Analysis/Synthesis Filter Bank Design Based on Time Domain Aliasing Cancellation

JOHN P. PRINCEN, STUDENT MEMBER, IEEE, AND ALAN BERNARD BRADLEY

Abstract—A single-sideband analysis/synthesis system is proposed which provides perfect reconstruction of a signal from a set of critically sampled analysis signals. The technique is developed in terms of a weighted overlap–add method of analysis/synthesis and allows overlap between adjacent time windows. This implies that time domain aliasing is introduced in the analysis; however, this aliasing is cancelled in the synthesis process, and the system can provide perfect reconstruction. Achieving perfect reconstruction places constraints on the time domain window shape which are equivalent to those placed on the frequency domain shape of analysis/synthesis channels used in recently proposed critically sampled systems based on frequency domain aliasing cancellation [7], [8]. In fact, a duality exists between the new technique and the frequency domain techniques of [7] and [8]. The proposed technique is more efficient than frequency domain designs for a given number of analysis/synthesis channels, and can provide reasonably band-limited channel responses. The technique could be particularly useful in applications where critically sampled analysis/synthesis is desirable, e.g., coding.

I. INTRODUCTION

ANALYSIS/SYNTHESIS techniques have found wide application and are particularly useful in speech coding [1]. The basic framework for an analysis/synthesis system is shown in Fig. 1. The analysis bank segments the signal $x(n)$ into a number of contiguous frequency bands, or channel signals $X_k(m)$. The synthesis bank forms a replica of the original signal based on the channel signals $\hat{X}_k(m)$, which are related to, but not always equal to, the signals $X_k(m)$. For example, implementation of a frequency domain speech coder [1] involves coding and decoding the channel signals, and this process generally introduces some distortion. An important requirement of any analysis/synthesis system used in coding is that, in the absence of any coding distortion, reconstruction of the signal $x(n)$ should be possible. Historically, there have been two approaches to the design and implementation of analysis/synthesis systems.

The first is based on a description of the system in terms of banks of bandpass filters, or modulators and low-pass filters [2]. Using this description, the design of the system so that reconstruction of $x(n)$ is achieved can be most eas-
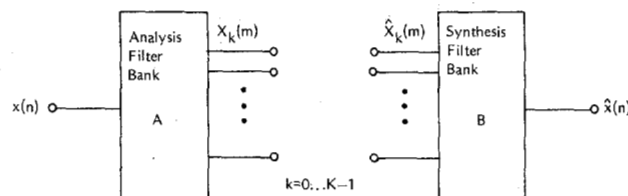
Fig. 1. Analysis/synthesis system framework.

ily expressed in the frequency domain. Reconstruction can be obtained if the composite analysis/synthesis channel responses overlap and add such that their sum is flat in the frequency domain. Any frequency domain aliasing introduced by representing the narrow-band analysis signals at a reduced sample rate must be removed in the synthesis bank [2], [4].

A second description of an analysis/synthesis system is in terms of a block transform approach (Fig. 2). Analysis involves windowing the signal and transforming the windowed signal using a "frequency domain transform." Note that the transform is restricted to be one which has an interpretation in the frequency domain [1]. This class includes the discrete Fourier transform (DFT), discrete cosine transform (DCT), and discrete sine transform (DST). The transform domain samples are the channel signals. To synthesize, the inverse transform is applied and the resulting time sequence is multiplied by a synthesis window and overlapped and added to a buffer of previously accumulated signal segments [2], [3], [5]. Again a straightforward statement of the requirements of the analysis/synthesis system, so that reconstruction of $x(n)$ is possible, can be made. Reconstruction of $x(n)$ will occur if the composite analysis/synthesis window responses overlap and add in the time domain so that the result is flat and any time domain aliasing introduced by the frequency domain representation is removed in the synthesis process.

It has been recognized that a formal duality exists between the two descriptions when the filter bank described has uniformly spaced channels and the shape of the channel responses are derived from a single low-pass prototype [2], [4]. This duality has been discussed with particular reference to filter banks which use a channel structure based on complex modulation and block transform approaches utilizing the DFT [1]–[4]. It has been shown that the approaches are in fact mathematically equivalent. This means that filter banks which use complex modulation can be designed and implemented using either interpretation.
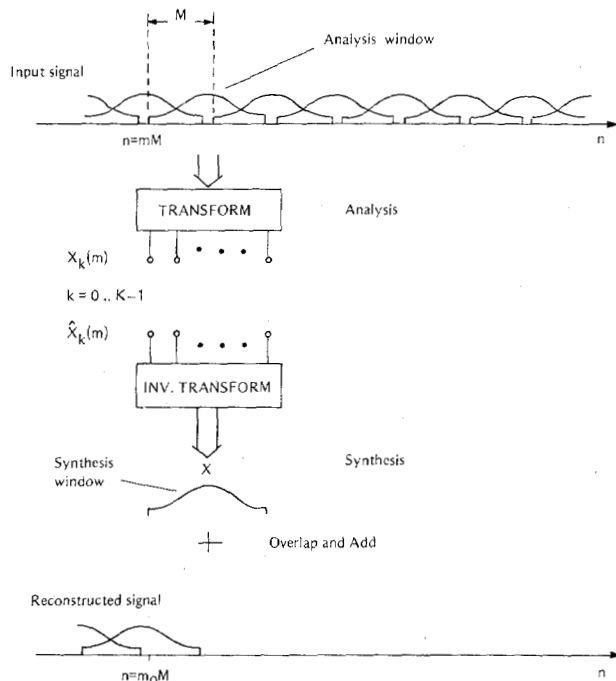
Fig. 2. A block transform description of an analysis/synthesis system.

In coding applications it is desirable that the analysis/synthesis system be designed so that the overall sample rate at the output of the analysis bank is equal to the sample rate of the input signal. In the case of a uniform filter bank with K unique channel signals all of equal bandwidth, each channel signal must be decimated by K. Systems which satisfy this condition are described as being critically sampled.

