ELSEVIER

Signal Processing 36 (1994) 163–174

SIGNAL PROCESSING

FIR approximations of inverse filters and perfect reconstruction filter banks

Michael Unser,*,† Murray Eden

Biomedical Engineering and Instrumentation Program, Bldg 13, Room 3W13, National Center for Research Resources, National Institutes of Health, Bethesda, MD 20892, USA

Received 31 October 1991; revised 24 February 1992, 16 July 1992 and 30 November 1992

Abstract

This paper first describes an algorithm that finds the approximate finite impulse response (FIR) inverse of an FIR filter by minimizing the inversion (or reconstruction) error constrained to zero-bias. The generalization of the inverse filtering problem in the two channel case is the design of perfect reconstruction filter banks that use critical sampling. These considerations lead to the derivation of an algorithm that provides a minimum error and unbiased FIR/FIR approximation of a perfect reconstruction IIR/FIR (or FIR/IIR) filter bank. The one-channel algorithm is illustrated with the design of an FIR filter to compute the B-spline coefficients for cubic spline signal interpolation. The two-channel algorithm is applied to the design of a FIR/FIR filter bank that implements the cubic B-spline wavelet transform. Finally, we consider a modification of this technique for the design of modulated-filter banks, which are better suited for subband coding.

Zusammenfassung

Dieser Artikel beschreibt zunächst einen Algorithmus, der näherungsweise die FIR-Inverse eines FIR-Filters durch Minimierung des Inversions- (oder Rekonstruktions-) Fehlers unter der Nebenbedingung eines verschwindenden mittleren Fehlers (Bias) findet. Die Verallgemeinerung des inversen Filterproblems im Zweikanal-Fall ist der Entwurf einer Filterbank mit perfekter Rekonstruktion, die kritische Abtastung benützt. Diese Überlegungen führen zur Ableitung eines Algorithmus, der einen minimalen Fehler und eine biasfreie FIR/FIR Näherung einer perfekten IIR/FIR (oder FIR/IIR) Rekonstruktionsfilterbank liefert. Der Einkanal-Algorithmus wird mit dem Entwurf eines FIR-Filters veranschaulicht, um die B-Spline Koeffizienten für die kubische Spline-Signalinterpolation zu berechnen. Der Zweikanal-Algorithmus wird zum Entwurf einer FIR/FIR Filterbank, die die kubische B-Spline Wavelet-Transformation implementiert, angewendet. Zum Schluß wird eine Modifikation dieser Technik für den Entwurf von modulierten Filterbänken betrachtet, die besser geeignet sind für Teilband-Codierung.

Résumé

Nous présentons d'abord un algorithme qui calcule une approximation à réponse impulsionnelle finie (FIR) de l'inverse d'un filtre FIR. Cette méthode se base sur la minimisation de l'erreur de reconstruction sous la contrainte d'un

* Corresponding author. Email: unser@helix.nih.gov
† On leave from INSERM U. 2, Hôpital Henri-Mondor, F-94010 Créteil, France.

Elsevier Science B.V.
SSDI 0165-1684(93)E0083-W

biais nul. La généralisation du problème d'inversion pour deux canaux est la synthèse de bancs de filtres à reconstruction parfaite. Ceci mène à la description d'un algorithme pour la determination d'une approximation FIR/FIR aux moindres carrés d'un banc de filtres IIR/FIR (ou FIR/IIR) à reconstruction parfaite. Le premier algorithme est illustré avec la synthèse d'un filtre FIR permettant le calcul directe des coefficients B-splines pour une interpolation du signal par splines cubiques. Le second algorithme est appliqué à la synthèse d'un banc de filtres FIR/FIR permettant le calcul rapide de la transformation d'ondelettes de type B-spline cubique. Finalement, nous présentons une variation de cette méthode pour la synthèse de banc de filtres modulés qui sont plus appropriés pour le codage en sous-bandes.

Keywords: Least-squares design; Inverse filter; Perfect reconstruction filler bank; Wavelet transform; Spline functions; Cubic splines; B-splines; Gabor transform; Interpolation

1. Introduction

There are many signal transformations that can be formulated in terms of shift-invariant convolution operators. It is often of interest to perform, or at least approximate, the inverse transformation. This task involves the design of inverse or pseudo-inverse filters. In the one-channel case, this problem is equivalent to designing a deconvolution filter, which is especially relevant for signal restoration. When the initial operator has a finite impulse response (FIR), the inverse filter that allows signal recovery is generally recursive which may lead to instabilities and limit cycles. To avoid such problems, Gülboy and Geçkinli [9] have proposed using an FIR filter approximation derived from the optimization of a least-squares criterion; a similar algorithm has also been described in [17]. The use of inverse filters can also be useful in other applications, for example, polynomial spline interpolation for which this technique provides an efficient way to solve the system of banded equations that determine the B-spline coefficients [21].

