



Audio Engineering Society

Convention Paper

Presented at the 121st Convention
2006 October 5–8 San Francisco, CA, USA

This convention paper has been reproduced from the author's advance manuscript, without editing, corrections, or consideration by the Review Board. The AES takes no responsibility for the contents. Additional papers may be obtained by sending request and remittance to Audio Engineering Society, 60 East 42nd Street, New York, New York 10165-2520, USA; also see www.aes.org. All rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.

Fast Complex Quadrature Mirror Filterbanks for MPEG-4 HE-AAC

Han-Wen Hsu, Chi-Min Liu, and Wen-Chieh Lee

PSPLab, Department of Computer Science, National Chiao Tung University, Hsinchu, 300, Taiwan
cmliu@csie.nctu.edu.tw

ABSTRACT

Spectral Band Replication (SBR) has been introduced in MPEG-4 HE-AAC as bandwidth extension tool. All the framework of SBR is on complex-value domain to avoid the aliasing effect, and hence results in considerable time complexity. This paper focuses on the complex Quadrature Mirror Filter (QMF) banks used in HE-AAC encoder and decoder, and proposes the two fast decomposition methods, that are based on DCT-IV and DFT respectively, for the time-consuming matrix operations in the filterbanks. Therefore, the time complexity can be effectively reduced by the fast algorithms for DCT-IV and FFT.

1. INTRODUCTION

Spectral Band Replication (SBR) [1]-[3] has been combined with MPEG AAC as bandwidth extension tool. With SBR module taking care of the HF contents, the conventional AAC encoder can compress the LF part using most of the available bits. The resulting scheme is referred to as the MPEG-4 High Efficient (HE) AAC or AACplus. Based on the harmonic phenomenon, the SBR reconstructs high bands by reproducing and adjusting the replicated low bands perceptually similar to original ones. There are three main reconstruction procedures in SBR process. At first, in the HF regeneration module, the low QMF bands analyzed from the decoded LF AAC signal are replicated to HF and further inversely filtered to clip the unwanted tones from LF. At the second phase, the envelope of the replicated bands is scaled to fit the original one that is recoded by the average energy data in the T/F grids. Finally, the additional compensation of

tones and noise can be added to adjust the tonality of the reconstructed signal.

However, the cosine QMF bank [4] that is widely used in perceptual audio coding is unsuitable for the SBR process. This is because any envelope adjustment on the cosine QMF bands will result in aliasing effect. Therefore, to mitigate the aliasing effect by the advantage of oversampling, the SBR adapts a complex analysis QMF bank to separate the input real signal into 64 complex-value subbands. In the decoder part, there is a complex synthesis QMF bank to combine 64 complex-value subbands back to a real decoded signal. On the other hand, to obtain low bands from the conventional AAC decoded LF signal, there is also a 32-band complex analysis QMF bank. Furthermore, there is a complex downsampled synthesis QMF bank that can synthesize only the 32 lower subbands and simultaneously decimate the synthesized decoded signal by the factor of two.

For the polyphase implementation of the four different QMF banks, the kernel time-consuming matrix operations determine dominantly the time complexity. Furthermore, all the matrix operations are the SDFTs (shifted discrete time transform) with different time/frequency shifts [5]. The four matrix operations are summarized. Also, in the following context, the M notation represents the number of the subbands split in SBR, that is $M = 64$.

Analysis matrix operation in encoder

For $2M$ -point real-value input $x[n]$, the M -point output $X[k]$ is computed by

$$X[k] = \sum_{n=0}^{2M-1} x[n] \exp\left(\frac{i\pi(k+0.5)(2n+1)}{2M}\right), \quad (1)$$

for $k = 0 \sim M-1$.

Analysis matrix operation in decoder

For M -point real-valued input $x[n]$, the $M/2$ -point output $X[k]$ is computed.

$$X[k] = \sum_{n=0}^{M-1} x[n] \exp\left(\frac{i\pi(k+0.5)(2n-0.5)}{M}\right), \quad (2)$$

for $k = 0 \sim \frac{M}{2} - 1$.

Synthesis matrix operation in decoder

For M -point complex-valued input $X[k]$, the $2M$ -point real-value output $x[n]$ is computed.

$$x[n] = \operatorname{Re}\left\{\sum_{k=0}^{M-1} X[k] \exp\left(\frac{i\pi(k+0.5)(2n-4M+1)}{2M}\right)\right\}, \quad (3)$$

for $n = 0 \sim 2M-1$.

Downsampled Synthesis matrix operation in decoder

For $M/2$ -point complex-valued input $X[k]$, the M -point real-value output $x[n]$ is computed.

$$x[n] = \operatorname{Re}\left\{\sum_{k=0}^{M-1} X[k] \exp\left(\frac{i\pi(k+0.5)(2n-2M+0.5)}{M}\right)\right\}, \quad (4)$$

for $n = 0 \sim M-1$.

The SDFT operations decide the efficiency of both encoder and decoder, and hence require effective techniques to reduce the time complexity. Comparing the direct matrix operation, a trivially more efficient computation method is to decompose the SDFT into three stages, pre-butterfly, DFT, and post-butterfly operation. However, the fact of the analyzed input signal or synthesized output signal is real is not exploited to further improve the efficiency. By the conjugate symmetry property of SDFT for real samples,

there are possible faster decomposition methods to reduce the computational complexity of the QMF banks.

In the open source FAAD [6], the fast algorithms based on DCT-IV (type IV of discrete cosine transform) for decoder analysis and decoder synthesis filterbanks have been implemented. This paper develops all algorithms and related derivations based on DCT-IV. On the other hand, to maintain the unified framework for the codec systems adapting FFT as the basic computation module, the paper develops the fast algorithms based on FFT (fast furrier transform). Especially, the point numbers of the FFT are only half of the ones in the simple decomposition method. Through the available fast methods for DCT and FFT, the complexity can be reduced from $O(N^2)$ to $O(N \cdot \log N)$.

The paper is organized as follows. In Section 2, the basic principle of complex QMF bank and the comparison with cosine QMF bank is reviewed. In Section 3, the fast algorithms based on DCT and FFT are presents. Furthermore, for convenience, some notations are defined in Table 1 for the following context.

W_M	$e^{-\frac{i2\pi}{M}}$, for arbitrarily positive integer M .
W_a^b	$e^{-\frac{i2\pi b}{a}}$, for arbitrarily real number a, b .
C_a^b	$\cos\left(\frac{2\pi \cdot b}{a}\right)$, al number a, b .
S_a^b	$\sin\left(\frac{2\pi \cdot b}{a}\right)$, al number a, b .
$\operatorname{Re}(\cdot)$	Return the real part of the input number.
$R_x[n], I_x[n]$	Return the real part and the image of $x[n]$, respectively.
$DCT_M^4(x[k])[n]$ $DST_M^4(x[k])[n]$	The n^{th} coefficient of the M -point DCT-IV and DST-IV of the input signal $x[n]$.
$DFT_M(x[k])[n]$ $IDFT_M(x[k])[n]$	The n^{th} coefficient of the M -point DFT and IDFT of the input signal $x[n]$.
$x^r[n]$	The reverse signal related to the original signal $x[n]$. Assume $x[n]$ is defined for $n = 0 \sim M-1$, then $x^r[n] = x[M-1-n]$ for $n = 0 \sim M-1$.

