Performance Evaluation of Subband Coding and Optimization of Its Filter Coefficients

Jiro Katto and Yasuhiko Yasuda

Institute of Industrial Science, University of Tokyo,
7-22-1 Roppongi, Minato-ku, Tokyo 106, JAPAN.

ABSTRACT

In this paper, two analytical methods to evaluate coding performance of subband coding are proposed, and optimization of its filter coefficients from the viewpoint of energy compaction property is considered. The first method is based on matrix representation of subband coding in time domain, where the coding gain given by Jayant and Noll is introduced as a performance measure for filter banks with orthogonal property. The second method is based on optimum bit allocation problem for subband coding (multirate filter bank), where the unified coding gain is derived as a new performance measure which can be applied to arbitrary transform techniques. We then try to find filter coefficients which maximize the unified coding gain according to input characteristics. This approach leads to optimization of filter coefficients from the viewpoint of energy compaction property.

1 INTRODUCTION

Recently, subband coding of images has attracted attention of many researchers [1]-[4], because it has more advantages than other coding schemes in addition to its good energy compaction property: e.g. perceptual improvement of picture quality without blocking distortions. Its relationship to transform coding or that to hierarchical coding [5][6] are also indicated individually as the multirate filter banks [7] or the wavelet transforms [8], respectively.

However there exist few methods, except simulations, to evaluate the coding performance according to the number of subbands or the way to split subbands, nor to compare those of the quadrature mirror filter (QMF) [9], the conjugate quadrature filter (CQF) [10], and the symmetric short kernel filter (SSKF) [3] etc. Recently, a paper with similar objectives to ours was presented by Akansu and Liu [11]. However they focused only on the orthogonal and one dimensional filter banks.

The purpose of this manuscript is therefore placed on theoretical evaluation of the energy compaction property of arbitrary subband coders [12]. Section 2 describes a method based on matrix representation of subband coding in time domain [2][7], where the coding gain [13] is introduced as a performance measure for the filter banks with orthogonal property such as QMF and CQF. Section 3 presents another method based on optimum bit allocation problem for subband coding (multirate filter bank), where the unified coding gain is derived as a new performance measure which can be applied to arbitrary (non-orthogonal) transform techniques. In section 4, performance comparison is carried out and it will be suggested that some non-orthogonal filter banks such as SSKF can outperform the orthogonal ones as long as the number of subbands is limited. In section 5, we then try to find filter coefficients which maximize the unified coding gain on the perfect reconstruction (PR) condition of filter banks. This approach leads to optimization of filter coefficients from the viewpoint of energy compaction property.
2 PERFORMANCE EVALUATION BASED ON MATRIX REPRESENTATION OF SUBBAND CODING IN TIME DOMAIN

2.1 Matrix representation of subband coding in time domain

Consider one-dimensional two-band analysis/synthesis system as shown in Fig. 1. Then, the analysis and synthesis operations can be represented by matrix operation in time domain as

\[ y = H_1 x \]  
\[ \hat{x} = G_1 y \] (1) (2)

respectively, where \( x, y \) and \( \hat{x} \) are defined by

\[ x = (\cdots, x(0), x(1), x(2), \cdots)^t \]
\[ y = (\cdots, y_0(0), y_0(1), \cdots, y_1(0), y_1(1), \cdots)^t \]
\[ \hat{x} = (\cdots, \hat{x}(0), \hat{x}(1), \hat{x}(2), \cdots)^t \] (3)

(\( t \) represents transpose) and the matrices \( H_1 \) and \( G_1 \) are denoted by

\[
\begin{bmatrix}
  h_0(L-1) & \cdots & \cdots & h_0(0) \\
  h_0(L-1) & h_0(2) & h_0(1) & h_0(0) \\
  \cdots & h_0(4) & h_0(3) & h_0(2) & h_0(1) \\
  h_1(L-1) & \cdots & \cdots & h_1(0) \\
  h_1(L-1) & h_1(2) & h_1(1) & h_1(0) \\
  \cdots & h_1(4) & h_1(3) & h_1(2) & h_1(1)
\end{bmatrix}
\]

and

\[
\begin{bmatrix}
  g_0(0) & \cdots & \cdots & g_0(0) \\
  g_0(1) & g_0(0) & \cdots & \cdots \\
  g_0(1) & g_0(2) & g_0(1) & \cdots \\
  \cdots & \cdots & \cdots & \cdots \\
  g_0(L-1) & g_0(L-2) & g_0(L-3) & g_0(L-1) \\
  g_0(L-1) & g_0(L-2) & g_0(L-3) & g_0(L-1)
\end{bmatrix}
\]