In the two (or more) channel case, this basic inversion problem is closely related to the design of multirate filter banks that permit perfect reconstruction [24, 26]. An example of a system that uses critical sampling and allows perfect signal recovery is the quadrature mirror filter (QMF) bank [3, 24]. Such filter banks can be applied hierarchically to produce subband signal decompositions. The application of this concept to image coding has led to very promising results [1, 19, 25, 28]. More recently, Mallat showed that the same computational technique could be used to implement the wavelet transform [12, 13]. This latter representation is the expansion of a signal in terms of hierarchical basis functions obtained by translation and dilation of a single template: the wavelet function [4, 13, 14].

In this paper, we investigate the special case in which the transformations to be inverted involve given FIR filters. One practical problem in this situation is that the corresponding inverse operators generally have an infinite impulse response (IIR). Our approach is to use instead a truncated FIR approximation with a prescribed error tolerance. For this purpose, we apply a general design principle based on the minimization of the reconstruction error. One variation of this procedure is a constrained optimization that demands perfect reconstruction of the DC signal component (zero-bias constraint). In the one-channel case, this formulation yields a matrix algorithm for the FIR approximation of an all-pole filter. In the two-channel case, it results in a least-squares design technique for perfect reconstruction filter banks; this procedure is applicable whenever the analysis (or synthesis) filters are certain specified FIR kernels. This direct method provides an alternative to the iterative LMS-based algorithm recently proposed by Paillard et al. [15]; it also offers the flexibility of adding constraints to the system. The same computational techniques may also be useful in designing fast algorithms for the determination of the expansion coefficients of a continuous signal representation in terms of shifted basis functions of compact support. Such representations are frequently used for signal interpolation [10, 16, 21]; they occur as well in the context of the wavelet transform [22, 27].

The presentation is organized as follows. Section 2 provides a precise statement of the perfect reconstruction (PR) problem in the one-channel (inverse filter) and two-channel case. Section 3 is concerned with the derivation of a constrained least-squares procedure for designing an FIR filter to approximate the inverse of a given convolution kernel. Section 4 extends this technique to the two-channel case and presents a general design technique for a minimum reconstruction error FIR/FIR filter bank. Section 5 presents a few experimental examples and compares the proposed approaches with a reference technique that uses simple truncation of the theoretical impulse responses. The one-channel algorithm is applied to the design of an inverse filter for efficient implementation of the direct cubic spline transform [21]. The two-channel algorithm is illustrated with the design of a fast algorithm for the cubic B-spline wavelet transform [2, 23]. It is worth noting that the corresponding decomposition provides a reversible multiresolution signal analysis that is close to optimal in terms of its time/frequency localization [22]. This representation is in many respects similar to a hierarchical or wavelet-like Gabor transform [5–7].

1.1. Notations and operators

A signal $\{a(k)\}_{k\in Z}$ is characterized by its z-transform, which we denote by a capital letter

$$A(z) = \sum_{k=-\infty}^{+\infty} a(k)z^{-k}. \tag{1.1}$$

This correspondence is also expressed as: $a(k) \overset{z}{\longleftrightarrow} A(z)$. The unit impulse at $k = i$ is represented by the symbol

$$\delta_i(k) = \begin{cases} 1, & k = i, \\ 0, & \text{otherwise}, \end{cases} \quad \overset{z}{\longleftrightarrow} \quad z^{-i}. \tag{1.2}$$

The down-sampling (or decimation) by a factor of two is defined by

$$[a]_{\downarrow 2}(k) = a(2k) \quad \overset{z}{\longleftrightarrow} \quad \tfrac{1}{2}(A(z^{1/2}) + A(-z^{1/2})). \tag{1.3}$$

The complementary operation is the up-sampling by a factor of two

$$[a]_{\uparrow 2}(k) = \begin{cases} a(k/2), & k \text{ even}, \\ 0, & k \text{ odd}, \end{cases} \quad \overset{z}{\longleftrightarrow} \quad A(z^2). \tag{1.4}$$

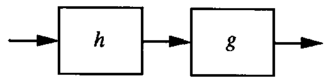

Fig. 1. Block diagram for the inverse filtering problem

2. Statement of the problem

2.1. Inverse filter

Let us consider the block diagram in Fig. 1, where the convolution operator g is an FIR filter of length $M = k_2 - k_1 + 1$:

$$g(k) \overset{z}{\longleftrightarrow} G(z) = \sum_{k=k_1}^{k_2} g(k)z^{-k}. \tag{2.1}$$

We seek an operator h that compensates exactly for the effect of g. This leads to the constraint:

$$\forall k \in Z, \quad h * g(k) = \delta_0(k), \tag{2.2}$$

where $\delta_0(k)$ is the unit impulse at the origin (cf. (1.2)). A filter satisfying (2.3) is referred to as the inverse filter of g and is characterized by

$$h(k) = (g)^{-1}(k) \overset{z}{\longleftrightarrow} H(z) = 1/G(z). \tag{2.3}$$

The operator is stable provided that the complex roots of $G(z)$ are not on the unit circle. In general the inverse filter has an IIR. In Section 3, we will consider the problem of finding an FIR approximation of $(g)^{-1}$ that satisfies (2.2) in the least-square sense.