Table 1: Notations used in this paper

2. BASIC PRINCIPLE OF COMPLEX QMF BANK

This section first reviews the basic concept of the maximal decimated filter bank, and three main distortion errors. Furthermore, by comparing cosine QMF bank and complex QMF bank, the advantage of the aliasing resisting of complex QMF bank is explained. Finally, the essential principle of the construction of complex QMF bank is present.

2.1. *M*-channel Maximal Decimated Filter Bank

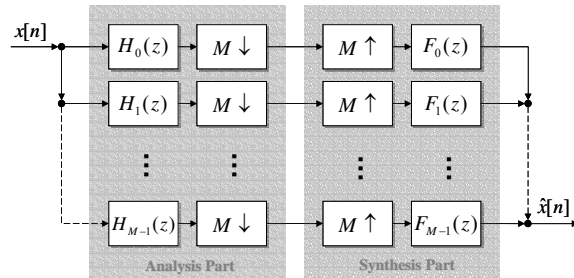


Figure 1: The *M*-channel QMF bank

The typical framework of the *M*-channel maximal decimated filter bank, also named *M*-channel QMF bank, is illustrated in Figure 1 [4]. In the analysis part, the input signal $x[n]$ is split into *M* channels which are decimated by *M* to maintain the critical sampling rates. In the synthesis part, after being expanded by *M*, the subbands are synthesized into the reconstructed signal $\hat{x}[n]$. By the advantage of the rate conversion, the multirate system supplies the efficient intermedium for signal processing between the analysis and synthesis parts.

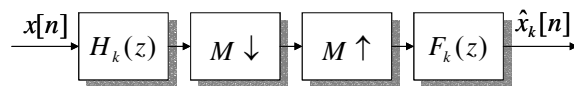


Figure 2: The k^{th} channel of the *M*-channel QMF bank

As shown in Figure 2, for the relation between the original subband and the k^{th} reconstructed subband with the same sampling rate is given by the following z-transform equation

$$\hat{X}_k(z) = \frac{1}{M} \cdot F_k(z) \cdot \sum_{l=0}^{M-1} \{H_k(zW_M^l) \cdot X(zW_M^l)\}. \quad (5)$$

Or, (5) can be equivalently decomposed into the distortion term and the aliasing term as (6)

$$\hat{X}_k(z) = T_k(z) \cdot X(z) + A_k(z) \cdot F_k(z), \quad (6)$$

where the k^{th} distortion function is given as (7)

$$T_k(z) = \frac{1}{M} \cdot F_k(z) \cdot H_k(z), \quad (7)$$

and the k^{th} aliasing function is given as (8)

$$A_k(z) = \frac{1}{M} \cdot \sum_{l=1}^{M-1} \{H_k(zW_M^l) \cdot X(zW_M^l)\}. \quad (8)$$

The affection of the aliasing terms that locate at the frequency can be eliminated mutually as far as possible in (9).

$$\sum_{k=0}^{M-1} F_k(z) H_k(zW_M^l) X(zW_M^l). \quad (9)$$

However, various signal processes, such as spectral band scaling in SBR, may destroy the effect of the aliasing concealing largely, and degrade the target outcome due to the aliasing interference.

Combine (6) of all the channels, the relation between the original and the resultant reconstructed signal is expressed by the following z-transform equation

$$\hat{X}(z) = T(z) \cdot X(z) + A(z), \quad (10)$$

where the distortion and the aliasing functions of the full system are given by (11) and (12) respectively,

$$T(z) = \sum_{k=0}^{M-1} T_k(z), \quad (11)$$

$$A(z) = \sum_{k=0}^{M-1} (A_k(z) \cdot F_k(z)). \quad (12)$$

For example, in an ideal case, the perfect reconstruction QMF bank, the conditions of the delay-only (13) and the aliasing-free (14) constraints should be satisfied.

$$T(z) = c \cdot z^{-n_0}, \quad (13)$$

$$A(z) = 0. \quad (14)$$

Another scheme, the pseudo QMF bank, is to eliminate the phase distortion completely and the aliasing distortion approximately, and to minimize the amplitude distortion.

2.2. Basic Concept of DFT, Cosine and Complex Modulated Filter Banks

All the three QMF banks are formed by a series of equidistant modulation of a prototype lowpass filter. Through numerous illustrations, this section will explain the essential concept of the tree QMF banks, and investigates the advantage of complex QMF on the elimination of aliasing suffering.

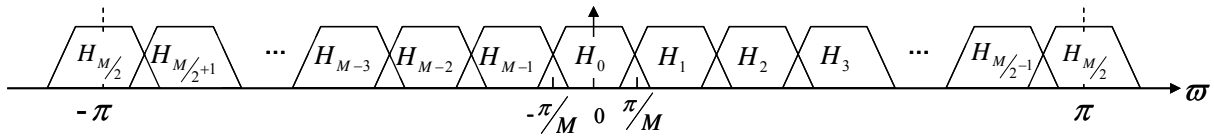


Figure 3: Illustration of M -channel DFT QMF bank

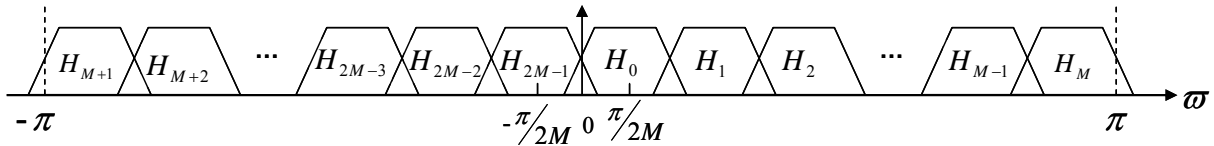


Figure 4: Shifted version of $2M$ -channel DFT QMF bank

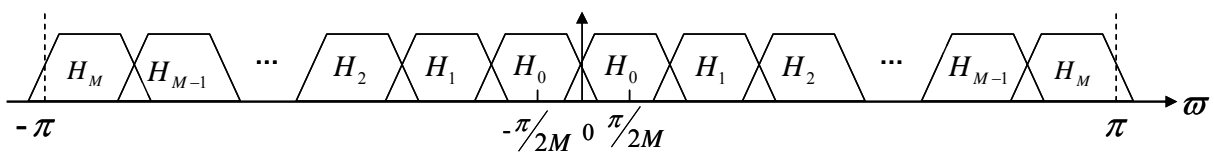


Figure 5: Illustration of M -channel cosine QMF bank

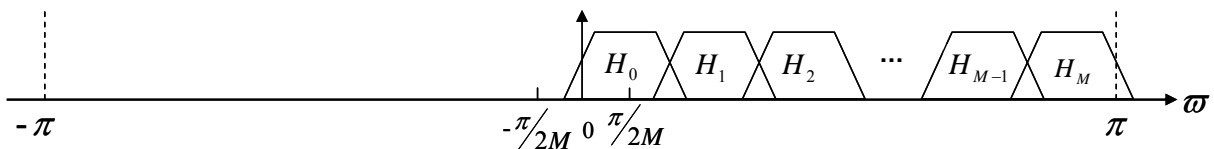


Figure 6: Illustration of M -channel complex QMF bank

The simplest modulated filter bank is DFT QMF bank, as shown in Figure 3. Based on a prototype lowpass filter, the $M-1$ modulated bands cover the interval $[-\pi, \pi]$. Except the 0^{th} and the $M/2^{th}$ bands, the remain $M-2$ bands occupy only either positive or negative frequency. The basic framework of the polyphase implementation of DFT QMF bank is illustrated in Figure 7, where possible phase adjustment is ignored for simplification. Furthermore, to construct a “real input real output” system, the M -band cosine QMF bank can be formed from $2M$ -band DFT QMF bank by combining the positive and negative band pair to conceal the image component. Also, for the symmetry to the zero frequency, the $2M$ -band DFT QMF bands need to be right shifted $\pi/2M$ as shown in Figure 4. The resultant cosine QMF bank is as shown in Figure 5, where each band occupies both positive and negative frequencies and is cosine modulated from the prototype filter. Figure 8 illustrates the basic polyphase implementation of cosine QMF bank modified directly from the DFT QMF bank’s, and also some possible phase adjustments after IDFT in analysis part and before the DFT in synthesis part is ignored for simplification.

