Design of Uniform DFT Filter Banks Optimized for Subband Coding of Speech

AHARON SATT, STUDENT MEMBER, IEEE, AND DAVID MALAH, FELLOW, IEEE

Abstract—**A new approach for designing uniform DFT analysis/synthesis filter banks, optimized for subband coding (SBC) of speech, is presented. A spectral-domain distortion measure, which consists of a weighted sum of error terms due to filtering, rate conversions, and quantization, is derived. The quantization is embedded by means of a statistical model. For the case of "fine" quantization, a simpler, deterministic, distortion function is derived. The design of the optimal filters is performed by an iterative algorithm, in which two sets of linear equations are solved in each iteration, aiming at minimizing the distortion function. A 16 kbit/s SBC is simulated, using a filter bank designed by the new approach, and is found to achieve similar subjective and objective (SNR) performance as that of a conventional QMF-based SBC, with only about half the computations.**

I. Introduction

SUBBAND coding is a well-known method for digital speech coding at medium rates (e.g., 16 kbit/s). In a subband coder (SBC), the signal is divided into separate bands (typically eight) by using an analysis filter bank. Usually, each band signal is quantized by a gain-adaptive scalar quantizer. The speech is reconstructed from the quantized band signals using a synthesis filter bank. A coder of this type is considered to be of "medium complexity," with its most complex part being the filter bank [1].

Two common types of filter banks are the Quadrature Mirror Filter (QMF) bank [2] and the uniform DFT filter bank [2]. The QMF bank is designed to completely cancel the aliasing due to the decimation of the band signals (in the absence of quantization), and it is widely used in subband speech coders. The QMF-based SBC obtains good quality at medium bit rates. Its drawback, however, is its relatively high implementation complexity.

On the other hand, the DFT bank can be implemented efficiently using FIR filters and the Weighted Overlap Add (WOLA) scheme [2], in which case it is of much lower complexity than the QMF bank, for similar band separation. However, because known design techniques aim at minimizing the overall response error of the filter bank, using either deterministic [12] or statistical [3] error measure, the performance of the DFT-based SBC (in terms of subjective quality) was found to be much lower.

In this work we present a new approach for designing filters for uniform DFT filter banks. A spectral-domain distortion function, which consists of a weighted sum of Euclidean norms of the various error terms in the output signal, is derived. The optimal filters are obtained by assigning an appropriate weight to each term, and minimizing the distortion function. The error terms are as follows: a) aliasing distortion and the error relative to unity transmission, b) output quantization noise, c) sidelobe (stop-band) energy of the analysis filter. A more detailed discussion of these error terms follows.

a) The frequency bands are required to be well separated (as explained in the sequel). Thus, the length (L) of the analysis (FIR) filter impulse-response must be longer than the number of bands (M). If $L > M$ and the subband signals are decimated by the critical ratio ($R = M$), the uniform DFT filter bank with FIR analysis/synthesis filters cannot realize an exact unity system (i.e., the overall analysis/synthesis transfer function—in the absence of quantization—is not a pure delay) [2]. The aliasing may be cancelled in the output signal, by using synthesis filters which have long duration [11], but this leads to an inefficient implementation. If a lower decimation ratio is applied ($R < M$), a unity system is achievable. However, the filter response is not simply controlled (in particular, the stop-band attenuation). Hence, in practice, aliasing cancellation and unity transmission are only approximated.

b) The quantization error is modeled as additive noise. Two terms of the noise in the output signal are considered: the term ($n1$) of noise components which are contained in the frequency bands at which they were produced; the term ($n2$) of noise leakage from one frequency band to other bands due to nonzero frequency response of the synthesis filter in the stop-band. The quantization noise in each frequency band is "masked" by a stronger band signal in the same band, i.e., reducing the loudness at which the noise is perceived (the auditory masking effect [3]). It is important to minimize the leakage of quantization noise from one frequency band to other bands, including adjacent bands if their bandwidth is not narrow [3], since its masking in other bands cannot be controlled as it depends on signal intensity in those bands.

c) It is required to have good band separation in the analysis stage, since redundancy removal is enhanced by

Manuscript received July 20, 1988; revised January 16, 1989.

A. Satt is with the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 32000, Israel.

D. Malah is with the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 32000, Israel, on sabbatical leave at the Department of Signal Processing Research, AT&T Bell Laboratories, Murray Hill, NJ 07974.

IEEE Log Number 8930538.

providing uncorrelated band signals. This allows the design of quantizers which are better matched to the nonstationary properties of the band signals (in particular, adaptation to the gain in each band). The band separation is controlled by introducing a stop-band energy term (to be defined later) into the distortion function.

The paper is organized as follows. Section II presents the distortion function to be minimized, and a discussion on the quantization noise effect on the optimal filters; Section III presents the iterative algorithm used; Section IV presents a design example; Section V presents the implementation of a 16 kbit/s SBC using the optimized filter banks; and Section VI presents simulation results with the new SBC system and, for comparison, also with a conventional QMF-based SBC.

II. Distortion Function

In this section, first, a statistical distortion criterion is developed, and then, for the case of "fine" quantization, a deterministic distortion function is derived.

The basic structure of a (complex) uniform DFT filter bank, with decimation and interpolation, is shown in Fig. 1. In the analysis stage, the input signal $x(n)$ is demodulated by $\exp(-j\omega_k n)$, filtered by the FIR analysis low-pass filter $h(n)$, and R:1 decimated to produce the M band signals $X_k(m)$, $k = 0, 1, \cdots, M-1$. Ideally, these M complex signals have a bandwidth of $2\pi/M$. In the synthesis stage, the band signals are interpolated by the synthesis low-pass filters $f(n)$, modulated by $\exp(j\omega_k n)$, and summed up to produce the output signal $\hat{x}(n)$.