Consider the design of critically sampled analysis/synthesis systems using the two approaches outlined above. For the frequency domain approach, if a system satisfies the overlap condition for reconstruction, then critical sampling will always introduce frequency domain aliasing, except in the case where the analysis and synthesis filters have rectangular responses of width equal to the channel spacing. A dual condition exists for the time domain approach. If the time domain overlap requirement is satisfied, a critically sampled system will always introduce time domain aliasing, except if the analysis and synthesis windows are rectangular and their widths (in samples) are equal to the decimation factor.

It is possible, however, to develop critically sampled analysis/synthesis systems which have overlapped channel frequency responses and also provide reconstruction. One such technique is known as quadrature mirror filtering (QMF) and was originally proposed by Croisier et al. [6] for the case of a two-channel system. Using the basic QMF principle, a number of authors have extended the result to allow an arbitrary number of channels [7], [8]. Overlap is restricted to occur only between adjacent channel filters. For these techniques, frequency domain aliasing is introduced in the channel signals; however, the systems are designed so that this distortion is cancelled in the synthesis filter bank. The techniques cannot be de-

scribed in terms of the complex channel structure. They produce real channel signals and can be described using channel structures based on single-sideband (SSB) modulation [2].

An SSB analysis/synthesis system is shown in Fig. 3. As will be shown, it is possible to develop a time domain description of the SSB system which results in a block transform implementation that is basically the same as the transform implementation of complex filter banks; however, the DFT is replaced by appropriately defined DCT and DST. Just as a duality exists between the time and frequency domain descriptions of complex analysis/synthesis systems, a duality also exists between time and frequency domain descriptions of SSB systems. This means that the form of the time domain aliasing introduced in a critically sampled SSB system with overlapped windows is the same as that of the frequency domain aliasing introduced in a critically sampled SSB system with overlapped channel responses.

In this paper a new critically sampled SSB analysis/synthesis system is described which allows overlapped windows to be used. Overlap occurs between adjacent windows. Time domain aliasing distortion is introduced in the analysis; however, this distortion is cancelled in the synthesis process.

Since the technique is a critically sampled SSB system, the second section of this paper defines such a system and states the efficient weighted overlap–add (OLA) [2], [3], [5] analysis and synthesis equations. The analysis and synthesis involve both sine and cosine transforms, and it is shown that representation and subsequent reconstruction of a finite sequence from the appropriately defined sine or cosine transforms results in time domain aliasing. The form of this aliasing and the effect of the introduction of a time offset in the modulation functions is described in Section III.

Based on this discussion, it is shown that perfect reconstruction by time domain aliasing cancellation is possible from a critically sampled set of analysis signals. The constraints on the time domain windows are emphasized, and the final section discusses the design of suitable windows. Two examples are given. The first is similar to the windows commonly used in transform coding [1] and has a small amount of overlap. The second contains maximum allowable overlap, and the resulting channel frequency response shows that the bandwidth of such designs is close to the bandwidth of designs used in subband coding of speech.

The advantage of the new technique is that critically sampled analysis/synthesis can be performed with overlap between adjacent analysis and synthesis windows. This is not possible using conventional block transform techniques and implies that narrower channel frequency response can be obtained, while allowing critical sampling.

II. SSB Analysis/Synthesis

Consider a uniformly spaced, evenly stacked, critically sampled SSB analysis/synthesis (Fig. 3) [1], [2]. The
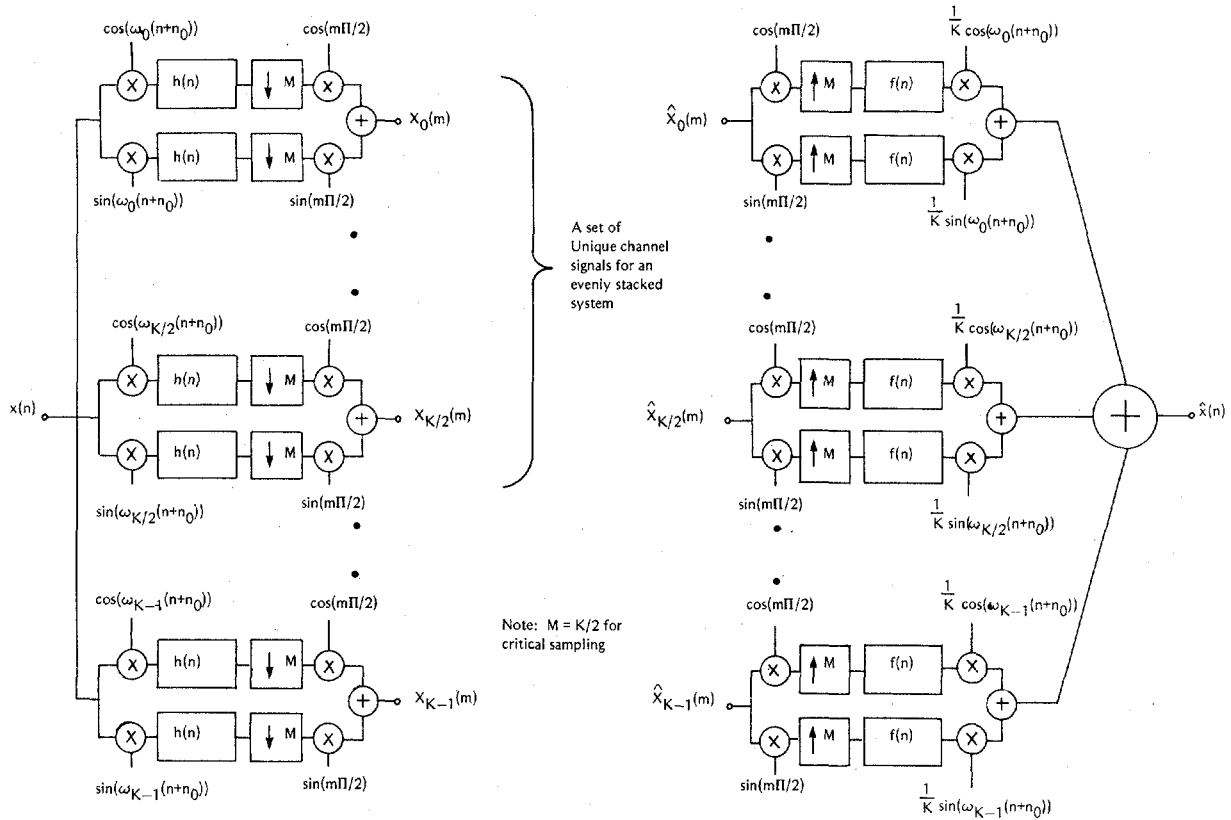
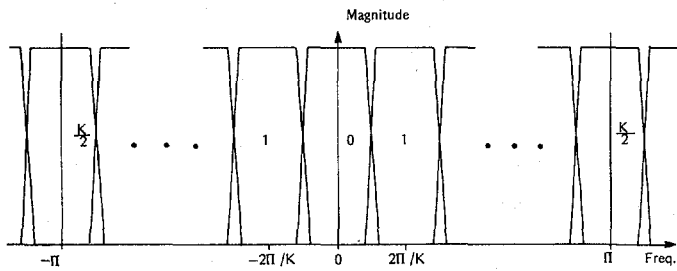Fig. 3. A single-sideband analysis/synthesis system.