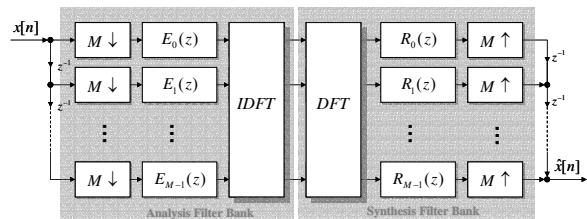


Figure 7: Basic polyphase implementation of DFT filter bank

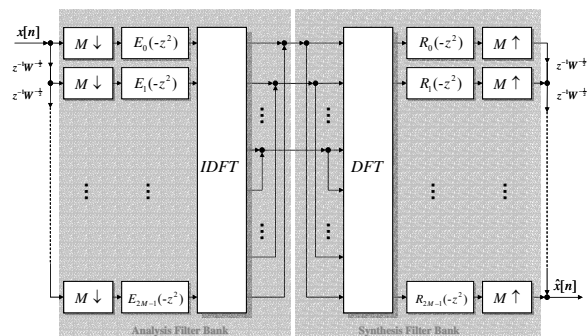


Figure 8: Basic polyphase implementation of cosine filter bank

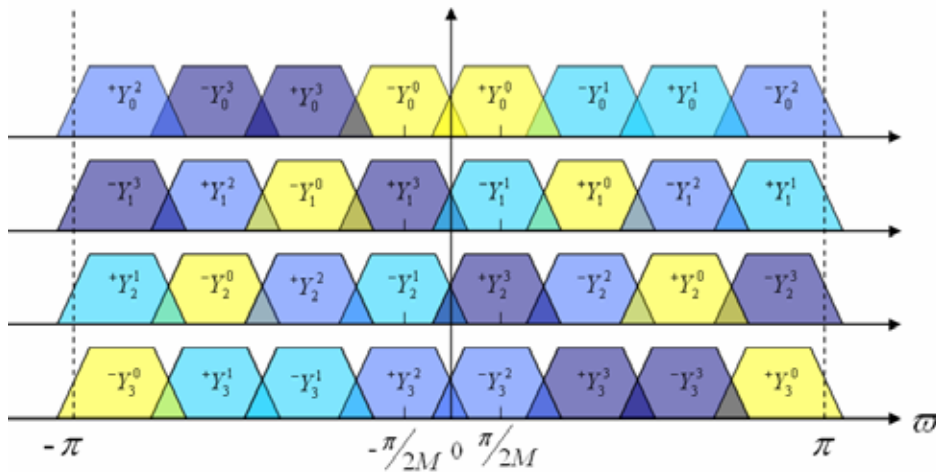


Figure 9: The aliasing terms of cosine QMF bank

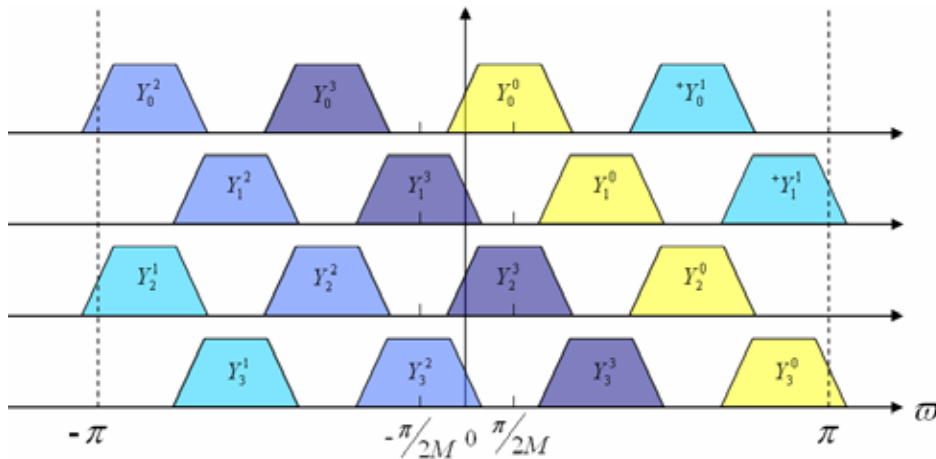


Figure 10: The aliasing terms of complex QMF bank

To consider the aliasing suffering of the cosine QMF bank, the aliasing functions (8) should be reviewed. Take for example, Figure 9 illustrates all the aliasing terms in the 4-band cosine QMF bank, where the notation $+Y_k^j$ and $-Y_k^j$ means the positive and the negative components of $H_k(zW_{2M}^j)X(zW_{2M}^j)$ respectively. Because the magnitude response of the synthesis filters are the same as the corresponding analysis filters, it is obvious by observing Figure 9 that several aliasing terms are not eliminated in (8). For example, the intersections between $+Y_0^0$, $-Y_0^1$ and $-Y_0^0$, $+Y_0^3$ cause the dominate aliasing component. The dominate aliasing component can be eliminated mutually by the appropriate phase adjustment on the filters. For instance, the aliasing of

$+Y_0^0$ and $-Y_0^1$ can be eliminated with the aliasing of $-Y_1^1$ and $+Y_1^0$. However, additional process, such as spectral equalization and the envelope adjustment process in the SBR, will destroy the original well balance of mutual elimination, and result in some annoying aliasing component. The main cause of the dominate aliasing components comes from the intersection of the positive part of the modulated band and the negative part of some aliasing term, or contrariwise. Therefore, the absence of the negative component in the complex QMF bank, as shown in Figure 6, exactly avoids the aliasing problem. Figure 10 illustrates all the aliasing terms in the 4-band complex QMF bank, where the notation Y_k^j means the aliasing term $H_k(zW_{2M}^j)X(zW_{2M}^j)$. Because the

modulation distance is $\pi/2M$ and the minimum distance between the modulate and the aliasing band is π/M , there is no dominate aliasing component captured by the related synthesis filter. Hence, the complex QMF bank has the inherent advantage to resist the aliasing interference, and is more suitable for some spectral process and application. Figure 11 illustrates the basic polyphase implementation of complex QMF bank modified directly from the DFT QMF bank's, and also some possible phase adjustments is ignored. Because the absence of the negative component, only the real part of the reconstructed signal need be captured.

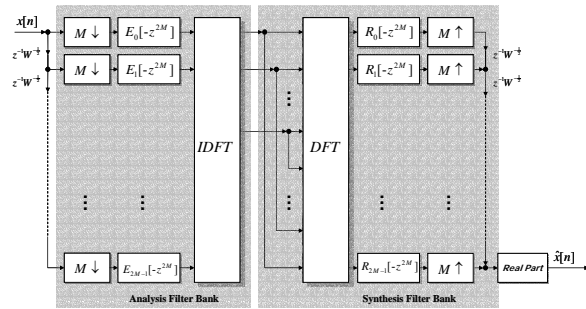


Figure 11: Basic polyphase implementation of complex filter bank

2.3. Complex Quadrature Mirror Filter Banks

In the following subsections, the basic construction principle is present for complex QMF bank. Furthermore, the close relation between cosine and complex QMF bank is also considered.