Using well-known z-transform relations for decimated and interpolated signals, the z-transform of the output signal is given by

$$\hat{X}(z) = \frac{1}{R}\sum_{k=0}^{R-1} X(zW_R^k)\,\frac{1}{M}\sum_{l=0}^{M-1} F(zW_M^{-l})\,H(zW_R^kW_M^{-l}) + \{\text{terms which are dependent on the quantization noise}\}$$

$$W_M \triangleq e^{j2\pi/M}, \qquad W_R \triangleq e^{j2\pi/R}. \tag{1}$$

The terms which include $X(zW_R^k)$, for $k \neq 0$, are the aliasing components in the output signal. Hence, we define the aliasing distortion as

$$E_{al} \triangleq \frac{1}{M^2}\sum_{k=1}^{R-1}\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{1}{R}\sum_{l=0}^{M-1} F(e^{j\omega}W_M^{-l}) \cdot H(e^{j\omega}W_R^kW_M^{-l})\right|^2 d\omega. \tag{2}$$

The mean squared error (MSE) relative to unity transmission, i.e., pure delay of n_0 samples, is given by

$$E_{tr} \triangleq \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|e^{-j\omega n_0} - \frac{1}{MR}\sum_{l=0}^{M-1} F(e^{j\omega}W_M^{-l}) \cdot H(e^{j\omega}W_M^{-l})\right|^2 d\omega. \tag{3}$$

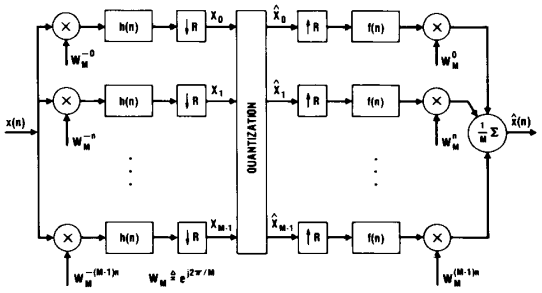


Fig. 1. Basic structure of the uniform DFT filter bank.

The quantization process is modeled by an additive wide-sense stationary noise, uncorrelated with the speech signal, and assumed to have zero mean. This model is adequate for medium and high bit-rate SBC (e.g., 16 kbit/s and above), where most of the frequency bands are allocated at least 2–3 bits. Because of the interpolation process, the output quantization noise is not wide-sense stationary. However, its power spectrum is periodic in time with a period R (the interpolation ratio). An average power spectrum is defined by averaging the (time-dependent) power spectrum over one period. An expression for the average spectrum is derived in Appendix A, and is given by

$$S_{\text{out}}(e^{j\omega}) = \frac{1}{M^2}\sum_{l=0}^{M-1} S_{vl}(e^{j\omega}W_M^{-lR})\left|F(e^{j\omega}W_M^{-l})\right|^2 \tag{4}$$

where $S_{vl}(e^{j\omega})$ is the power spectral density of the noise source in the lth (complex) band. The power of the quantization noise components contained in their original frequencies, normalized by the input signal power σ_x^2, is

$$E_{n1} \triangleq \frac{1}{\sigma_x^2}\frac{1}{M^2}\sum_{l=0}^{M-1}\frac{1}{2\pi}\int_{2\pi l/M-\pi/M}^{2\pi l/M+\pi/M} \cdot S_{vl}(e^{j\omega}W_M^{-lR})\left|F(e^{j\omega}W_M^{-l})\right|^2 d\omega \tag{5}$$

and the normalized power of the leakage noise components is

$$E_{n2} \triangleq \frac{1}{\sigma_x^2}\frac{1}{M^2}\sum_{l=0}^{M-1}\frac{1}{2\pi}\int_{2\pi l/M-\pi/m}^{2\pi l/M+\pi/M}\sum_{\substack{k=0\\k\neq l}}^{M-1} \cdot S_{vk}(e^{j\omega R}W_M^{-kR})\left|F(e^{j\omega}W_M^{-k})\right|^2 d\omega. \tag{6}$$

Both output noise power terms, E_{n1} and E_{n2}, are normalized by the input signal power rather than the output signal power, to simplify the design equations, assuming approximately unity transmission.

The ideal frequency response of the analysis filter is

$$H_{\text{ideal}}(e^{j\omega}) = \begin{cases} M, & 0 < |\omega| < \pi/M \\ 0, & \pi/M < |\omega| < \pi \end{cases} \tag{7}$$

where the gain factor M compensates for the division of the output signal by M (see Fig. 1). The sidelobe (stop-

band) energy is defined as

$$E_h \triangleq \frac{1}{2\pi}\int_{-\pi}^{\pi} U_h(e^{j\omega})\left|H(e^{j\omega})\right|^2 d\omega$$

$$U_h(e^{j\omega}) \triangleq \begin{cases} 0, & 0 < |\omega| < \pi/M \\ 1, & \pi/M < |\omega| < \pi. \end{cases} \quad (8)$$

The first distortion function we introduce is given by the following weighted sum:

$$\hat{D} = w_{al}E_{al} + E_{tr} + w_{n1}E_{n1} + w_{n2}E_{n2} + w_hE_h \quad (9)$$

where the weight factors w_{al}, w_h, w_{h1}, w_{h2} are nonnegative constants. $\hat{D}$ may be used for designing the optimal filters, in which case it is required to estimate the autocorrelation sequence of the noise signal in each band. In the particular case of equal weights: $w_{al} = w_h = w_{n1} = w_{n2} = 1$, the distortion function $\hat{D}$ is similar to a particular case of the time-domain error measure in [4, eqns. 30, B3] (when the input signal is white and the weight function in [4, eq. B3] is constant).