Fig. 4. A bank of real bandpass filters.

channel responses are the low-pass equivalent of the bandpass filter bank shown in Fig. 4. The analysis channel signals are of course real, and the center frequencies of the K channels are

$$\omega_k = \frac{2\pi k}{K}, \qquad k = 0 \cdots K - 1. \tag{1}$$

The modulation functions have been generalized to include a time shift (n_0). The reason for this will become apparent later.

As shown in Fig. 3, each analysis and synthesis channel has two paths. These paths are time multiplexed [1], [2] since

$$\sin \left(\frac{m\pi}{2} \right) = 0, \qquad m \text{ even}$$

$$\cos \left(\frac{m\pi}{2} \right) = 0, \qquad m \text{ odd}.$$

Only $K/2 + 1$ of the K analysis channels are unique; however, in the following development the system will be described as a K-channel system, and the full K channels will be retained. This is more convenient when considering a block transform implementation. As shown in Fig. 4, half-width channels occur at $k = 0$ and $k = K/2$. If the full-width channels are decimated by M, then the half-width channels can be decimated by $2M$. Critical sampling implies that the overall output sample rate equals the input sample rate, so

$$M = \frac{K}{2}. \tag{2}$$

From Fig. 3 and (1) the analysis signals can be written as

$$X_k(m) = \cos \left(\frac{m\pi}{2} \right) \sum_{n=-\infty}^{\infty} x(n) \, h(P - 1 + mM - n)$$

$$\cdot \cos \left(\frac{2\pi k}{K} (n + n_0) \right) + \sin \left(\frac{m\pi}{2} \right) \sum_{n=-\infty}^{\infty} x(n)$$

$$\cdot h(P - 1 + mM - n) \sin \left(\frac{2\pi k}{K} (n + n_0) \right). \tag{3}$$

Both $h(n)$ and $f(n)$ are assumed to be finite impulse response filters and exist only for $0 \le n \le P - 1$. The delay of $P - 1$ has been added to make the development straightforward.

To express the analysis as a block transform operation, make the substitution $-r = mM - n$ [2], [3], [5] and use

(2) to give the OLA analysis equations

$$X_k(m) = \cos\left(\frac{m\pi}{2}\right) \sum_{r=0}^{P-1} x(mM + r) \, h(P - 1 - r)$$

$$\cdot \cos\left(\frac{2\pi}{K}\left(r + n_0 + \frac{mK}{2}\right)\right) + \sin\left(\frac{m\pi}{2}\right)$$

$$\cdot \sum_{r=0}^{P-1} x(mM + r) \, h(P - 1 - r)$$

$$\cdot \sin\left(\frac{2\pi k}{K}\left(r + n_0 + \frac{mK}{2}\right)\right). \tag{4}$$

At this point it is useful to introduce the notation

$$x_m(r) = x(mM + r)$$

$$\tilde{x}_m(r) = x_m([r]_K)$$

where $[r]_K$ represents the residue of r modulo K.
  Note that

$$\tilde{x}_m(r) = x_m(r), \qquad r = 0 \cdots K - 1.$$

Using the above notation and recognizing that $2\pi k/K \times mK/2 = \pi mk$, (4) can be reduced to

$$X_k(m) = (-1)^{mk} \cos\left(\frac{m\pi}{2}\right) \sum_{r=0}^{P-1} x_m(r) \, h(P - 1 - r)$$

$$\cdot \cos\left(\frac{2\pi k}{K}(r + n_0)\right) + (-1)^{mk} \sin\left(\frac{m\pi}{2}\right)$$

$$\cdot \sum_{r=0}^{P-1} x_m(r) \, h(P - 1 - r) \sin\left(\frac{2\pi k}{K}(r + n_0)\right). \tag{5}$$

It can be seen that for a given value of the block time m only one of the terms in (5) is nonzero. For m even, the channel signals $X_k(m)$ are the appropriately defined DCT of the windowed signal segment near time $n = mM$, modified by $(-1)^{mk}$. Similarly, for m odd, the channel signals are the modified DST of the windowed sequence near $n = mM$.