2.3.1. Construction of Complex QMF Bank

Assume the prototype lowpass filter $P(z)$ with cutoff frequency $\pi/2M$ is real-value, FIR and linear phase, and has order N . Its impulse response must have the symmetry as (15), and its frequency response can be expressed as (16) with the amplitude function $P_R(\varpi)$,

$$p[n] = p[N - n], \quad (15)$$

$$P(e^{j\varpi}) = e^{-j\varpi \frac{N}{2}} P_R(\varpi). \quad (16)$$

Furthermore, the related k^{th} modulated filter $Q_k(z)$ is given by

$$Q_k(z) = P(zW_{2M}^{k+0.5}), \quad k = 0 \sim M-1. \quad (17)$$

To make the design easier by maintaining the same phase as the prototype filter, the phase adjustment is

introduced into the modulated filters by the c_k constants, that is

$$U_k(z) = c_k \cdot Q_k(z), \quad k = 0 \sim M-1. \quad (18)$$

Moreover, the c_k constant should be chosen as

$$c_k = W_{2M}^{(k+0.5)\frac{N}{2}}, \quad k = 0 \sim M-1, \quad (19)$$

and it implies

$$U_k(e^{j\varpi}) = e^{-j\varpi \frac{N}{2}} \cdot P_R\left(\varpi - \frac{\pi \cdot (k + 0.5)}{M}\right). \quad (20)$$

for $k = 0 \sim M-1$. Because of linear phase, a well relation between U_k and its tilde pair U_k^* is implied as (21)

$$\tilde{U}_k(z) = z^N U_k^*(z). \quad (21)$$

To increase the design flexibility, the constants a_k, b_k are further introduced to generate the analysis and synthesis filters as

$$H_k(z) = a_k U_k(z), \quad (22)$$

$$F_k(z) = b_k U_k^*(z). \quad (23)$$

In order to avoid the phase distortion to archive (13), the synthesis filter can be chosen as the tilde of the related analysis filter from (7), that is

$$F_k(z) = z^{-N} \cdot \tilde{H}_k(z). \quad (24)$$

Therefore, from (21)-(24), a constraint of the a_k, b_k is the conjugate relation,

$$b_k = a_k^*. \quad (25)$$

As a result, the final distortion function T is given by

$$T(z) = \frac{1}{M} \sum_{k=0}^{M-1} |a_k|^2 U_k^2(z). \quad (26)$$

Also, once the value $e^{j\theta_k}$ on the unit cycle is chosen for a_k , the resultant distortion function T is simplified as

$$T(z) = \frac{1}{M} \sum_{k=0}^{M-1} U_k^2(z). \quad (27)$$

In other word, because both the aliasing suffering and the phase distortion have been eliminated, the main effort of the design can only focus on the U_k , or equivalently on the prototype filter, to conceal or minimize the amplitude distortion.

To summarize, under the prototype filter $P(z)$, the k^{th} modulated analysis and synthesis filters can be chosen

as (28) and (29) respectively to eliminate the aliasing suffering and the phase distortion, and only the amplitude distortion needs to consider by the design of the prototype filter.

$$H_k(z) = e^{j\theta_k} \cdot W_{2M}^{(k+0.5)\frac{N}{2}} \cdot P(zW_{2M}^{k+0.5}), \quad (28)$$

and

$$F_k(z) = e^{-j\theta_k} \cdot W_{2M}^{(k+0.5)\frac{N}{2}} \cdot P(zW_{2M}^{k+0.5}). \quad (29)$$

Also, their impose response can be expressed by

$$h_k(n) = p(n) \exp\left(i\left(\frac{\pi}{2M} \cdot (k+0.5)(2n-N) + \theta_k\right)\right), \quad (30)$$

and

$$f_k(n) = p(n) \exp\left(i\left(\frac{\pi}{2M} \cdot (k+0.5)(2n-N) - \theta_k\right)\right). \quad (31)$$

Another representation of (30)(31) is to merge the θ_k term into a close form as (32)(33)

$$h_k(n) = p(n) \exp\left(i\left(\frac{\pi}{2M} \cdot (k+0.5)(2n+\alpha)\right)\right), \quad (32)$$

and

$$f_k(n) = p(n) \exp\left(i\left(\frac{\pi}{2M} \cdot (k+0.5)(2n+\beta)\right)\right), \quad (33)$$

The relation between θ_k and α, β are as following

$$\theta_k = \frac{\pi}{2M} (k+0.5)(\alpha+N) + 2\pi\rho_1, \quad (34)$$

and

$$\theta_k = \frac{-\pi}{2M} (k+0.5)(\beta+N) - 2\pi\rho_2, \quad (35)$$

for some integers ρ_1, ρ_2 . Furthermore, from (34) and (35) it gives

$$(2k+1)(\alpha+\beta+2N) = -8M(\rho_1+\rho_2), \quad (36)$$

Therefore, the pair relation of α, β is as

$$\alpha+\beta+2N = 8M\rho. \quad (37)$$

2.3.2. Comparison to Cosine QMF Bank

Cosine QMF bank has been widely used in perceptual audio coding, such as MP3 [7] and low power SBR [1]. The construction of cosine QMF bank is similar to the above derivation for the complex QMF bank. In [4], there has been the detail description of the cosine QMF bank. This subsection indicates the main similarities and dissimilarities between cosine and complex QMF banks.

Under the prototype filter $P(z)$, the impulse response of the k^{th} modulated analysis and synthesis filters of

cosine QMF bank can viewed as the real parts of the ones from the complex QMF bank by taking advantage of the concealment of the image parts by combining the negative bands.

$$h_k(n) = p(n) \cos\left(\frac{\pi}{2M} \cdot (k+0.5)(2n-N) + \theta_k\right), \quad (38)$$

$$f_k(n) = p(n) \cos\left(\frac{\pi}{2M} \cdot (k+0.5)(2n-N) - \theta_k\right). \quad (39)$$

Similarly, to keep the same phase as the prototype filter and avoid the phase distortion, the constraints (19)(25) need to be satisfied. However, to eliminate the aliasing terms, there is an addition constraint (40) involving the constant a_k, b_k .

$$a_k b_k^* + a_{k-1} b_{k-1}^* = 0, \quad k = 1 \sim M-1. \quad (40)$$

Also, to make the amplitude distortion easy to control, the two constraints (41)(42) are imposed.

$$|a_k| = 1, \quad k = 0 \sim M-1. \quad (41)$$

$$a_0^4 = a_{M-1}^4 = -1. \quad (42)$$

For all the constraints (19),(25),(40)-(42), a solution is used commonly as

$$a_k = e^{j(-1)^k \frac{\pi}{4}}, \quad k = 0 \sim M-1. \quad (43)$$

The resultant impulse response of the k^{th} modulated analysis and synthesis filters are

$$h_k(n) = p(n) \cos\left(\frac{\pi}{2M} \cdot (k+0.5)(2n-N) + (-1)^k \frac{\pi}{4}\right), \quad (44)$$

$$f_k(n) = p(n) \cos\left(\frac{\pi}{2M} \cdot (k+0.5)(2n-N) - (-1)^k \frac{\pi}{4}\right). \quad (45)$$

Furthermore, there is another familiar solution that is used in MP3 and low power SBR as (46)

$$a_k = e^{-j\frac{\pi}{2}(k+0.5)}, \quad k = 0 \sim M-1, \quad (46)$$

and the resultant impulse response of the k^{th} modulated analysis and synthesis filters are

$$h_k(n) = p(n) \cos\left(\frac{\pi}{2M} \cdot (k+0.5)(2n-N-M)\right), \quad (47)$$

$$f_k(n) = p(n) \cos\left(\frac{\pi}{2M} \cdot (k+0.5)(2n-N+M)\right). \quad (48)$$

In other words, if the special constraint, such as (43),(46), is used for complex QMF bank, the image parts of the subband signal can be scarified and the original signal still can be synthesized by the synthesis filter bank of the cosine QMF bank, instead of complex QMF bank. On the other hand, besides the aliasing term concealment, the other advantage of complex QMF

bank is that all the involving filters are linear phase as Table 2.