Next, we concentrate on the case of "fine" quantization, and derive a simpler, deterministic measure. Consider the case in which the number of quantization levels allocated to each nonzero-energy band is not too low (e.g., eight levels and above), hence, the quantization noise characteristics are similar to white noise. Let η_l^2 (a constant) denote the power spectrum density of the quantization noise in the lth (complex) band. Expression (5) is therefore reduced to

$$E_{n1} = \frac{1}{\sigma_x^2}\frac{1}{M^2}\frac{1}{2\pi}\int_{-\pi/M}^{\pi/M}\left|F(e^{j\omega})\right|^2 d\omega \cdot \sum_{l=0}^{M-1}\eta_l^2. \quad (10)$$

Note that the output signal-to-noise ratio with respect to the noise term $n1$ is not affected by the synthesis filter, since by attenuating the frequency response $F(e^{j\omega})$ both the output signal and the output noise levels are reduced. Hence, the term E_{n1} can be omitted from (9).

If optimal bit allocation is applied, the noise power is the same in each band, i.e., $\eta_l^2 = \eta^2$, $l = 0, \cdots, M-1$ [1]. E_{n2} in (6) is then reduced to

$$E_{n2} = \frac{1}{\sigma_x^2}\frac{1}{M}\frac{1}{2\pi}\int_{\omega\notin[-(\pi/M),\pi/M]}\left|F(e^{j\omega})\right|^2 d\omega \cdot \frac{\eta^2}{\sigma_x^2}. \quad (11)$$

Thus, E_{n2} is proportional to the sidelobe energy of the synthesis filter. Hence, we can replace E_{n2} in (9) by a sidelobe energy term, defined below.

To define the sidelobe energy, we consider the following ideal frequency response of the synthesis filter, as shown in Fig. 2:

$$F_{\text{ideal}}(e^{j\omega}) = \begin{cases} R, & 0 < |\omega| < \pi/M \\ 0, & 2\pi k/R - \pi/M < |\omega| < 2\pi k/R + \pi/M \\ \phi, & 2\pi(k-1)/R + \pi/M < |\omega| < 2\pi k/R - \pi/M \end{cases}$$
$$k = 1, 2, \cdots, R/2. \quad (12)$$

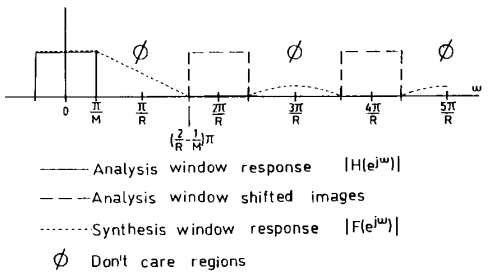


Fig. 2. Frequency responses of the ideal filters.

F_{ideal} is an ideal interpolation filter (cf. [2]), and in the absence of quantization, the regions denoted by ϕ are "don't care" regions. In an SBC, the ϕ regions corresponding to $k = 2, \cdots, R/2$ must be zero (ideally), in order to prevent the leakage of quantization noise. Thus, the sidelobe energy of the synthesis filter is defined here by

$$E_f \triangleq \frac{1}{2\pi}\int_{-\pi}^{\pi} U_f(e^{j\omega})\left|F(e^{j\omega})\right|^2 d\omega$$

$$U_f(e^{j\omega}) \triangleq \begin{cases} 0, & 0 < |\omega| < 2\pi/R - \pi/M \\ 1, & 2\pi/R - \pi/M < |\omega| < \pi. \end{cases} \quad (13)$$

Defining E_f as in (13), rather than extending the stop-band definition as in (8), results in a lower distortion filter bank (when $R < M$). The distortion function (9) is now replaced by

$$D = w_{al}E_{al} + E_{tr} + w_hE_h + w_fE_f \quad (14)$$

where w_f is a nonnegative constant.

III. Iterative Algorithm for Optimal Filter Bank Design

Using Parseval's equality D can be represented as a positive-semidefinite (PSD) quadratic form, in terms of the analysis and synthesis filter coefficients:

$$\begin{aligned} D &= 1 - 2\boldsymbol{f}^T\boldsymbol{q} + \boldsymbol{f}^T\boldsymbol{Q}\boldsymbol{f} + \text{terms which are independent of } \boldsymbol{f} \\ &= 1 - 2\boldsymbol{h}^T\tilde{\boldsymbol{q}} + \boldsymbol{h}^T\tilde{\boldsymbol{Q}}\boldsymbol{h} + \text{terms which are independent of } \boldsymbol{h} \end{aligned} \quad (15)$$

where $\boldsymbol{h}$ and $\boldsymbol{f}$ are vectors with elements obtained from consecutive terms of $h(n)$ and $f(n)$, respectively. The derivation of (15) and the expressions for the PSD matrices $\boldsymbol{Q}$, $\tilde{\boldsymbol{Q}}$ and the vectors $\boldsymbol{q}$, $\tilde{\boldsymbol{q}}$ are given in Appendix B.

When the analysis filter $h(n)$ is given, the synthesis filter $f(n)$ is computed by minimizing D, and vice versa. Minimization of the corresponding PSD quadratic forms in (15) with respect to $f(n)$ or $h(n)$ gives the following linear equations for solving for the optimal filters, respectively:

$$\boldsymbol{Q}\boldsymbol{f}_{\text{opt}} = \boldsymbol{q} \qquad \tilde{\boldsymbol{Q}}\boldsymbol{h}_{\text{opt}} = \tilde{\boldsymbol{q}}. \quad (16)$$

Assuming $w_h > 0$ and $w_f > 0$ [since we wish to limit the frequency response in the stop-band, both for $h(n)$ and

$f(n)$], and following the derivation in Appendix B, we conclude that $\boldsymbol{Q}$ and $\tilde{\boldsymbol{Q}}$ are PD matrices, yielding therefore unique solutions to (16).