![Diagram](image)

**Fig 1:** One-dimensional two-band analysis/synthesis system
Fig 2: Matrix representation of subband coding: (a) $H_1$, (b) $H_2$ in case of the full band splitting, (c) $H_2$ in case of the pyramidal splitting and (d) $H$ of block transforms.

respectively, where $L$ represents filter length.

Next, consider tree-structured subband coders with $K$ stages. Let $H$ be total band splitting (analysis) matrix. Then $H$ is given by

$$H = H_K \cdots H_2 H_1$$

where $H_k$ is the band splitting matrix for the $k$-th stage. Fig.2 shows examples of $H_k$, where (a) is $H_1$, (b) is $H_2$ in case of full band splitting (four bands), and (c) is $H_2$ in case of pyramidal splitting (three bands) corresponding to the wavelet transforms. Similarly, the synthesis matrix $G$ is denoted by

$$G = G_1 G_2 \cdots G_K$$

where $H_k$ and $G_k$ have to satisfy $G_k \cdot H_k = I$ to realize exact reconstruction.

Meanwhile, Fig.2(d) shows the analysis matrix for block transforms such as discrete cosine transform (DCT). Therefore, linkage of transform coder to subband coder can be also represented by the Eq.(6).

2.2 Verification of orthogonality

When the analysis matrix $H$ satisfies orthogonal condition

$$H \cdot H^t = H^t \cdot H = I$$

this subband coder can be considered as a special version of orthogonal transforms, where its block size is equal to the total number of inputs. Examples of filter banks which satisfy this orthogonality condition are QMF and CQF, where the synthesis matrix $G$ is given by $G = H^t$. On the other hand, SSKF never satisfies this orthogonal condition though it exactly realizes perfect reconstruction ($G \cdot H = I$). (Strictly speaking, QMF is quasi-orthogonal though CQF is exactly orthogonal as related to the orthonormal wavelet basis. QMF’s quasi-PR property is related with improvement of the degree of its orthogonality.)

Therefore, typical filter banks can be classified as follows:

- CQF (··· orthonormal wavelet): orthogonal, perfect reconstruction
- QMF: quasi-orthogonal, quasi-perfect reconstruction
- SSKF: non-orthogonal, perfect reconstruction
2.3 Application of the coding gain

When $H$ is the orthogonal matrix, its energy compaction property can be evaluated quantitatively by the coding gain [13], which is defined by

$$G_{TC} = \frac{1}{N} \sum_{k=0}^{N-1} \sigma_k^2 \left( \prod_{k=0}^{N-1} \sigma_k^2 \right)$$

(9)

where $\sigma_k$ is variance of the $k$-th transform coefficient, which is obtained as a diagonal component of $R_{xx} = H R_{xx} H^t$ where $R_{xx}$ is the autocorrelation matrix of inputs. For AR(1) inputs, this $R_{xx}$ is given by

$$R_{xx} = \begin{bmatrix}
1 & \rho & \rho^2 & \cdots & \rho^{N-1} \\
\rho & 1 & \rho & & \\
\rho^2 & \rho & 1 & & \\
\vdots & & & \ddots & \\
\rho^{N-1} & & & & 1
\end{bmatrix}$$

(10)

where $\rho$ denotes the correlation coefficient of inputs, and $N$ does the total number of inputs (size of the analysis matrix $H$).

We give results based on this approach in section 4 (Fig.4(f)). This approach can reflect the effect of border problem owing to the finite number of inputs, which is generally dealt with by the circular convolution method [4]. However, it cannot be applied to non-orthogonal filter banks such as SSKF.