	$P(z)$	$T(z)$	$H_k(z)$	$F_k(z)$
Cosine	$-\frac{N}{2}\varpi$	$-N\varpi$	Not linear phase	Not linear phase
Complex	$-\frac{N}{2}\varpi$	$-N\varpi$	$-\frac{N}{2}\varpi + \theta_k$	$-\frac{N}{2}\varpi - \theta_k$

Table 2: The phase functions of the involving filters in cosine and complex QMF banks

3. FAST COMPLEX QMF BANKS IN HE-AAC

This section presents two kinds of fast computing method based on DCT-IV and FFT for each matrix operation in (1)-(4).

3.1. Fast Algorithm based on DCT-IV

The close relation between DCT and DST will make the design of the fast algorithm easily, and result in a unified framework with DCT-IV only. Hence, before proceeding to the following subsections, the useful relation is reviewed.

Lemma 1 For an input signal $x[k]$ for $k = 0 \sim N-1$, the relation between its DST-IV coefficients and its reverse signal's DCT-IV coefficients is given as

$$DST_N^4(X[k])[n] = (-1)^n \cdot DCT_N^4(X^r[k])[n]. \quad (49)$$

<proof>

It is easy to prove the relation by direct derivation,

$$\begin{aligned} DST_N^4(X[k])[n] &= \sum_{k=0}^{N-1} X[k] S_{2N}^{(k+0.5)(n+0.5)} \\ &= \sum_{k=0}^{N-1} X[N-1-k] S_{2N}^{((N-1-k)+0.5)(n+0.5)} \\ &= \sum_{k=0}^{N-1} X^r[k] S_{2N}^{(N-(k+0.5))(n+0.5)} \\ &= (-1)^n \cdot \sum_{k=0}^{N-1} X^r[k] C_{2N}^{(k+0.5)(n+0.5)} \end{aligned} \quad (50)$$

3.1.1. Fast Analysis Matrix Operation in Encoder

A fast computing method by two M -point DCT-IV for (1) is present in the follow theorem.

Theorem 1 For $n = 0 \sim M-1$, let

$$\phi[n] = x[n] - x^r[n], \quad (51)$$

$$\theta[n] = x[n] + x^r[n]. \quad (52)$$

Then, for $k = 0 \sim M-1$

$$X[k] = DCT_M^4(\phi[n])[k] + i \cdot (-1)^k \cdot DCT_M^4(\theta^r[n])[k]. \quad (53)$$

<Proof>

From (1), it can be separated into two parts,

$$\begin{aligned} X[k] &= \sum_{n=0}^{M-1} x[n] W_{2M}^{-(k+0.5)(n+0.5)} \\ &+ \sum_{n=0}^{M-1} x^r[n] W_{2M}^{-(k+0.5)((2M-1-n)+0.5)} \\ &= \sum_{n=0}^{M-1} x[n] W_{2M}^{-(k+0.5)(n+0.5)} \\ &- \sum_{n=0}^{M-1} x^r[n] W_{2M}^{(k+0.5)(n+0.5)} \end{aligned} \quad (54)$$

By combining the real part and the image part of the two summations on RHS(right hand side) respectively, it gives

$$\begin{aligned} X[k] &= \sum_{n=0}^{M-1} (x[n] - x^r[n]) C_{2M}^{(k+0.5)(n+0.5)} \\ &+ i \cdot \sum_{n=0}^{M-1} (x[n] + x^r[n]) S_{2M}^{(k+0.5)(n+0.5)} \\ &= DCT_M^4(x[n] - x^r[n])[k] \\ &+ i \cdot DST_M^4(x[n] + x^r[n])[k] \end{aligned} \quad (55)$$

By Substituting (49),(51),(52) into (55), (53) is derived. ■

3.1.2. Fast Analysis Matrix Operation in Decoder

A fast computing method by one M -point DCT-IV for (2) is present in the follow theorem.

Theorem 2 For $n = 0 \sim M-1$, define $\phi[n]$ as following,

$$\phi[n] = \begin{cases} x[0], & \text{if } n = 0 \\ x[k+1], & \text{if } n = 2k, k = 1 \sim \frac{M}{2} - 1 \\ -x^r[k-1], & \text{if } n = 2k, k = 1 \sim \frac{M}{2} - 1 \\ x^r[\frac{M}{2} - 1], & \text{if } n = M - 1 \end{cases} \quad (56)$$

Also, let

$$\psi[k] = DCT_M^4(\phi[n])[k], \quad (57)$$

for $k = 0 \sim M-1$. Then,

$$X[k] = \psi[k] - i \cdot \psi^r[k], \quad (58)$$

for $k = 0 \sim \frac{M}{2} - 1$.

<Proof>

From the symmetry property as (59), for $k = 0 \sim M-1$,

$$\begin{aligned} X^r[k] &= \sum_{n=0}^{M-1} x[n] W_{2M}^{-(M-1-k+0.5)(2n-0.5)} \\ &= -i \cdot \sum_{n=0}^{M-1} x[n] W_{2M}^{(k+0.5)(2n-0.5)} \\ &= -i \cdot X^*[k] \end{aligned} \quad (59)$$

Then, the relation between the image and the real parts of $X[k]$ is given as (60), for $k = 0 \sim \frac{M}{2} - 1$,

$$I_X[k] = -R_X^r[k]. \quad (60)$$

Therefore, only the real part of $X[k]$ need to be considered,

$$R_X[k] = \sum_{n=0}^{M-1} x[n] C_{2M}^{(k+0.5)(2n-0.5)}. \quad (61)$$

It can be separated into two halves as (62),

$$\begin{aligned} R_X[k] &= \sum_{n=0}^{\frac{M}{2}-1} x[n] C_{2M}^{(k+0.5)(2n-0.5)} \\ &\quad + \sum_{n=\frac{M}{2}}^{M-1} x[n] C_{2M}^{(k+0.5)(2n-0.5)} \\ &\equiv S_1 + S_2 \end{aligned} \quad (62)$$

For the first summation on RSH in (62), if its 0th term is drew out, and it can be rewritten as

$$S_1 = x[0] C_{2M}^{(k+0.5)0.5} + \sum_{n=0}^{\frac{M}{2}-2} x[n+1] C_{2M}^{(k+0.5)(2n+1+0.5)}. \quad (63)$$