If the given FIR filter is symmetric (i.e., linear phase), the corresponding computed filter from (16) is also symmetric. In this case, as shown in Appendix B, the dimension of the equation set can be reduced from $L \times L$ to $(L/2) \times (L/2)$, where L is the length of the computed filter.

The combined design of a pair of optimal analysis and synthesis filters is performed by an iterative algorithm, similar to [4], which converges to a local minimum of D, as follows.

a) Initialization: Let $\boldsymbol{f}^{(0)}$ be an initial synthesis filter, $\epsilon > 0$ a threshold constant, $k = 1$, $D^{(0)} = \infty$.
b) Given $\boldsymbol{f}^{(k-1)}$, compute $\boldsymbol{h}^{(k)}$ from (16): $\tilde{\boldsymbol{Q}}\boldsymbol{h}^{(k)} = \tilde{\boldsymbol{q}}$.
c) Given $\boldsymbol{h}^{(k)}$, compute $\boldsymbol{f}^{(k)}$ from (16): $\boldsymbol{Q}\boldsymbol{f}^{(k)} = \boldsymbol{q}$.
d) Compute the distortion D using (15). If $(D^{(k-1)} - D^{(k)})/D^{(k)} < \epsilon$, then go to step e. Otherwise, $k \leftarrow k + 1$, and return to step b.
e) Normalize $\boldsymbol{h}^{(k)}$ and $\boldsymbol{f}^{(k)}$ (as explained in the sequel). Stop.

By increasing the weights w_{al}, w_h, w_f in the distortion function defined in (14), the amplitude of the windows obtained by minimizing D decreases. In order to restore the amplitude of the synthesized signal, without affecting the output SNR, $f(n)$ is scaled by a factor c_0, which minimizes the MSE relative to unity transmission:

$$E_{tr}(c) \triangleq \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|e^{-j\omega n_0} - \frac{c}{MR}\sum_{l=0}^{M-1} F(e^{j\omega}W_M^{-l}) \cdot H(e^{j\omega}W_M^{-l})\right|^2 d\omega$$

$$c_0 = \arg\left\{\min_c E_{tr}(c)\right\}; \qquad \boldsymbol{f}^{(k)} \leftarrow c_0\boldsymbol{f}^{(k)}. \quad (17)$$

An expression for the factor c_0, in terms of the filter coefficients, is given in (B13).

The iterative algorithm consists of computing the analysis and synthesis filters alternatively, minimizing the same (positive) distortion function. Hence, D is monotonically decreasing from iteration to iteration, and the algorithm converges to a local minimum of D, which depends on the initial filter chosen.

IV. Design Example

In this section, we present a filter design example for a subband coder with $M = 16$ complex-bands.

To achieve maximum coding efficiency, it is desirable to decimate the band signals by the critical ratio $R = M$. However, the designed filters for the critical ratio were not found to achieve both a low distortion filter bank (when no coding is applied) and sufficiently high stop-band attenuation, unless they are of very long duration. Decreasing slightly the decimation ratio, using $R = (M - 1) = 15$, enables the design of filters which results in high performance SBC, but still of quite lower complexity than a conventional QMF coder of similar performance.

The impulse and frequency responses of the optimal filters, with $M = 16$ bands, decimation ratio $R = 15$, and length of 256 taps each, are shown in Fig. 3. The weight factors were chosen to be $w_{al} = 10$, $w_h = 10$, $w_f = 20$, and the resulting filters provide good band separation and a high performance SBC. The "transfer-function" of the system in the above example (obtained from (1) by excluding the aliasing components) is given by

$$T(\omega) \triangleq \frac{1}{MR}\sum_{l=0}^{M-1} H(e^{j\omega}W_M^{-l})\,F(e^{j\omega}W_M^{-l}) \quad (18)$$

and is plotted in Fig. 4.

The initial filter $\boldsymbol{f}^{(0)}$ was obtained in the following way. First, a symmetric low-pass filter with cutoff frequency π/M is designed via any conventional method. This low-pass filter is then used as an initial analysis filter for designing a synthesis filter having high attenuation in the stop-band by using the iterative algorithm with $w_{al} = 1$, $w_h = w_f = 1000$. The resulting synthesis filter is then used as an initial filter for the final design process. The design process can, of course, begin with an initial analysis filter $\boldsymbol{h}^{(0)}$, using a similar procedure to find it.

V. SBC Implementation

The proposed SBC consists of the above analysis and synthesis filter banks and adaptive scalar quantizers with dynamic bit allocation.

The filter banks are implemented efficiently, using the WOLA scheme [2]. The complex modulations following the DFT in the analysis stage, and prior to the IDFT in the synthesis stage [2, Fig. 7.19, 7.20] may be omitted. The resulting phase modification of the quantization noise has no effect on the synthesized speech quality. The DFT is performed by the Decimation-In-Time FFT algorithm [5]. This algorithm, when used for transforming real sequences, has many redundancies. The FFT of a real sequence of length 16 requires only 12 real multiplies, 62 real adds, and the storage of only 3 constants ($\cos \pi/4$, $\cos \pi/8$, $\sin \cos \pi/8$) [6]. Similarly, the IDFT is performed by the Decimation-In-Frequency FFT algorithm [6].

The M complex signals are uniquely represented by the following M real signals:

$$Y_0(m) = X_0(m), \quad Y_{M-1}(m) = X_{M/2}(m),$$
$$Y_{2k-1}(m) = \mathrm{Re}\,X_k(m),$$
$$Y_{2k}(m) = \mathrm{Im}\,X_k(m),$$
$$k = 1, 2, \cdots, M/2 - 1. \quad (19)$$

The real signals $Y_i(m)$, $i = 0, \cdots, M - 1$ quantized independently, using forward gain-adaptive quantizers [1]. Gain adaptation in the ith band is based on the estimated and quantized variance $\hat{\sigma}_i^2$ of the signal $Y_i(m)$, which is updated every $N = 16$ samples of $Y_i(m)$ and transmitted as side information. The corresponding N samples of $Y_i(m)$ are quantized by a uniform quantizer [1] and optimized for a zero-mean Gaussian PDF with variance $\hat{\sigma}_i^2$.