3 PERFORMANCE EVALUATION BASED ON OPTIMUM BIT ALLOCATION PROBLEM FOR MULTIRATE FILTER BANK

3.1 Optimum bit allocation problem for multirate filter bank

Fig.3 shows the multirate filter bank where subsampling ratios are not restricted to $K : 1$, i.e.

- input $x(n)$ is separated into $K$ subbands, $y_k(n)$ ($k = 0, 1, \cdots, K - 1$)
- $y_k(n)$ is quantized into $u_k(n)$
- output $\hat{x}(n)$ is reconstructed from every $u_k(n)$ which includes quantization errors.

Here, we consider optimum bit allocation problem for this codec as follows.

A reconstruction error $r(n)$ is given by $x(n) - \hat{x}(n)$ and a quantization error of the $k$-th subband $q_k(n)$ is given by $y_k(n) - u_k(n)$, respectively. Then, define parameters $A_k$ and $B_k$ which satisfy

$$\sigma_{y_k}^2 = A_k \cdot \sigma_x^2$$

(11)
\[ \sigma_r^2 = \sum_{k=0}^{K-1} B_k \cdot \sigma_{q_k}^2 \]  

(12)

where \( \sigma \) represents variance. Every \( A_k \) is determined from both the correlation coefficients of inputs and the filter coefficients \( \{ h_k(n) \} \) at a transmitter side. On the other hand, every \( B_k \) is determined from both the correlation coefficients of quantization errors and the filter coefficients \( \{ g_k(n) \} \) at a receiver side. (Notice that, assuming that quantization errors have no correlations each other, \( B_k \) is given by square-sum of \( g_k(n) \) multiplied by \( \alpha_k \) shown below) Next, let \( N \) and \( N_k \) be the total number of inputs and that of \( y_k(n) \), respectively, and then define a parameter \( \alpha_k \) by \( N_k/N \), where

\[ \sum_{k=0}^{K-1} \alpha_k = 1 \]  

(13)

is satisfied due to the critical sampling.

Based on the notations above, consider optimum bit allocation problem as follows: under the constant rate constraint

\[ \sum_{k=0}^{K-1} \alpha_k R_k = R(\text{const}) \]  

(14)

, minimize

\[ \sigma_r^2 = \sum_{k=0}^{K-1} B_k \sigma_{q_k}^2 \]  

(15)

where \( R_k \) is bit rate for the \( k \)-th subband. By substituting the approximate relationship [13]

\[ \sigma_{q_k}^2 \simeq \epsilon^2 \sigma_{y_k}^2 \]  

(16)

where \( \epsilon \) is a constant depending on the input characteristics and using the Lagrange multiplier method, minimum value of the reconstruction error variance is given by

\[ \min \{ \sigma_r^2 \} = \prod_{k=0}^{K-1} \left( \frac{A_k B_k}{\alpha_k} \right)^{\alpha_k} \cdot \epsilon^2 \sigma_{y_k}^2, \]  

(17)
3.2 Derivation of the unified coding gain

The coding gain $G_{TC}$ is defined by the ratio of reconstruction error variance of orthogonal transform to that of PCM. Similarly, since the reconstruction error variance of PCM is given by $e^{2 - 2R} \sigma_x^2$ [13], a new performance measure $G_{SBC}$ for Fig.3 is given by

$$G_{SBC} = \frac{1}{\prod_{k=0}^{K-1} \left( \frac{A_k B_k}{\alpha_k} \right)^{\alpha_k}}.$$  \hspace{1cm} (18)

Note that this performance measure never depends on filters' orthogonality (i.e. the effect of $g_k(n)$ is reflected on $G_{SBC}$ dissimilar to the case of $G_{TC}$) and can be applied to all the linear transforms including DPCM and orthogonal transforms. For example, in case of the first order closed-loop DPCM, $G_{DPCM} = (1 - \rho^2)^{-1}$ is obtained because $K = 1$, $\alpha_0 = 1$, $A_0 = 1 - \rho^2$ and $B_0 = 1$. Therefore, we call this measure "unified coding gain".