A block transform implementation of the synthesis procedure is also possible. Assuming that $\hat{X}_k(m) = X_k(m)$ and using Fig. 3 and (1), the synthesis procedure can be written as

$$\hat{x}(n) = \frac{1}{K} \sum_{k=0}^{K-1} \left\{ \cos\left(\frac{2\pi k}{K}(n + n_0)\right) \sum_{m=-\infty}^{\infty} X_k(m) \right.$$

$$\cdot \cos\left(\frac{m\pi}{2}\right) f(n - mM) + \sin\left(\frac{2\pi k}{K}(n + n_0)\right)$$

$$\left. \cdot \sum_{m=-\infty}^{\infty} X_k(m) \sin\left(\frac{m\pi}{2}\right) f(n - mM) \right\}. \tag{6}$$

Interchanging the order of summations gives

$$\hat{x}(n) = \sum_{m=-\infty}^{\infty} f(n - mM) \left\{ \frac{1}{K} \sum_{k=0}^{K-1} X_k(m) \cos\left(\frac{m\pi}{2}\right) \right.$$

$$\cdot \cos\left(\frac{2\pi k}{K}(n + n_0)\right) + \frac{1}{K} \sum_{k=0}^{K-1} X_k(m) \sin\left(\frac{m\pi}{2}\right)$$

$$\left. \cdot \sin\left(\frac{2\pi k}{K}(n + n_0)\right) \right\}. \tag{7}$$

Substituting $r = n - m_0 M$, the synthesized signal at output block time m_0 can be determined as the windowed inverse transform of $X_k(m_0)$ plus overlapping terms from other time segments [2], [3], [5], i.e.,

$$\hat{x}_{m_0}(r) = \sum_{m=-\infty}^{\infty} f((m_0 - m)M + r) \left\{ \cos\left(\frac{m\pi}{2}\right) \frac{1}{K} \right.$$

$$\cdot \sum_{k=0}^{K-1} X_k(m) \cos\left(\frac{2\pi k}{K}(r + m_0 M + n_0)\right)$$

$$+ \sin\left(\frac{m\pi}{2}\right) \frac{1}{K} \sum_{k=0}^{K-1} X_k(m) \sin\left(\frac{2\pi k}{K}\right.$$

$$\left. \left. \cdot (r + m_0 M + n_0) \right) \right\}. \tag{8}$$

To interpret the term between the braces as a block transform, let

$$\tilde{y}_m(r) = \cos\left(\frac{m\pi}{2}\right) \frac{1}{K} \sum_{k=0}^{K-1} X_k(m) \cos\left(\frac{2\pi k}{K}\right.$$

$$\left. \cdot (r + mM + n_0) \right) + \sin\left(\frac{m\pi}{2}\right) \frac{1}{K}$$

$$\cdot \sum_{k=0}^{K-1} X_k(m) \sin\left(\frac{2\pi k}{K}(r + mM + n_0)\right). \tag{9}$$

This can be simplified as was (4) using (2) to give

$$\tilde{y}_m(r) = \cos\left(\frac{m\pi}{2}\right) \frac{1}{K} \sum_{k=0}^{K-1} (-1)^{mk} X_k(m)$$

$$\cdot \cos\left(\frac{2\pi k}{K}(r + n_0)\right) + \sin\left(\frac{m\pi}{2}\right) \frac{1}{K}$$

$$\cdot \sum_{k=0}^{K-1} (-1)^{mk} X_k(m) \sin\left(\frac{2\pi k}{K}(r + n_0)\right). \tag{10}$$

That is, for m even, $\tilde{y}_m(r)$ is the inverse DCT of the modified channel signals; while for m odd, it is the inverse DST of the modified signals. The index r can be interpreted as reduced modulo K.

Using (9) or (10), we can write

$$\tilde{y}_m((m_0 - m)M + r)$$

$$= \cos\left(\frac{m\pi}{2}\right) \frac{1}{K} \sum_{k=0}^{K-1} X_k(m) \cos\left(\frac{2\pi k}{K}\right.$$

$$\left. \cdot (r + m_0 M + n_0) \right) + \sin\left(\frac{m\pi}{2}\right) \frac{1}{K} \sum_{k=0}^{K-1} X_k(m)$$

$$\cdot \sin\left(\frac{2\pi k}{K}(r + m_0 M + n_0)\right). \tag{11}$$

This can be identified as the transform operation in (8), and hence, from (8) and (11)

$$\hat{x}_{m_0}(r) = \sum_{m=-\infty}^{\infty} f((m_0 - m)M + r)$$
$$\cdot \tilde{y}_m((m_0 - m)M + r), \quad r = 0 \cdots M - 1.$$

(12)

It can be seen that since $f(n)$ is finite, the summation in (12) includes only a small number of terms. In this instance we limit both the analysis and synthesis windows to be of length P where $K/2 \le P \le K$. This implies that overlap occurs only between adjacent time segments, and at most two terms in (12) are nonzero and contribute to the final output at a given block time, i.e.,

$$\hat{x}_{m_0}(r) = f(r + M) \tilde{y}_{m_0} - 1(r + M) + f(r) \tilde{y}_{m_0}(r),$$
$$r = 0 \cdots M - 1.$$

(13)

Assuming that $P = K$, i.e., maximum overlap, makes the analysis straightforward. The results for shorter-length windows will follow by assuming that the length K window has an appropriate number of zero-valued coefficients at either end. This means that the channel signals are the modified $(\times(-1)^{mk})$ DCT or DST of the sequence [see (5)]

$$g_m(r) = h(K - 1 - r) x_m(r), \quad r = 0 \cdots K - 1.$$

(14)

Also, $\tilde{y}_m(r)$ is the inverse DCT or DST of the modified channel signals, i.e., $(-1)^{mk} X_k(m)$. Since $(-1)^{2mk} = 1$, $\tilde{y}_m(r)$ is the sequence obtained from the forward and then the reverse DCT or DST of (14). Synthesis can be implemented using the block transform approach shown in Fig. 2. The channel signals are modified by $(-1)^{mk}$, and an inverse DCT or DST transform is applied to form the sequence $\tilde{y}_m(r)$. This sequence is multiplied by the synthesis window $f(n)$ and overlapped and added to shifted output from the previous block time. In most applications, the analysis/synthesis procedure can be simplified by removing the modifications $(-1)^{mk}$ [3].