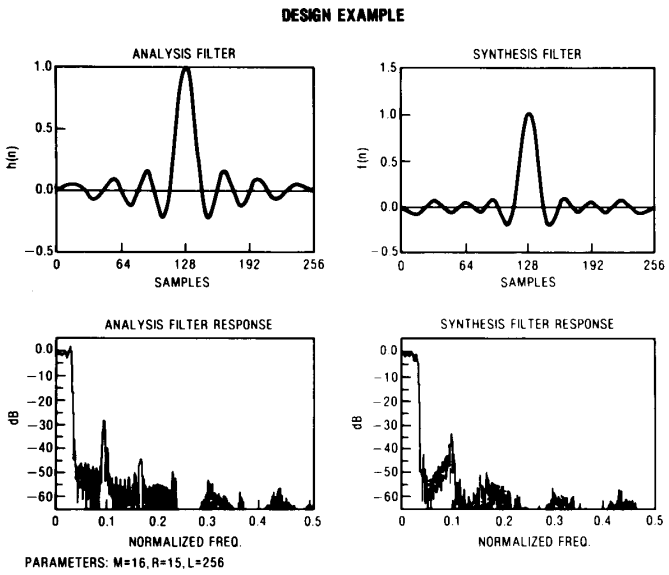


Fig. 3. Impulse and frequency responses of the optimal filters.

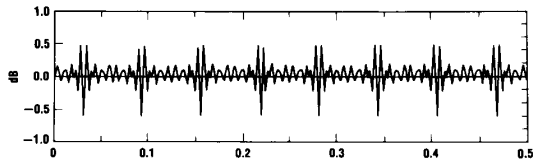


Fig. 4. The system's "transfer function" $|T(\omega)|$.

Based on the estimated variances $\hat{\sigma}_i^2$, the bit allocation is computed to minimize the MSE:

$$E\left[\sum_{k=0}^{M-1} |X_k - \hat{X}_k|^2\right] = E[Y_0 - \hat{Y}_0]^2 + E[Y_{M-1} - \hat{Y}_{M-1}]^2 + 2E\left[\sum_{i=1}^{M-2} (Y_i - \hat{Y}_i)^2\right] \quad (20)$$

where $\hat{X}$ and $\hat{Y}$ are the quantized versions of X and Y. The double weight of the errors in Y_i, $i = 1, \cdots, M-2$, is accounted for by multiplying the estimated variances σ_i^2, $i = 1, \cdots, M-2$, by 2 before computing the allocation. A procedure for bit allocation is given in [7] (after correcting some typographical errors). The allocation is also updated every $N = 16$ samples of each of $Y_i(m)$, $i = 0, 1, \cdots, M-1$.

The computation rate required for implementing the analysis filter bank is as follows: L multiplies, $L - M$ adds, and an FFT of length M for the analysis stage for producing M complex samples [2, Fig. 7.19], where L is the filter length. The same number of operations is required for the synthesis stage [2]. For a decimation ratio of $R = 15$, the implementation of both the analysis and synthesis filter banks requires 36 multiplies and 40 adds per input sample.

VI. Simulation Results and Discussion

An SBC which utilizes the above optimized filters, and an SBC which is based on an 8-band QMF bank, were simulated for a transmission rate of 16 kbits/s. The implementation of the QMF bank requires 96 multiplies and 102 adds per input sample (using 32-tap filters in a tree structure). The DFT-based coder and the QMF-based coder were found to have similar subjective performance. Both coders yield high quality synthesized speech, degraded only by very slight hoarseness. The objective performance of the QMF-based coder is slightly higher: SNR = 19.8 dB, as compared to 18.9 dB for the DFT-based coder. An attempt to reduce the QMF bank complexity to 48 multiplies per input sample, using shorter filters, caused noticeable degradation in the synthesized speech quality.

Another filter bank structure is the generalized parallel QMF bank [8]–[10], which can be implemented more efficiently than a tree-structured QMF. Still, the DFT-based SBC is expected to be more efficient. The example given in [8] is that of a filter bank having 8 real bands, which requires about 75 percent of the number of multiplies and 150 percent of the additions (using 40 tap filters) as compared to the above uniform DFT filter bank. The filter bank described in [9] splits the signal into 16 bands and requires 42 multiplies and 82 adds per input sample. In both of the above examples, quite short duration filters are used (40 taps and 80 taps, respectively), and the transition-band region of the filter in each band extends up to the center of adjacent bands. Since in the presence of quantization the power of the noise leaking to adjacent bands is proportional to the width of the transition band, subjective performance similar to the above DFT-based SBC could be expected only if longer duration filters are used, as done in [10].

VII. Summary

A new approach for designing FIR filters for uniform DFT filter banks, optimized for subband coding of speech, is presented. The optimal filters are designed by an iterative algorithm, aiming at minimizing a spectral-domain distortion function.

The distortion function takes into account both the error due to the filtering and rate conversions and the error due to quantization. Another approach for defining the distortion function, in which the quantization effect is also accounted for, was presented in [4]. In contrast to [4], the above frequency-domain distortion function enables us to separate between the error components and apply a different weight to each one of them, thereby allowing enhancement of the subjective quality of the SBC. In simulations, the leakage noise was found to be a main factor in the quality degradation of the synthesized speech.