On the other hand, the second summation on RSH in (62) can be permuted and simplified as (64)

$$\begin{aligned} S_2 &= \sum_{n=0}^{\frac{M}{2}-1} x^r[n] C_{2M}^{(k+0.5)(2(M-1-n)-0.5)} \\ &= - \sum_{n=0}^{\frac{M}{2}-1} x^r[n] C_{2M}^{(k+0.5)(2(n+1)+0.5)} \end{aligned} \quad (64)$$

By drawing out the last term in (64), it is rewritten as

$$\begin{aligned} S_2 &= - \sum_{n=0}^{\frac{M}{2}-2} x^r[n] C_{2M}^{(k+0.5)(2(n+1)+0.5)} \\ &\quad - x^r[\frac{M}{2}-1] C_{2M}^{(k+0.5)(M+0.5)} \\ &= - \sum_{n=1}^{\frac{M}{2}-1} x^r[n-1] C_{2M}^{(k+0.5)(2n+0.5)} \\ &\quad + x^r[\frac{M}{2}-1] C_{2M}^{(k+0.5)(2M-(M+0.5))} \end{aligned} \quad (65)$$

Substituting (63), (65) into (62), it gives

$$\begin{aligned} R_X[k] &= x[0] C_{2M}^{(k+0.5)0.5} \\ &\quad + \sum_{n=0}^{\frac{M}{2}-2} x[n+1] C_{2M}^{(k+0.5)(2n+1+0.5)} \\ &\quad - \sum_{n=1}^{\frac{M}{2}-1} x^r[n-1] C_{2M}^{(k+0.5)(2n+0.5)} \\ &\quad + x^r[\frac{M}{2}-1] C_{2M}^{(k+0.5)(M-1+0.5)} \end{aligned} \quad (66)$$

(66) is exactly the permutation of the DCT-IV transform of $\phi[n]$ defined in (56). Furthermore, according to (60), (58) is also derived. ■

3.1.3. Synthesis Filter Bank in Decoder

A fast computing method by two M -point DCT-IV for (3) is present in the follow theorem.

Theorem 3 Let $\phi[n]$ and $\psi[n]$ be the DCT-IV transform coefficients of the real part signal of $X[k]$ and the image part signal of $X^r[k]$, respectively.

$$\phi[n] = DCT_M^4(R_X[k])[n], \quad (67)$$

$$\psi[n] = DCT_M^4(I_X^r[k])[n]. \quad (68)$$

Then,

$$\begin{cases} x[n] = (-1)^n \psi[n] - \phi[n] \\ x^r[n] = (-1)^n \psi[n] + \phi[n] \end{cases} \quad (69)$$

for $n = 0 \sim M-1$.

<Proof>

From (3), for $k = 0 \sim M-1$, it can be simplified as

$$\begin{aligned}
x[n] &= \operatorname{Re} \left\{ \sum_{k=0}^{M-1} X[k] W_{4M}^{-(k+0.5)(2n-4M+1)} \right\} \\
&= \operatorname{Re} \left\{ - \sum_{k=0}^{M-1} X[k] W_{2M}^{-(k+0.5)(n+0.5)} \right\} \\
&= - \left\{ \begin{aligned} &\sum_{k=0}^{M-1} R_X[k] C_{2M}^{(k+0.5)(n+0.5)} + \\ & - \sum_{k=0}^{M-1} I_X[k] S_{2M}^{(k+0.5)(n+0.5)} \end{aligned} \right\} \quad (70)
\end{aligned}$$

For $n = 0 \sim M-1$, the two summations in (70) are exactly the DCT-IV and the DST-IV, that is

$$x[n] = \operatorname{DST}_M^4(I_X[k])[n] - \operatorname{DCT}_M^4(R_X[k])[n]. \quad (71)$$

Similarly, consider the first half of the reverse signal $x^r[n]$ by (70), it gives

$$\begin{aligned}
x^r[n] &= - \left\{ \begin{aligned} &\sum_{k=0}^{M-1} R_X[k] C_{2M}^{(k+0.5)(2M-1-n+0.5)} \\ & - \sum_{k=0}^{M-1} I_X[k] S_{2M}^{(k+0.5)(2M-1-n+0.5)} \end{aligned} \right\} \\
&= \sum_{k=0}^{M-1} R_X[k] C_{2M}^{(k+0.5)(n+0.5)} + \sum_{k=0}^{M-1} I_X[k] S_{2M}^{(k+0.5)(n+0.5)} \quad (72)
\end{aligned}$$

Equivalently, for $n = 0 \sim M-1$

$$x^r[n] = \operatorname{DST}_M^4(I_X[k])[n] + \operatorname{DCT}_M^4(R_X[k])[n]. \quad (73)$$

Substituting (49) into (71) and (73) lead to (69). ■

3.1.4. Downsampled Synthesis Filter Bank in Decoder

A fast computing method by one M -point DCT-IV for (4) is present in the follow theorem.

Theorem 4 For $n = 0 \sim M-1$, let

$$\phi[k] = \begin{cases} -R_X[k], & \text{for } k = 0 \sim \frac{M}{2} - 1 \\ I_{X^r}[k - \frac{M}{2}], & \text{for } k = \frac{M}{2} \sim M - 1 \end{cases}, \quad (74)$$

and for $n = 0 \sim M-1$, let

$$\psi[n] = \operatorname{DCT}_M^4(\phi[k])[n]. \quad (75)$$

Then,

$$\begin{aligned}
x[n] &= \psi[2n] \\
x^r[n] &= \psi[2n+1] \quad (76)
\end{aligned}$$

for $n = 0 \sim \frac{M}{2} - 1$.

<Proof>

From (4), it can be simplified to (77),

$$\begin{aligned}
x[n] &= \operatorname{Re} \operatorname{al} \left\{ - \sum_{k=0}^{\frac{M}{2}-1} X[k] W_{2M}^{-(k+0.5)(2n+0.5)} \right\} \\
&= - \sum_{k=0}^{\frac{M}{2}-1} R_X[k] C_{2M}^{(k+0.5)(2n+0.5)} \\
&\quad + \sum_{k=0}^{\frac{M}{2}-1} I_X[k] S_{2M}^{(k+0.5)(2n+0.5)} \\
&= -S_1 + S_2 \quad (77)
\end{aligned}$$

Furthermore, the second summation in (77) can be rewritten with cosine coefficients,

$$\begin{aligned}
S_2 &= \sum_{k=\frac{M}{2}}^{M-1} I_X[M-1-k] S_{2M}^{(M-1-k+0.5)(2n+0.5)} \\
&= \sum_{k=\frac{M}{2}}^{M-1} I_X[M-1-k] C_{2M}^{(k+0.5)(2n+0.5)} \quad (78)
\end{aligned}$$

Substituting (78) into (77) leads to

$$\begin{aligned}
x[n] &= \sum_{k=0}^{\frac{M}{2}-1} (-R_X[k]) C_{2M}^{(k+0.5)(2n+0.5)} \\
&\quad + \sum_{k=\frac{M}{2}}^{M-1} I_{X^r}[k - \frac{M}{2}] C_{2M}^{(k+0.5)(2n+0.5)} \quad (79)
\end{aligned}$$