In order to determine the relationship between $\hat{x}(n)$ and $x(n)$, it is necessary to investigate the properties of a sequence recovered from the cosine and sine transforms of a real sequence.

III. PROPERTIES OF SINE AND COSINE TRANSFORMS

If we have a K-point real sequence, it is possible to represent it by K unique points in the frequency domain and recover the original sequence from these frequency domain points. However, if a K-point sequence is represented by either the sine or cosine transforms used in the SSB analysis/synthesis procedures, fewer than K unique frequency domain points are determined, and hence, the original sequence cannot be recovered from the frequency domain samples. The inverse transform yields a distorted replica of the original sequence. In fact, the distortion can

be thought of as time domain aliasing which results from an inadequately sampled frequency domain. The time domain aliasing which is introduced can be controlled to some extent by introducing a phase factor in the transform kernels.

Consider first the DCT of a K-point sequence, defined as

$$X_k = \sum_{n=0}^{K-1} x(n) \cos \left(\frac{2\pi k}{K} (n + n_0) \right),$$
$$k = 0 \cdots K - 1,$$

(15)

and the inverse transform, which gives a distorted replica of $x(n)$,

$$\bar{x}'(n) = \frac{1}{K} \sum_{k=0}^{K-1} X_k \cos \left(\frac{2\pi k}{K} (n + n_0) \right).$$

(16)

Equations (15) and (16) can be recognized as the transform operations required, when the block time is even, for the analysis and synthesis, respectively (ignoring the modifications $(-1)^{mk}$).

Substitution of (15) into (16) and use of simple trigonometry gives

$$\bar{x}'(n) = \frac{1}{2K} \sum_{r=0}^{K-1} x(r) \left\{ \sum_{k=0}^{K-1} \cos \left(\frac{2\pi k}{K} (r - n) \right) \right.$$
$$\left. + \sum_{k=0}^{K-1} \cos \left(\frac{2\pi k}{K} (r + n + 2n_0) \right) \right\}.$$

(17)

Now

$$\sum_{k=0}^{K-1} \cos \left(\frac{2\pi k}{K} n \right)$$
$$= \begin{cases} K, & n = sK, s \text{ integer} \\ 0, & \text{other integer } n, \end{cases}$$

hence,

$$\bar{x}'(n) = \bar{x}(n)/2 + \bar{x}(K - n - 2n_0)/2, \quad 2n_0 \text{ integer}$$

(18)

where the time indexes are interpreted as reduced modulo K.

Equation (18) shows that the resulting sequence is equal to the original sequence plus an aliasing term which is a shifted replica of the original sequence, time reversed. The relative shift between the two sequences can be controlled by the phase factor n_0.

Similarly, for the DST let

$$X_k = \sum_{n=0}^{K-1} x(n) \sin \left(\frac{2\pi k}{K} (n + n_0) \right),$$
$$k = 0 \cdots K - 1$$

(19)

and

$$\bar{x}'(n) = \frac{1}{K} \sum_{k=0}^{K-1} X_k \sin \left(\frac{2\pi k}{K} (n + n_0) \right).$$

(20)

Equations (19) and (20) are the transform operations re-

quired for analysis and synthesis when the block time is odd.

It can be shown that

$$\tilde{x}'(n) = \tilde{x}(n)/2 - \tilde{x}(K - n - 2n_0)/2, \quad 2n_0 \text{ integer.}$$
$$(21)$$

Again, time indexes are interpreted as reduced modulo K.

In this case the aliasing term is equal in magnitude to that in the cosine transform; however, it has opposite sign.

It can be seen that for a critically sampled SSB analysis/synthesis system, time domain aliasing is introduced. The next section shows that careful design of the windows $h(n)$ and $f(n)$ and selection of a particular value of n_0 allows this distortion to be cancelled in the OLA synthesis.

IV. Perfect Reconstruction from Critically Sampled Analysis Signals

As discussed, $\tilde{y}_m(r)$ is the sequence obtained by a forward and then inverse DCT or DST of the sequence (14). Using the reconstruction equations developed in the previous section, i.e., (18) and (21), and substituting (14), $\tilde{y}_{m_0}(r)$ can be expressed in terms of the windowed sequence as

$$\tilde{y}_{m_0}(r) = \tilde{g}_{m_0}(r)/2 + (-1)^{m_0} \tilde{g}_{m_0}(K - r - 2n_0)/2. \quad (22)$$

The recovered sequence is the sum of the original windowed sequence and an alias term. The sign of the alias term is dependent on the block time m_0 (whether a DCT or DST is applied).

For the case where m_0 is even, (22) and (14) can be used in (13) to give the recovered signal segment at $n = m_0 M$ in terms of the input signal.

$$\hat{x}_{m_0}(r) = f(r + M) \{\tilde{h}(K - 1 - r - M)$$
$$\cdot \tilde{x}_{m_0 - 1}(r + M) - \tilde{h}(r + M + 2n_0 - 1)$$
$$\cdot \tilde{x}_{m_0 - 1}(K - r - 2n_0)\}/2$$
$$+ f(r) \{\tilde{h}(K - 1 - r) \tilde{x}_{m_0}(r)$$
$$+ \tilde{h}(r + 2n_0 - 1) \tilde{x}_{m_0}(K - r - 2n_0)\}/2.$$
$$(23)$$

Noting that $M = K/2$ and

$$\tilde{x}_{m_0 - 1}(r + M) = \tilde{x}_{m_0}(r), \quad r = 0 \cdots M - 1,$$

(23) becomes

$$\hat{x}_{m_0}(r) = \tilde{x}_{m_0}(r) \{f(r + M) \tilde{h}(M - 1 - r)$$
$$+ \tilde{h}(2M - r - 1) f(r)\}/2$$
$$+ \tilde{x}_{m_0}(K - r - 2n_0) \{f(r) \tilde{h}(r + 2n_0 - 1)$$
$$- f(r + M) \tilde{h}(M + r + 2n_0 - 1)\}/2,$$
$$r = 0 \cdots M - 1. \quad (24)$$

The first term in (24) is the desired signal component, while the second term is due to time domain aliasing.