We give results based on this approach in the next section. This approach can be applied not only to one dimensional arbitrary filter banks, but also to two dimensional ones without any modifications.

4 RESULTS AND DISCUSSIONS

4.1 One dimensional case

Fig.4 shows energy compaction properties of several one-dimensional subband coders. Here we assume AR(1) inputs, assume that quantization errors have no correlations one another, and fix the correlation coefficient $\rho$ at 0.95. Horizontal axes represent the number of stages $k$ for tree-structured subband coder: e.g. in case of full band splitting, $k = 2$ corresponds to four subbands and $k = 3$ does to eight, respectively. The number in the parenthesis indicates to filter length: in case of SSKF($m \times n$), $m$ means that of its lowpass filter and $n$ does that of its highpass filter, respectively.

Fig.4(a) shows performance comparison among CQF(16), QMF(16), orthonormal wavelet(16), SSKF(5×3) and SSKF(3×5), where the pyramidal splitting is assumed and coding gains of DCTs are appended for reference. From this figure, we can recognize that

- As band splitting is repeated, energy compaction properties of subband coders outperform those of DCTs except SSKF(3×5).
- As long as the number of subbands is limited, SSKF(5×3) outperforms the orthogonal filter banks although it is non-orthogonal and has far shorter taps.

The latter results can be explained qualitatively as follows: since a highpass filter of the SSKF(5×3) is given by $\{-1/2, 1, -1/2\}$, this structure can be considered as interpolative DPCM for high correlated inputs which performs better than the extrapolative case. Therefore, SSKF(5×3) progressively removes correlations as band splitting is repeated. On the contrary, energy compaction property of the SSKF(3×5) is not so good because its highpass filter is not fit for the input characteristics.
Fig 4: Energy compaction properties of one dimensional subband coders for AR(1) inputs ($\rho=0.95$): (a) CQF(16), QMF(16), wavelet(16), SSKFs and DCTs, (b) full band splitting vs. pyramidal splitting in case of CQF(16), (c) in case of SSKF(5x3), (d) subband coders followed by DCT(8) in case of CQF(16), (e) in case of SSKF(5x3), and (f) the border effect in case of CQF(16) ($N=256$).
Fig 5: Proposed two-dimensional perfect-reconstruction non-separable filter bank

Fig.4(b) and (c) shows performance comparison between the full band splitting and the pyramidal splitting for CQF(16) and SSKF(5×3), respectively. (We use CQF(16) since it performs best among the orthogonal filter banks) Fig.4(b) illustrates that the pyramidal splitting performs almost similar to the full band splitting in case of the CQF(16). This result convinces us of validity of the hierarchical coding or the wavelet transforms from the viewpoint of energy compaction. In addition, Fig.4(c) suggests that the pyramidal splitting is absolutely effective in case of the SSKF(5×3). This is because its highpass signal can be considered as a prediction error (almost no correlation) in a sense as described above, and its additive band splitting causes negative gains because of its non-orthogonality. Quantitatively speaking, $G_{SBC}$ of SSKF(5×3) for $\rho=0$ is -0.65dB (negative) though those of all the orthogonal filter banks are always 0dB.

Fig.4(d) and (e) shows the effect of DCT which follows subband splitting by CQF(16) and SSKF(5×3), respectively. Here DCT(8) is considered, and 1. the full band splitting vs. the pyramidal splitting, and 2. DCT to all subbands vs. that to only the lowest frequency, are compared. Fig.4(d) shows that coding gains of CQF + DCT approach the theoretical optimum value $(1-\rho^2)^{-1}$ as the number of subbands increases, but Fig.4(e) shows that those of SSKF + DCT saturate around 9.7dB. In either case, it is not necessary to apply additive data compression techniques like DCT to higher frequency subbands.