Examining the effect of the quantization error on the filter design, we arrive at the conclusion that in medium-to-high rate SBC, minimizing the quantization effect (actually the leakage noise) by optimal filter design is approximately equivalent to attenuating the stop-band magnitude response of the synthesis filter. The above conclusion, however, need not be true for low-rate SBC (e.g., 9.6 kbits/s and below), as the additive white noise model for the quantization process is no longer valid. The latter case is considered in [4].

A 16 kbit/s SBC which utilizes the optimized filters was simulated, and was found to achieve similar subjective performance as that of a QMF-based SBC, but effecting more than 60 percent reduction in computations required for the filter bank implementation, as compared to the QMF bank.

Appendix A
Average Power Spectrum of the Output Noise

The quantization process is modeled here as additive (complex) noise, which is wide-sense stationary. The processing done to the noise in the lth band is shown in Fig. 5. The noise source power spectrum density is denoted by $S_{vl}(e^{j\omega})$, $l = 0, \cdots, M-1$, and its autocorrelation by $\phi_l(k)$. We define

$$\tilde{v}_l(n) \triangleq \begin{cases} v_l(n/R), & n \bmod R = 0 \\ 0, & \text{otherwise} \end{cases}$$

$$\tilde{\phi}_l(k) \triangleq \begin{cases} \phi_l(k/R), & k \bmod R = 0 \\ 0, & \text{otherwise.} \end{cases} \quad \text{(A1)}$$

$\tilde{v}_l(n)$ is the interpolated noise signal prior to the filtering by $f(n)$. The (time-dependent) power spectrum in the output noise of the lth band is

$$\begin{aligned} S_{yl}(n, e^{j\omega}) &= \mathfrak{F}\left\{E\left[y_l(n)\, y_l^*(n+k)\right]\right\} \\ &= \sum_k \sum_s \sum_t f(n-s) f(n+k-t) \\ &\quad \cdot W_M^{ln} W_M^{-l(n+k)} E\left[\tilde{v}_l(s)\, \tilde{v}_l^*(t)\right] e^{-j\omega k} \end{aligned}$$

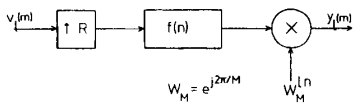


Fig. 5. Processing of the quantization noise in the lth band.

$$\begin{aligned} &= \sum_k f(k)\, W_M^{-lk} e^{-j\omega k} \sum_t \tilde{\phi}_l(tR) \\ &\quad \cdot W_M^{-ltR} e^{-j\omega tR} \sum_s f(n - sR) \\ &\quad \cdot W_M^{l(n-sR)} e^{j\omega(n-SR)} \end{aligned} \quad \text{(A2)}$$

where $\mathfrak{F}$ denotes the Fourier transform operator.

From (A2), the power spectrum is clearly periodic in time, with a period $n = R$. Defining an average power spectrum as a sum over one period, we obtain

$$\begin{aligned} \bar{S}_{yl}(e^{j\omega}) &\triangleq \sum_{n=0}^{R-1} S_{yl}(n, e^{j\omega}) \\ &= \left|F(e^{j\omega} W_M^l)\right|^2 S_{vl}(e^{j\omega R} W_M^{lR}). \end{aligned} \quad \text{(A3)}$$

We proceed by making the following two assumptions.

a) The noise sources in each pair of bands, which are not complex conjugates, are uncorrelated.

b) The real and the imaginary components of each noise source are uncorrelated, and their autocorrelation sequences are identical.

These two assumptions are approximately satisfied even by the speech band-signals, as verified by simulations. In the DFT filter bank, the noise sources v_0, $v_{M/2}$ are real, and

$$v_k = v_{M-k}^*, \qquad k = 1, 2, \cdots, M/2 - 1. \quad \text{(A4)}$$

Based on assumption b, it is easily verified that the power spectrum of the sum of the output noise terms $y_l(n) + y_{M-l}(n)$ is

$$\begin{aligned} &S_{vl}(e^{j\omega R} W_M^{lR}) \left|F(e^{j\omega} W_M^l)\right|^2 \\ &\quad + S_{v(M-l)}(e^{j\omega R} W_M^{-(M-l)R}) \left|F(e^{j\omega} W_M^{-(M-l)})\right|^2 \end{aligned} \quad \text{(A5)}$$

and following assumption a, the power spectrum of the output noise is given by the sum of the band components as in (4), where the factor $1/M^2$ is due to the division of the output signal by M (see Fig. 1).

Appendix B
Derivation of the Design Equations

A. Aliasing Distortion

From (2), using Parseval's equality,

$$E_{al} = \frac{1}{M^2 R^2} \sum_{k=1}^{R-1} \sum_n \left| \sum_{l=0}^{M-1} \left\{ f(n)\, W_M^{ln} \right\} * \left\{ h(n)\, W_M^{ln} W_R^{-kn} \right\} \right|^2 \quad \text{(B1)}$$

where $*$ denotes "convolution."

Using the identity

$$\sum_{k=0}^{M-1} W_M^{kn} = \sum_{k=0}^{M-1} e^{j2\pi kn/M} = M\,\delta(n \bmod M) = \begin{cases} M, & n = 0 \\ 0, & n \neq 0 \end{cases} \tag{B2}$$

the following expression for E_{al} is obtained:

$$E_{al} = \frac{1}{R^2} \sum_s \sum_t \sum_n f(s)\, h(n-s)\, h(n-t)\, f(t) \cdot \left[R\delta\big((s-t) \bmod R\big) - 1\right] \delta(n \bmod M). \tag{B3}$$