For reconstruction we require

$$f(r + M) \tilde{h}(M - 1 - r) + f(r)$$
$$\cdot \tilde{h}(2M - 1 - r) = 2, \quad r = 0 \cdots M - 1$$
$$(25a)$$

and

$$f(r) \tilde{h}(r + 2n_0 - 1) - f(r + M)$$
$$\cdot \tilde{h}(M + r + 2n_0 - 1) = 0, \quad r = 0 \cdots M - 1.$$
$$(25b)$$

Equation (25a) is the well-known condition that adjacent analysis/synthesis windows must add so that the result is flat [2], [3]. This is the time domain dual of the condition that the frequency response of the filter bank be flat.

Equation (25b) must be satisfied so that time domain aliasing introduced by the frequency domain representation is cancelled between adjacent time segments. This condition is the dual of the condition for adjacent band frequency domain aliasing cancellation in the frequency domain approaches [7], [8]. To satisfy both these conditions, choose

$$f(r) = h(r).$$

Note that

$$h(r) = \tilde{h}(r), \quad r = 0 \cdots K - 1,$$

and if

$$h(K - 1 - r) = h(r), \quad r = 0 \cdots K - 1,$$

then (25a) becomes

$$f^2(r + M) + f^2(r) = 2, \quad r = 0 \cdots M - 1 \quad (26a)$$

and (25b) becomes

$$f(r) \tilde{f}(-2M - 1 + r + 2n_0)$$
$$- f(r + M) \tilde{f}(-M - 1 + r + 2n_0) = 0,$$
$$r = 0 \cdots M - 1. \quad (26b)$$

If we let

$$2n_0 = M + 1, \quad (27)$$

then

$$\tilde{f}(-M - 1 + r + 2n_0) = f(r), \quad r = 0 \cdots M - 1,$$

and

$$\tilde{f}(-2M - 1 + r + 2n_0) = f(r + M),$$
$$r = 0 \cdots M - 1,$$

satisfying (26b).

A similar development is possible for the case when m_0 is odd. Equation (25a) would remain the same, and the signs on the aliasing components of (25b) would be inverted.
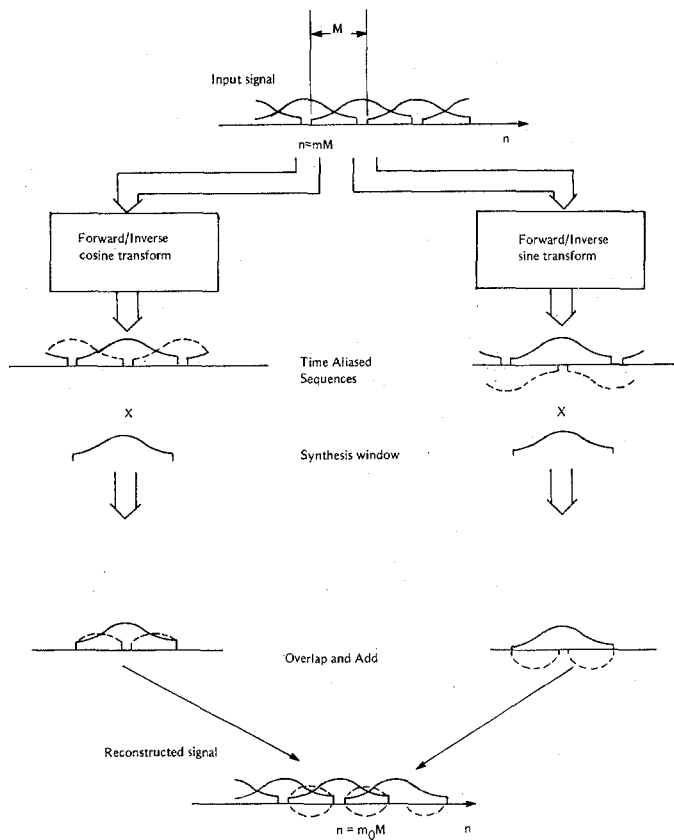
Fig. 5. The mechanism of aliasing cancellation.

The mechanism for time domain aliasing cancellation can be made obvious by the use of some simple diagrams, shown in Fig. 5. Assume that the input signal is flat in the time domain. Windowing results in a finite signal. Consider first that the block time m is even. The recovered sequence after a forward and inverse DCT contains aliasing distortion. The dashed line represents the aliased terms, which are reversed in time. Note that the alias terms are shown separated, but of course the actual signal is the sum of the original windowed sequence and the aliased sequence. The sequence can be interpreted as periodic with period K. The synthesis window extracts a portion of the sequence of length K. In the next block time, the window is shifted by $M = K/2$ and a forward and an inverse DST are performed on the windowed sequence and give the periodic sequence shown. Note that the time-reversed aliased sequence has opposite sign to that in the previous block time. The synthesis window extracts a K-point sequence, which is added to the sequence obtained in the previous block time. The aliasing terms at the upper edge of the segment from block time m, and at the lower edge of the segment from block time $m + 1$, are equal in magnitude but have opposite sign and will cancel when the time segments are overlapped and added. These diagrams may be compared to those in [8] for even and odd channel frequency responses of the dual frequency domain design.

Hence, it is possible to design an analysis/synthesis system based on an SSB structure which allows perfect reconstruction of the original signal from a set of criti-

cally sampled analysis signals by applying time domain constraints. The technique is the time domain dual of the frequency domain design techniques based on cancellation of frequency domain aliasing between adjacent channels. The analysis bank described is even and provides $K/2 + 1$ unique outputs which are critically sampled.