According to the definition (74), it is equivalent to

$$x[n] = \operatorname{DCT}_M^4(\phi[k])[2n]. \quad (80)$$

for $n = 0 \sim \frac{M}{2} - 1$. On the other hand, for the another half of the x signal, it can be derived as follow,

$$\begin{aligned}
x^r[n] &= x[M-1-n] \\
&= \sum_{k=0}^{\frac{M}{2}-1} (-R_X[k]) C_{2M}^{(k+0.5)(2(M-1-n)+0.5)} \\
&\quad + \sum_{k=\frac{M}{2}}^{M-1} I_{X^r}[k - \frac{M}{2}] C_{2M}^{(k+0.5)(2(M-1-n)+0.5)} \\
&= \sum_{k=0}^{\frac{M}{2}-1} R_X[k] C_{2M}^{(k+0.5)(2n+1+0.5)} \\
&\quad - \sum_{k=\frac{M}{2}}^{M-1} I_{X^r}[k - \frac{M}{2}] C_{2M}^{(k+0.5)(2n+1+0.5)} \quad (81)
\end{aligned}$$

Also, it is exactly the negative DCT coefficient of ϕ at odd index,

$$x^r[n] = -DCT_M^4(\phi[k])[2n+1]. \quad (82)$$

3.2. Fast Algorithm based on FFT

To maintain the unified framework for the codec systems adapting FFT as the kernel transform tool, the subsection presents the fast algorithms based on FFT.

The relation of IDFT and DFT is useful to the algorithms. Before proceeding to the following subsections, the relation is also reviewed.

Lemma 2 For the input signal $x[k]$ for $k = 0 \sim N-1$, the relation of its IDFT coefficients and DFT coefficients of its reverse signal is given as

$$IDFT(x[n])[k] = W_N^k \cdot DFT(x^r[n])[k]. \quad (83)$$

<Proof>

$$\begin{aligned} \sum_{n=0}^{N-1} x[n] W_N^{-kn} &= \sum_{n=0}^{N-1} x^r[n] W_N^{-k(N-1-n)} \\ &= W_N^k \sum_{n=0}^{N-1} x^r[n] W_N^{kn} \end{aligned} \quad (84)$$

3.2.1. Fast Analysis Matrix Operation in Encoder and Decoder

For (1)(2), it can be viewed identically as

$$X[k] = \sum_{n=0}^{\psi-1} x[n] W_{2\psi}^{-(k+0.5)(2n+\alpha)}, \quad (85)$$

for $k = 0 \sim \frac{\psi}{2} - 1$. The related notations α and ψ are “1, 128” and “0.5, 64” for the two cases. Also, for convenience, we extended the range of the index k to $\psi - 1$ for (85), and thus $X^r[k]$ means $X[\psi - 1 - k]$.

At first, the symmetry property of (85) is shown as the following theorem.

Theorem 5 For the output signal of (85), there is a skew symmetric-conjugation relation.

$$X^r[k] = -X^*[k]. \quad (86)$$

for $k = 0 \sim \psi - 1$.

<proof>

It is easy to prove the relation by direct derivation,

$$\begin{aligned} X^r[k] &= \sum_{n=0}^{\psi-1} x[n] W_{2\psi}^{-(\psi-1-k+0.5)(2n+\alpha)} \\ &= \sum_{n=0}^{\psi-1} x[n] W_{2\psi}^{(-\psi+k+0.5)(2n+\alpha)} \\ &= -\sum_{n=0}^{\psi-1} x[n] W_{2\psi}^{(k+0.5)(2n+\alpha)} \\ &= -\left(\sum_{n=0}^{\psi-1} x[n] W_{2\psi}^{-(k+0.5)(2n+\alpha)} \right)^* = -X^*[k] \end{aligned} \quad (87)$$

Furthermore, there is a fast method by FFT to compute the output sample of the even index.

Theorem 6 For $n = 0 \sim \frac{\psi}{2} - 1$, define

$$\phi[n] = x^r\left[n + \frac{\psi}{2}\right] + i \cdot x^r[n]. \quad (88)$$

Moreover, for $k = 0 \sim \psi - 1$, let

$$Y[k] = W_{\psi}^{-k\alpha+2k-\alpha} \cdot DFT_{\frac{\psi}{2}}\left(\phi[n] W_{\frac{\psi}{2}}^{n+1}\right). \quad (89)$$

Then, for $k = 0 \sim \frac{\psi}{2} - 1$,

$$X[2k] = Y[k]. \quad (90)$$

<proof>

For the output sample of the even index from (85), it is given as

$$\begin{aligned} X[2k] &= \sum_{n=0}^{\psi-1} x[n] W_{2\psi}^{-(2k+0.5)(2n+\alpha)} \\ &= \sum_{n=0}^{\frac{\psi}{2}-1} x[n] W_{2\psi}^{-(2k+0.5)(2n+\alpha)} \\ &\quad + \sum_{n=0}^{\frac{\psi}{2}-1} x\left[n + \frac{\psi}{2}\right] W_{2\psi}^{-(2k+0.5)(2(n+\frac{\psi}{2})+\alpha)} \\ &\equiv S_1 + S_2 \end{aligned} \quad (91)$$

The second summation in (91) can be simplified as

$$S_2 = i \cdot \sum_{n=0}^{\frac{\psi}{2}-1} x\left[n + \frac{\psi}{2}\right] W_{2\psi}^{-(2k+0.5)(2n+\alpha)}. \quad (92)$$

Substituting (92) into (91) gives

$$\begin{aligned} X[2k] &= \sum_{n=0}^{\frac{\psi}{2}-1} (x[n] + i \cdot x\left[n + \frac{\psi}{2}\right]) W_{2\psi}^{-(2k+0.5)(2n+\alpha)} \\ &= W_{\psi}^{-k\alpha} \sum_{n=0}^{\frac{\psi}{2}-1} \left\{ x[n] + i \cdot x\left[n + \frac{\psi}{2}\right] \right\} W_{\psi}^{-(2n+\alpha)} W_{\frac{\psi}{2}}^{-nk} \end{aligned} \quad (93)$$

Changing the IDFT involved in (93) to DFT by the relation (83) gives

$$X[2k] = W_{\psi}^{-k\alpha} W_{\frac{\psi}{2}}^k \cdot \sum_{n=0}^{\frac{\psi}{2}-1} \left\{ x[\frac{\psi}{2}-1-n] + i \cdot x[\psi-1-n] \right\} W_{\psi}^{-(2(\frac{\psi}{2}-1-n)+\alpha)} W_{\frac{\psi}{2}}^{nk} \quad (94)$$

(94) can be further simplified as

$$X[2k] = W_{\psi}^{-k\alpha+2k-\alpha} \cdot \sum_{n=0}^{\frac{\psi}{2}-1} \left\{ \phi[n] W_{\frac{\psi}{2}}^{n+1} \right\} W_{\frac{\psi}{2}}^{nk}, \quad (95)$$

The RHS of (95) is the same as the RHS of (89). Hence, (90) holds.

Based on the symmetric property in (86), for the output sample of the odd index from (85), it can be given by the related even-indexed sample as (96)

$$X[2k+1] = -X^{*r}[2k+1] = -X^{*}[\psi-2k]. \quad (96)$$

In other words, once $Y[k]$ is computed by (89) and ψ is even number, then

$$\begin{aligned} X[2k] &= Y[k] \\ X[2k+1] &= -Y^{*}[\frac{\psi}{2}-k], \end{aligned} \quad (97)$$

for $k = 0 \sim \frac{\psi}{2}-1$.