For simplicity, we choose the filters $h(n)$ and $f(n)$ to be of equal length L, where L is an integer multiple of the transform size M. Equation (B3) can then be rearranged in either of the following two PSD quadratic forms:

$$E_{al} = \boldsymbol{f}^T \boldsymbol{Q}_{al} \boldsymbol{f} = \boldsymbol{h}^T \tilde{\boldsymbol{Q}}_{al} \boldsymbol{h}$$

$$\boldsymbol{f} = \big[f(0), f(1), \cdots, f(L-1)\big]^T$$

$$\boldsymbol{h} = \big[h(1), h(2), \cdots, h(L)\big]^T$$

$$\boldsymbol{Q}_{al}(s,t) = \frac{1}{R^2} \sum_n h(nM-s)\, h(nM-t) \cdot \left[R\delta\big((s-t) \bmod R\big) - 1\right], \quad 0 \le s, \quad t \le L-1$$

$$\tilde{\boldsymbol{Q}}_{al}(s,t) = \frac{1}{R^2} \sum_n f(nM-s)\, f(nM-t) \cdot \left[R\delta\big((s-t) \bmod R\big) - 1\right], \quad 1 \le s, \quad t \le L \tag{B4}$$

where $\boldsymbol{Q}_{al}$ and $\tilde{\boldsymbol{Q}}_{al}$ are matrices of dimension $L \times L$. Since $E_{al} \ge 0$, both symmetric matrices are PSD.

B. MSE Relative to Unity Transmission

If $h(n)$ and $f(n)$ are linear phase filters [both of length L, as defined in (B4)], the delay of the analysis–synthesis system is $n_0 = L$ samples. Using Parseval's equality and the identity (B2), we get from (3)

$$E_{tr} = \sum_n \left| \delta(n-n_0) - \frac{1}{MR} \sum_{l=0}^{M-1} \left\{ f(n)\, W_M^{ln} \right\} * \left\{ h(n)\, W_M^{ln} \right\} \right|^2$$

$$= 1 - \frac{2}{R} \sum_s f(s)\, h(n_0 - s)\, \delta(n_0 \bmod M) + \frac{1}{R^2} \sum_s \sum_t \sum_n f(s)\, h(nM-S)\, h(nM-t)\, f(t) \tag{B5}$$

where $n_0 = L$. Rearranging (B5) in a quadratic form

$$E_{tr} = 1 - 2\boldsymbol{f}^T\boldsymbol{q} + \boldsymbol{f}^T\boldsymbol{Q}_{tr}\boldsymbol{f} = 1 - 2\boldsymbol{h}^T\tilde{\boldsymbol{q}} + \boldsymbol{h}^T\tilde{\boldsymbol{Q}}_{tr}\boldsymbol{h}$$

$$\boldsymbol{q} = \frac{1}{R}\big[h(L), h(L-1), \cdots, h(1)\big]^T$$

$$\tilde{\boldsymbol{q}} = \frac{1}{R}\big[f(L-1), f(L-2), \cdots, f(0)\big]^T$$

$$\boldsymbol{Q}_{tr}(s,t) = \frac{1}{R^2} \sum_n h(nM-s)\, h(nM-t), \quad 0 \le s, \quad t \le L-1$$

$$\tilde{\boldsymbol{Q}}_{tr}(s,t) = \frac{1}{R^2} \sum_n f(nM-s)\, f(nM-t), \quad 1 \le s, \quad t \le L. \tag{B6}$$

The symmetric matrices $\boldsymbol{Q}_{tr}$ and $\tilde{\boldsymbol{Q}}_{tr}$ are PSD due to the following equality:

$$\boldsymbol{f}^T\boldsymbol{Q}_{tr}\boldsymbol{f} = \boldsymbol{h}^T\tilde{\boldsymbol{Q}}_{tr}\boldsymbol{h} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \left| \frac{1}{MR} \sum_{l=0}^{M-1} F(e^{j\omega} W_M^{-l}) \cdot H(e^{j\omega} W_M^{-l}) \right|^2 d\omega \ge 0. \tag{B7}$$

C. Sidelobe Energy

From (8) and (13), using Parseval's equality,

$$E_h = \boldsymbol{h}^T\boldsymbol{Q}_h\boldsymbol{h}; \quad E_f = \boldsymbol{f}^T\boldsymbol{Q}_f\boldsymbol{f}$$

$$\boldsymbol{Q}_h(s,t) = u_h(s-t); \quad 1 \le s, t \le L; \quad u_h(n) = \mathfrak{F}^{-1}\big\{U_h(e^{j\omega})\big\}$$

$$\boldsymbol{Q}_f(s,t) = u_f(s-t); \quad 0 \le s, t \le L-1; \quad u_f(n) = \mathfrak{F}^{-1}\big\{U_f(e^{j\omega})\big\} \tag{B8}$$

where $\mathfrak{F}^{-1}$ denotes the inverse Fourier transform. Since $h(n)$ and $f(n)$ are FIR filters, $E_h, E_f > 0$, so both symmetric matrices $\boldsymbol{Q}_h$ and $\boldsymbol{Q}_f$ are PD. Combining (14), (B4), (B6), and (B8) we get (15), where $\boldsymbol{Q}$ and $\tilde{\boldsymbol{Q}}$ are given by

$$\boldsymbol{Q} = w_{al}\boldsymbol{Q}_{al} + \boldsymbol{Q}_{tr} + w_f\boldsymbol{Q}_f$$

$$\tilde{\boldsymbol{Q}} = w_{al}\tilde{\boldsymbol{Q}}_{al} + \tilde{\boldsymbol{Q}}_{tr} + w_h\boldsymbol{Q}_h. \tag{B9}$$

D. Case of Symmetric Filters

If the given filter is symmetric, i.e., $h(t) = h(L+1-t)$, $t = 1, \cdots, L$, or $f(t) = f(L-1-t)$, $t = 0, \cdots, L-1$, then the matrices $\boldsymbol{Q}$ and $\tilde{\boldsymbol{Q}}$ in (B9) satisfy

$$\boldsymbol{Q}(s,t) = \boldsymbol{Q}(L-1-s, L-1-t), \quad s,t = 0, 1, \cdots, L-1$$

$$\tilde{\boldsymbol{Q}}(s,t) = \tilde{\boldsymbol{Q}}(L+1-s, L+1-t), \quad s,t = 1, 2, \cdots, L. \tag{B10}$$