V. IMPLEMENTATION AND WINDOW DESIGN

Implementation of the technique can be efficiently achieved using the OLA structure shown in Fig. 2. Analysis involves windowing and then a DCT or DST as defined in Section II. Synthesis requires an inverse DCT or DST, windowing the recovered sequence, and overlapping and adding to a buffer. The data in the input and output buffers are then shifted relative to the analysis and synthesis windows, and the process is repeated for the next block time. In terms of efficiency, the method has a characteristic common to most time domain designs in that it is more efficient for a given number of bands than corresponding frequency domain designs. For example, a critically sampled design based on the methods described by Rothweiler [7] and Chu [8], which has greater than 40 dB stopband attenuation and less than 0.2 dB of passband ripple, would require an 80-sample window [8] to implement a 32-band system (16 or 17 unique bands). The new technique, with maximum overlap, uses 32-sample windows. The transform operations required are similar in both cases. It should be noted that the adjacent band overlap conditions, which are a requirement for aliasing cancellation in the frequency domain techniques, can only be approximated using finite impulse response filters. One could, of course, use shorter-length filters to implement the frequency domain techniques; however, this would increase the residual frequency domain aliasing and magnitude distortion, which exists even in the absence of coding. The time domain designs presented here have restrictions on the window length and shape which can be met exactly and, in the absence of coding, will give perfect reconstruction. The frequency domain designs presented in [7] and [8] are more efficient than the designs presented here for a given channel bandwidth. For example, the designs presented here do not closely satisfy the adjacent band overlap condition required in the frequency domain designs.

The proposed technique is, in some ways, more flexible than existing transform-type analysis/synthesis systems used in coding schemes. Using overlapped windows with conventional techniques implies that the systems must be noncritically sampled. This generally means that the windows used in transform coders are limited to have a small amount of overlap, and hence, the systems have a necessarily broad channel frequency response. Using the technique presented in this paper, significant overlap can be introduced, allowing narrower channel frequency response and critical sampling to be achieved simultaneously.

Two possible window designs are shown in Fig. 6, and the coefficient values are given in Table I. Both are for
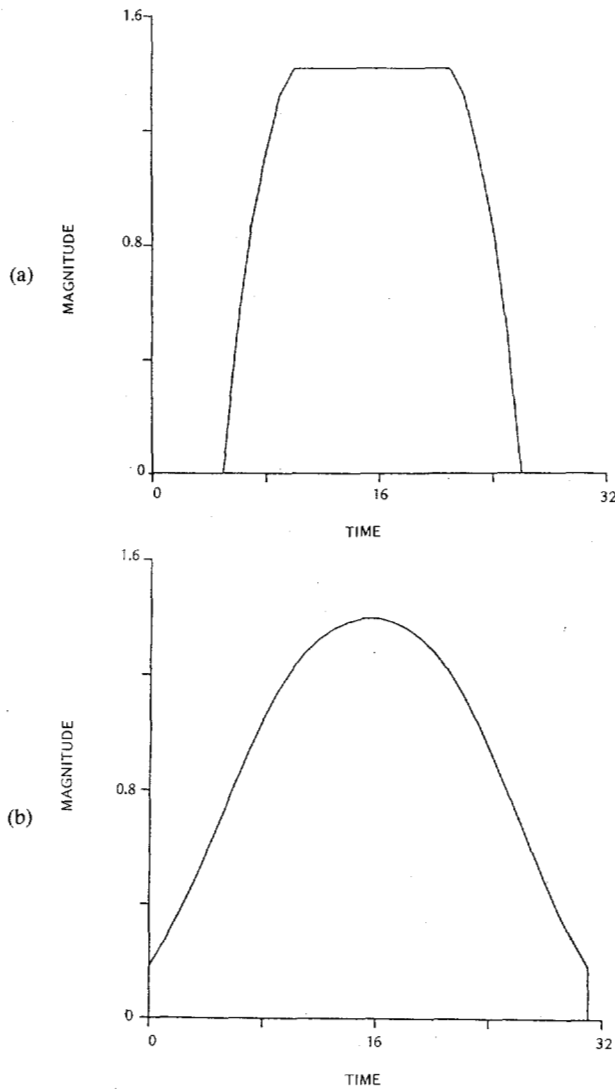
Fig. 6. (a) A 20-point window for a 32-band system. (b) A 32-point window for a 32-band system.

TABLE I
WINDOW COEFFICIENTS FOR TWO POSSIBLE DESIGNS WHICH GIVE PERFECT RECONSTRUCTION IN A 32-BAND SYSTEM

r	DESIGN 1	DESIGN 2
0	0	0.18332
1	0	0.26722
2	0	0.36390
3	0	0.47162
4	0	0.58798
5	0	0.70847
6	0.50000	0.82932
7	0.86603	0.94533
8	1.11803	1.05202
9	1.32287	1.14558
10	1.41421	1.22396
11	1.41421	1.28610
12	1.41421	1.33307
13	1.41421	1.36655
14	1.41421	1.38866
15	1.41421	1.40212