3.2.2. Synthesis and Downsampled Synthesis Filter Banks in Decoder

For(3)(4), it can be viewed identically as

$$x[n] = \text{Re} \left\{ \sum_{k=0}^{\frac{\psi}{2}-1} X[k] W_{2\psi}^{-(k+0.5)(2n+\alpha)} \right\}, \quad (98)$$

for $n = 0 \sim \psi-1$. The related notations α and ψ are “-255, 128” and “-127.5, 64” for the two cases, respectively.

A fast computing method by one $\frac{\psi}{2}$ -point FFT for both (3)(4) is present in the follow theorem.

Theorem 7 Let the two intermediums be

$$\Phi[k] = X[k] W_{2\psi}^{-(k+0.5)\alpha}, \quad (99)$$

$$\Psi[k] = X[k] W_{2\psi}^{-(k+0.5)(\alpha-2)}, \quad (100)$$

for $k = 0 \sim \frac{\psi}{2}-1$. Also, let

$$\Theta[k] = \Phi^{*}[k] + \Phi^r[k] - i(\Psi[k] + \Psi^{*r}[k]). \quad (101)$$

Furthermore, let

$$Y[n] = \frac{1}{2} \text{DFT}_{\frac{\psi}{2}}(\Theta[k])[n], \quad (102)$$

for $n = 0 \sim \psi-1$. Then,

$$x[2p] = R_Y[n], \quad (103)$$

$$x^r[2p] = I_Y[n], \quad (104)$$

for $n = 0 \sim \frac{\psi}{2}-1$.

<Proof>

Consider the even-indexed output sample, (98) can be simplified as

$$\begin{aligned} x[2p] &= \text{Re} \left\{ W_{\psi}^{-p} \sum_{k=0}^{\frac{\psi}{2}-1} X[k] W_{2\psi}^{-(k+0.5)\alpha} W_{\frac{\psi}{2}}^{-kp} \right\} \\ &= \text{Re} \left\{ W_{\psi}^{-p} \sum_{k=0}^{\frac{\psi}{2}-1} \Phi[k] W_{\frac{\psi}{2}}^{-kp} \right\} \end{aligned} \quad (105)$$

Also, consider the even-indexed reverse output sample,

$$\begin{aligned} x^r[2p] &= \text{Re} \left\{ \sum_{k=0}^{\frac{\psi}{2}-1} X[k] W_{2\psi}^{-(k+0.5)(2(\psi-1-2p)+\alpha)} \right\} \\ &= \text{Re} \left\{ - \sum_{k=0}^{\frac{\psi}{2}-1} X[k] W_{2\psi}^{(k+0.5)(4p+2-\alpha)} \right\} \end{aligned} \quad (106)$$

The exponential function in (106) can be separated as

$$\begin{aligned} x^r[2p] &= -\text{Re} \left\{ W_{\psi}^p \sum_{k=0}^{\frac{\psi}{2}-1} X[k] W_{2\psi}^{(k+0.5)(2-\alpha)} W_{\frac{\psi}{2}}^{kp} \right\} \\ &= \text{Re} \left\{ - W_{\psi}^p \sum_{k=0}^{\frac{\psi}{2}-1} \Psi[k] W_{\frac{\psi}{2}}^{kp} \right\} \end{aligned} \quad (107)$$

Now, combine $x[2p]$ and $x^r[2p]$ as a complex number as

$$\begin{aligned} x[2p] + ix^r[2p] &= \frac{1}{2} \left\{ \begin{aligned} &W_{\psi}^{-p} \sum_{k=0}^{\frac{\psi}{2}-1} \Phi[k] W_{\frac{\psi}{2}}^{-kp} \\ &+ W_{\psi}^p \sum_{k=0}^{\frac{\psi}{2}-1} \Phi^{*}[k] W_{\frac{\psi}{2}}^{kp} \end{aligned} \right\}, \\ &-i \cdot \frac{1}{2} \left\{ \begin{aligned} &W_{\psi}^p \sum_{k=0}^{\frac{\psi}{2}-1} \Psi[k] W_{\frac{\psi}{2}}^{kp} \\ &+ W_{\psi}^{-p} \sum_{k=0}^{\frac{\psi}{2}-1} \Psi^{*}[k] W_{\frac{\psi}{2}}^{-kp} \end{aligned} \right\} \end{aligned} \quad (108)$$

Combine the terms in (108), it gives

$$\begin{aligned}
 2(x[2p] + ix^r[2p]) &= W_\psi^{-p} \sum_{k=0}^{\frac{\psi}{2}-1} (\Phi[k] - i\Psi^*[k])W_{\frac{\psi}{2}}^{-kp} \\
 &\quad + W_\psi^p \sum_{k=0}^{\frac{\psi}{2}-1} (\Phi^*[k] - i\Psi[k])W_{\frac{\psi}{2}}^{kp} \\
 &\equiv S_1 + S_2
 \end{aligned} \tag{109}$$

From (83), the first summation involving a IDFT can be derived as

$$\begin{aligned}
 S_1 &= W_\psi^{-p} W_{\frac{\psi}{2}}^p \sum_{k=0}^{\frac{\psi}{2}-1} (\Phi^r[k] - i\Psi^{*r}[k])W_{\frac{\psi}{2}}^{kp} \\
 &= W_\psi^p \sum_{k=0}^{\frac{\psi}{2}-1} (\Phi^r[k] - i\Psi^{*r}[k])W_{\frac{\psi}{2}}^{kp}
 \end{aligned} \tag{110}$$

Substituting (110) back to (109) gives

$$x[2p] + ix^r[2p] = \frac{1}{2} \cdot W_\psi^p \sum_{k=0}^{\frac{\psi}{2}-1} \Theta[k]W_{\frac{\psi}{2}}^{kp} \tag{111}$$

■

4. CONCLUSION

This paper has explained the inherent capacity of aliasing suffering of complex QMF bank used in MPEG-4 HE-AAC. Furthermore, this paper has present the two different kind of fast algorithms based on FFT and DCT-IV for the kernel matrix operations in the complex QMF bank. Through the available fast methods for DCT and FFT, the complexity can be reduced from $O(N^2)$ to $O(N \cdot \log N)$.

5. ACKNOWLEDGEMENTS

This work was supported by National Science Council under contract NSC94-2213-E-009-128.

6. REFERENCES

[1] ISO/IEC, “Text of ISO/IEC 14496-3:2001/FDAM1, Bandwidth Extension,” ISO/IEC JTC1/SC29/WG11/N5570, March 2003, Pattaya, Thailand.

[2] M. Dietz, L. Liljeryd, K. Kjörling and O. Kunz, “Spectral Band Replication, A Novel Approach in Audio Coding,” presented at the AES112nd Convention, Munich, 2002 May 10–13.

[3] M. Wolters, K. Kjörling, D. Himm and H. Purnhagen, “A Closer Look into MPEG-4 High Efficiency AAC,” presented at the AES115th Convention, New York, USA, 2003 October 10–13.

[4] P. P. Vaidyanthan, “Multirate digital filters,” Prentice Hall Inc., 1993.

[5] Y. Wang and M. Vilermo, “Modified Discrete Cosine Transform— Its Implications for Audio Coding and Error Concealment,” J. Audio Eng. Soc., vol. 51, No. 1/2, 2003 Jan./Feb.

[6] FAAD, website <http://www.audiocodind.com>

[7] ISO/IEC JTC1/SC29, “Information technology-coding of moving pictures and associated audio for digital storage media at up to 1.5 mps-CD11172(part3, audio),” Dos. ISO/IEC JTC1/SC29 NO71.