Finally, Fig.4(f) shows the effect of border problem for CQF(16) which is generally dealt with by the circular convolution method. Here, both the tools in the Eq.(9) and the Eq.(18) are used, where the number of inputs, i.e. matrix size, is fixed at 256 in the former tool. Also the pyramidal splitting is assumed and DCT(8) is connected to only the lowest frequency. This figure gives quantitative evaluation of the performance degradation which results from the circular convolution.

4.2 Two dimensional case

We first propose a new PR non-separable filter bank as shown in Fig.5. Here, $h_1(m,n)$ is organized as a two dimensional interpolation filter similar to the SSKF(5×3) on the PR condition for the quincunx subsampling, which is given by

\[ \sum_k \sum_l h_0(k,l) g_0(2i-k, 2j-l) = \delta_i \delta_j \]  \hspace{1cm} (19)

where $h_1(m,n) = (-1)^{m+n} g_0(m,n)$ and $g_1(m,n) = (-1)^{m+n} h_0(m,n)$ are formed for aliasing cancellation.
Fig 6: Energy compaction properties of two dimensional subband coders: (a) for the isotropic inputs, and (b) for the separable inputs ($\rho=0.95$).

Table 1: Energy compaction properties of two dimensional subband coders followed by DCT(8)

<table>
<thead>
<tr>
<th></th>
<th>CQF+DCT(8)</th>
<th>SSKF+DCT(8)</th>
<th>DCT(8)</th>
<th>DCT(16)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>all</td>
<td>low</td>
<td>all</td>
<td>low</td>
</tr>
<tr>
<td>isotropic</td>
<td>12.70</td>
<td>12.11</td>
<td>11.86</td>
<td>11.83</td>
</tr>
<tr>
<td>separable</td>
<td>19.57</td>
<td>15.42</td>
<td>18.83</td>
<td>15.69</td>
</tr>
</tbody>
</table>

Next, we consider two input models assuming the rectangular sampling as follows:

1. isotropic $\cdots \rho_{x,y} = \rho \sqrt{x^2+y^2}$
2. separable $\cdots \rho_{x,y} = \rho |x| + |y|$

where $\rho_{x,y}$ is the two dimensional correlation coefficient. It was made clear that the latter can be more compressed than the former with DPCM [13].

Assuming $\rho=0.95$, Fig.6 shows performance comparison among CQF(16), SSKF(5x3) and the non-separable filter bank shown in Fig.5 for the two input models. Horizontal axes represent the number of stages for the separable 2-D filters (CQF and SSKF), and results of the non-separable filter are superposed so that their lowest subbands coincide when ideal filters are used. Fig.6(a) shows that, for the isotropic inputs, the pyramidal splitting is effective for non-orthogonal filter banks (SSKF and Fig.5) similar to the one dimensional case. On the contrary, Fig.6(b) shows that, for the separable inputs, the pyramidal splitting is not effective and energy compaction property of the non-separable filter bank is not so good.

These results suggest that, interpolation filters (highpass filters) of non-orthogonal filter banks cannot
remove directional correlations sufficiently such as the separable model. The non-separable filter is not fit for
the separable model particularly in the even order’s stages because their predictive direction is sloped at the
angle of 45° to the input.

Finally, Table 1 shows performance comparison of two dimensional subband coders followed by DCT(8),
where only separable filters CQF(16) and SSKF(5×3) are considered, and band splitting is carried out only
one times (four bands). In this table, all means that DCT(8) is linked to all subbands and low means that it
is linked to only the lowest frequency.

5 OPTIMIZATION OF FILTER COEFFICIENTS FROM THE VIEWPOINT OF ENERGY COMPACTION PROPERTY

5.1 Perfect reconstruction conditions in time domain

The PR condition of the one-dimensional two-band analysis/synthesis system (Fig.1) is given, with aliasing
cancellation conditions $G_0(z) = H_1(-z)$ and $G_1(z) = -H_0(-z)$, by $H_0(z)G_0(z) - H_0(-z)G_0(-z) = 2 \cdot z^{-m}$.
This equation can be rewritten in time domain as

$$\sum_n h_0(n)g_0(2k - 1 - n) = \delta_k$$  \hspace{1cm} (20)

which is called biorthogonality. When we add $H_1(z) = -H_0(-z^{-1})z^{-L}$ to the above conditions ($L$ is filter
length), the PR condition with orthogonal property is given by

$$\sum_n h_0(n)h_0(n + 2k) = \delta_k.$$  \hspace{1cm} (21)