Following (B10), we conclude that the computed filter obtained via (16) is also symmetric. Equation (16) can therefore be reduced to compute half the filter coefficients:

$$\begin{aligned}
&\boldsymbol{P}\boldsymbol{f}' = \boldsymbol{p};\ \tilde{\boldsymbol{P}}\boldsymbol{h}' = \tilde{\boldsymbol{p}} \\
&\boldsymbol{P}(s,t) = \boldsymbol{Q}(s,t) + \boldsymbol{Q}(s, L-1-t), \\
&\qquad s,t = 0, \cdots, L/2-1 \\
&\tilde{\boldsymbol{P}}(s,t) = \tilde{\boldsymbol{Q}}(s,t) + \tilde{\boldsymbol{Q}}(s, L+1-t), \\
&\qquad s,t = 1, \cdots, L/2 \\
&\boldsymbol{h}' = \left[h(1), h(2), \cdots, h(L/2)\right]^T; \\
&\qquad \boldsymbol{f}' = \left[f(0), f(1), \cdots, f(L/2-1)\right]^T \\
&\boldsymbol{p} = \left[q(0), q(1), \cdots, q(L/2-1)\right]^T; \\
&\qquad \tilde{\boldsymbol{p}}\left[\tilde{q}(1), \tilde{q}(2), \cdots, \tilde{q}(L/2)\right]^T. \quad \text{(B11)}
\end{aligned}$$

E. Expression for c_0 (17)

Using Parseval's equality and (B2), $E_{tr}(c)$ can be expressed as

$$E_{tr}(c) = 1 - 2c\boldsymbol{f}^T\boldsymbol{q} + c^2\boldsymbol{f}^T\boldsymbol{Q}_{tr}\boldsymbol{f}. \quad \text{(B12)}$$

Minimizing $E_{tr}(c)$ results in

$$c_0 = (\boldsymbol{f}^T\boldsymbol{Q}_{tr}\boldsymbol{f})/(\boldsymbol{f}^T\boldsymbol{q}). \quad \text{(B13)}$$

References

[1] N. S. Jayant and P. Noll, *Digital Coding of Waveforms*. Englewood Cliffs, NJ: Prentice-Hall, 1984, Ch. 4, 11.

[2] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1983, Ch. 7.

[3] M. A. Kranser, "The critical band coder—Digital encoding of speech signal based on the perceptual requirements of the auditory systems," in *Proc. ICASSP*, 1980, pp. 327-331.

[4] A. Dembo and D. Malah, "Statistical design of analysis/synthesis systems with quantization," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 328-341, Mar. 1988.

[5] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1975, Ch. 6.

[6] H. V. Sorensen *et al.*, "Real-valued fast Fourier transforms algorithms," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 849-863, June 1987.

[7] R. V. Cox and R. E. Crochiere, "Real time simulation of adaptive transform coding," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 147-154, Apr. 1981.

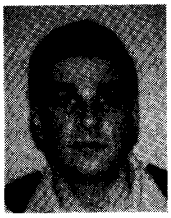
[8] J. H. Rothweiler, "Polyphase quadrature filters—A new subband coding technique," in *Proc. ICASSP*, 1983, pp. 1280-1283.

[9] P. L. Chu, "Quadrature mirror filter design for an arbitrary number of equal bandwidth channels," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 203-218, Feb. 1985.

[10] R. V. Cox, "The design of uniformly and nonuniformly spaced pseudoquadrature mirror filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 1090-1096, Oct. 1986.

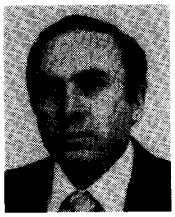
[11] K. Swaminathan and P. P. Vaidyanathan, "Theory and design of uniform DFT, parallel, quadrature mirror filter banks," *IEEE Trans. Circuits Syst.*, vol. CAS-33, pp. 1170-1191, Dec. 1986.

[12] V. K. Jain and R. E. Crochiere, "A novel approach to the design of analysis/synthesis filter banks," in *Proc. ICASSP*, 1983, pp. 228-231.



Aharon Satt (S'86) received the B.Sc. and M.Sc. degrees from the Technion—Israel Institute of Technology, Haifa, Israel, in 1986 and 1988, respectively, both in electrical engineering.

He is currently pursuing his studies toward the Ph.D. degree at the Technion. His areas of interest include digital speech processing and information theory.



David Malah (S'67-M'71-SM'84-F'87) received the B.Sc. and M.Sc. degrees in 1964 and 1967, respectively, from the Technion—Israel Institute of Technology, Haifa, Israel, and the Ph.D. degree in 1971 from the University of Minnesota, Minneapolis, all in electrical engineering.

During 1971-1972 he was an Assistant Professor at the Electrical Engineering Department of the University of New Brunswick, Fredericton, N.B., Canada. In 1972 he joined the Electrical Engineering Department of the Technion, where he is presently an Associate Professor. From 1979 to 1981 he was on sabbatical leave at the Acoustic Research Department of AT&T Bell Laboratories, Murray Hill, NJ, and was a consultant during the Summers of 1983 and 1986. Currently he is on sabbatical leave at the Signal Processing Research Department of AT&T Bell Laboratories in Murray Hill. Since 1975 (except during 1979-1981 and 1988-1989) he has been in charge of the Signal Processing Laboratory, at the EE Department, which is active in speech and image communication research and real-time hardware development. His main research interests are in image and speech coding, image and speech enhancement, and digital signal processing techniques.