we have shown how to compute the weightings of the quantization distortions of each subband in order to evaluate the global reconstruction distortion for the compression of quincunx sampled images. We have used the proposed weightings in the bit allocation process of a codec using the quincunx lifting scheme. The results for the (4,2) and (6,2) quincunx lifting schemes show SNR improvements belonging to the range 0.1 to 0.2 dB. Furthermore, our method does not increase the codec complexity.

7. REFERENCES