We try to find filter coefficients which maximize the unified coding gain under the Eq.(20) or the Eq.(21)

5.2 Biorthogonal cases

We focus our attention on the odd-tap and linear-phase $(4n-3)\times 3$ type’s PR filter banks where \{g_0(n)\} is
fixed at \{q, 1, q\} as an analogy to the SSKF(5×3). (Note that (4n-1)×3 type reduces to (4n-3)×3 type due
to the Eq.(20)) For reference, G_SBC of QMF(16) for $p = 0.95$ is 5.920(dB), that of CQF(16) is 5.922(dB),
that of wavelet(16) is 5.891(dB), and that of SSKF(5×3) is 6.277(dB).

At $n=1$, \{h_0(n)\} which satisfies Eq.(20) is only \{1\}. Then varying the value of $q$, $G_{1×3}$ for $p = 0.95$ is
maximized at $q = 0.490$ to 5.586(dB). When $n=2$ and \{h_0(n)\} = \{c, b, a, b, c\}, $G_{5×3}$ is maximized at $q=0.51$
and $b=0.29$ to 6.307(dB) as shown in Fig.7(a) where \(a, c\) are determined automatically from \(q, b\) based on
the Eq.(20). When $n=3$ and \(h_0(n) = \{c, d, c, b, a, b, c, d, e\\}$, $G_{9×3}$ is maximized at $q = 0.51$, $b = 0.29$ and $d = -0.04$ to 6.322(dB). Fig.7(b) illustrates the optimum coding gains for the $(4n-3)\times 3$ type’s PR filter banks.
From this figure, we can recognize that

- Coding gains almost saturate even though filter length is increased.
- Coding gain of the SSKF(5×3) is pretty high.
Fig 7: Optimization of filter coefficients in the biorthogonal case: (a) 5×3 type's PR filter banks, and (b) (4n−3)×3 type's PR filter banks

5.3 Orthogonal cases

Under the Eq.(21), the CQF is obtained with the spectral factor decomposition techniques. On the other hand, the orthonormal wavelet basis is obtained with the so called regularity condition. Our approach is another one because maximization of the unified coding gain is imposed in place of these constraints.

Let L be filter length, where L must be even from the Eq.(21). In case of L=2, \( h_0(n) = \{1/\sqrt{2},1/\sqrt{2}\} \) is obtained, where no input characteristics are reflected. When \( L=4 \) and \( h_0(n) = \{a,b,c,d\} \), the optimization problem is summarized as follows: under

\[
a^2 + b^2 + c^2 + d^2 = 1
\]

\[
ac + bd = 0
\]

find \( \{a,b,c,d\} \) which maximize

\[
(ab + bc + cd) + ad \cdot \rho^2
\]

where AR(1) input is assumed (This problem is now being continued). One result is obtained, when the equal contribution constraints \( a + c = b + d \) [6] is appended, as

\[
a = \frac{\sqrt{2} + \sqrt{6}}{8}, \quad b = \frac{3\sqrt{2} + \sqrt{6}}{8}, \quad c = \frac{3\sqrt{2} - \sqrt{6}}{8}, \quad d = \frac{\sqrt{2} - \sqrt{6}}{8}
\]

which coincides to the orthonormal wavelet basis in case of \( L=4 \) [8].
6 CONCLUSIONS

This paper clarifies the energy compaction property of subband coding with a new performance measure called unified coding gain. Main results are

1. Both the iterative subband partitioning and the linkage of DCT outperform the single DCT.

2. Some biorthogonal short tap filter banks can outperform orthogonal ones as far as the number of subbands is limited.

These theoretical evaluations will help us to design an efficient subband codec.

As future works, optimization of the orthogonal filter banks from the viewpoint of energy compaction property in section 5.3 should be solved.

REFERENCES