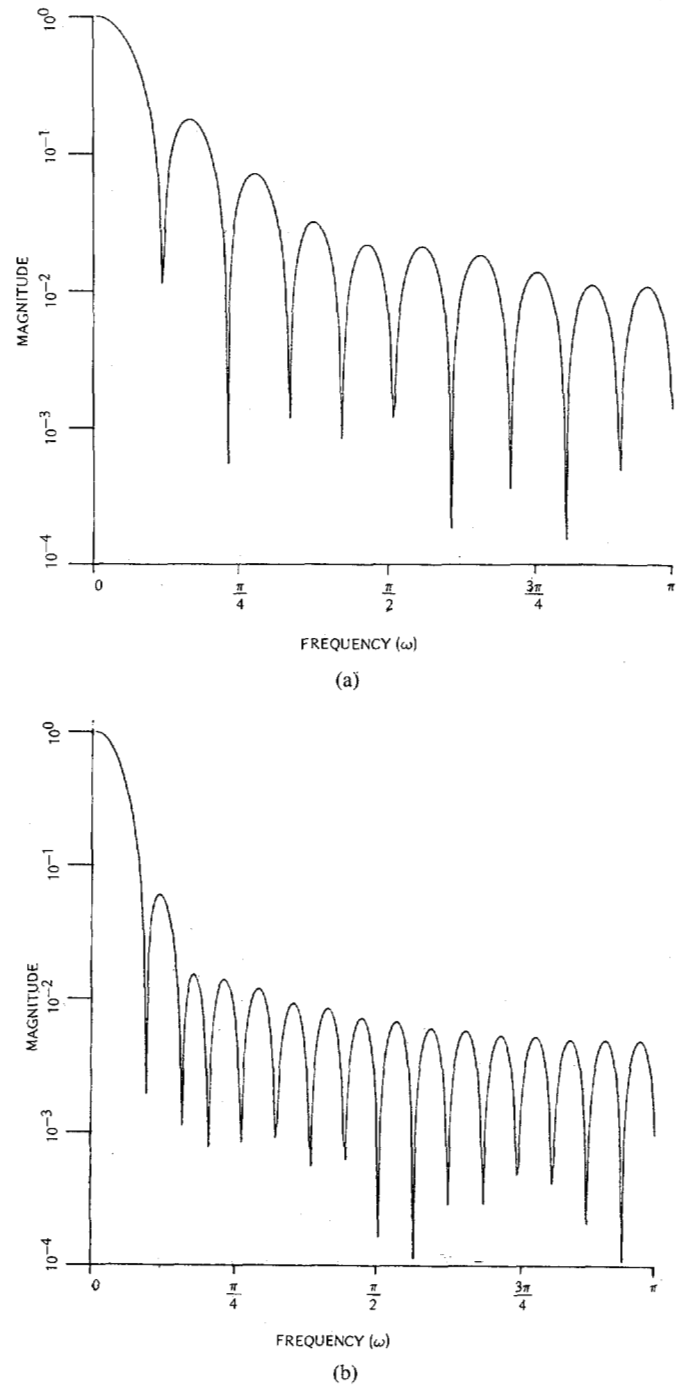Windows symmetric, ie,
$h(K-1-r) = h(r)$

Fig. 7. (a) The frequency response of an analysis (synthesis) channel using the 20-point window. (b) The frequency response for the 32-point window.

32-band evenly stacked (i.e., 17 unique bands) systems. The designs represent two extremes. The first is similar to designs used in current transform coders and gives only a small amount of overlap between adjacent windows. The corresponding channel response is shown in Fig. 7(a). Fig. 6(b) shows a design which exhibits maximum overlap. This design represents an analysis/synthesis system with properties in between those currently used in subband coding and those used in transform coders. The frequency response is shown in Fig. 7(b).

It is not suggested that these window designs are optimum in any sense for use in coding systems. An investigation of the performance of the new technique in a cod-

ing system and appropriate window designs for these systems is a subject for future research. We can say, however, that the technique presented allows a number of design choices not possible with other critically sampled analysis/synthesis techniques.

VI. Conclusion

In this paper a new critically sampled SSB analysis/synthesis system which provides perfect reconstruction has been developed. The technique allows overlap between adjacent analysis and synthesis windows, and hence, time domain aliasing is introduced in the analysis. However, this aliasing is cancelled in the synthesis process. Achieving aliasing cancellation places time domain constraints on the analysis and synthesis windows. The proposed system was developed using the OLA method of analysis/synthesis, which also provides an efficient implementation. The technique is the dual of recently developed critically sampled frequency domain designs based on aliasing cancellation. It adds considerable flexibility in the design of transform coders, can provide reasonably band-limited channel responses, and is more efficient than corresponding frequency domain designs for a given number of bands.

Acknowledgment

References

[1] J. M. Tribolet and R. E. Crochiere, "Frequency domain coding of speech," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-27, pp. 512-530, Oct. 1979.

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

[3] R. E. Crochiere, "A weighted overlap-add method of short time Fourier analysis/synthesis," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-28, pp. 99-102, Feb. 1980.

[4] J. B. Allen and L. R. Rabiner, "A unified approach to short-time Fourier analysis and synthesis," *Proc. IEEE*, vol. 65, pp. 1558-1564, Nov. 1977.

[5] M. R. Portnoff, "Implementation of the digital phase vocoder using the fast Fourier transform," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-24, pp. 243-248, June 1976.

[6] A. Croisier et al., "Perfect channel splitting by use of interpolation/decimation/tree decomposition techniques," presented at the Int. Conf. Inform. Sci./Syst., Patras, Greece, Aug. 1976.

[7] J. Rothweiler, "Polyphase quadrature filters—A new subband coding technique," presented at the ICASSP, Boston, MA, 1983.

[8] 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.
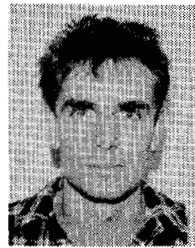
John P. Princen (S'83) was born in Melbourne, Australia, on September 12, 1961. He received the B.Eng. degree in communications engineering with distinction from the Royal Melbourne Institute of Technology (R.M.I.T.), Melbourne, in 1984 and has completed the M.Eng. program in the area of analog speech scrambling, also at R.M.I.T.

His research interests include digital signal processing and digital communciations.

Alan Bernard Bradley was born in Melbourne, Australia, in 1946. He received the B.E. degree in electrical engineering from Melbourne University in 1968 and the M.E.Sc. degree in electrical engineering from Monash University in 1972.

Since 1973 he has been working as Lecturer and Senior Lecturer in the Department of Communication and Electronic Engineering at the Royal Melbourne Institute of Technology, Melbourne. His research interests are in the field of digital signal processing, with particular emphasis on speech analysis and coding.