2.2. IIR/FIR perfect reconstruction filter banks

Fig. 2 represents the block diagram of a standard two-channel PR filter bank. Note that two delays have been included in the definition of the filters in the lower branch in order to simplify some of the notation. Here, we will consider the special situation in which the reconstruction (or synthesis) filters are two FIR filters given a priori:

$$g_i(k) \overset{z}{\longleftrightarrow} G_i(z) = \sum_{k=k_{1i}}^{k_{2i}} g_i(k)z^{-k}, \quad i = 1, 2. \tag{2.4}$$

By expressing the effect of down-sampling and up-sampling in the z-transform domain (cf. (1.3) and

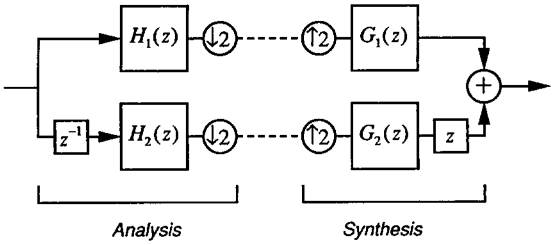

Fig. 2. Block diagram for the two channel perfect reconstruction problem.

(1.4)), we get the following input/output relation:

$$B(z) = G_1(z)\tfrac{1}{2}(A(z)H_1(z) + A(-z)H_1(-z) + G_2(z)\tfrac{1}{2}(A(z)H_2(z) - A(-z)H_2(z)),$$

where $A(z)$ and $B(z)$ denote the z-transform of the input and output signal, respectively. Regrouping the terms, we get

$$B(z) = A(z)\tfrac{1}{2}(H_1(z)G_1(z) + H_2(z)G_2(z)) + A(-z)\tfrac{1}{2}(H_1(-z)G_1(z) - H_2(-z)G_2(z))$$

From this last equation, we can easily obtain the conditions for a perfect reconstruction (i.e. the global output of the system is equal to its input)

$$\begin{aligned} &H_1(z)G_1(z) + H_2(z)G_2(z) = 2, \\ &H_1(-z)G_1(z) - H_2(-z)G_2(z) = 0, \end{aligned} \tag{2.5}$$

which are well known in multirate filter bank theory [24, 26]. We note that the modulated version of the last equation, obtained by replacing z by $-z$, must also be satisfied. This property implies that perfect reconstruction is also achieved when the location of the analysis and synthesis filters are interchanged. Therefore, we can write the system of equations in terms of the unknowns $H_1(z)$ and $H_2(z)$,

$$\begin{bmatrix} G_1(z) & G_2(z) \\ G_1(-z) & -G_2(-z) \end{bmatrix} \begin{bmatrix} H_1(z) \\ H_2(z) \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}. \tag{2.6}$$

Assuming that the polynomial in z corresponding to the determinant

$$G_{12}(z) = -(G_1(z)G_2(-z) + G_1(-z)G_2(z)) \tag{2.7}$$

has no complex root on the unit circle, the solution is given by

$$\begin{aligned} H_1(z) &= -2G_2(-z)/G_{12}(z), \\ H_2(z) &= -2G_1(-z)/G_{12}(z). \end{aligned} \tag{2.8}$$

The filters $H_1(z)$ and $H_2(z)$ are formed from the ratio of two polynomials, and therefore their impulse responses are generally infinite (IIR). The problem that will be studied in Section 4 is to obtain good FIR approximations of both h_1 and h_2 for a given set of FIR reconstruction or synthesis filters g_1 and g_2.

2.3. Approximation errors

When an FIR approximation h of the inverse filter $(g)^{-1}$ is used, the inversion criterion (2.2) cannot be satisfied perfectly. This discrepancy can be measured by the error criterion:

$$\varepsilon_0 = \sqrt{\sum_{k=-\infty}^{+\infty} [h * g(k) - \delta_0(k)]^2}; \tag{2.9}$$

$\varepsilon_0 = 0$ if and only if $h(k) = (g)^{-1}(k)$. A systematic error is introduced when the zero frequency gain of $h(k)$ is not precisely equal to the reciprocal of that of $g(k)$. This type of error – also referred to as bias – is defined as

$$\text{bias}_0 = |1 - H(z)G(z)|_{z=1}|, \tag{2.10}$$

where the zero frequency (i.e. $z = 1$) gains $H(1)$ and $G(1)$ also correspond to the summation of the corresponding impulse responses over the integers. The bias distortion will affect the DC component of the signal and can therefore account for a very large fraction of the total energy. This is especially true in image processing applications where the pixel values are positive by definition; this implies that the average gray level value is always greater than the standard deviation of the signal. Therefore, one should seek a design that reduces the bias insofar as possible.

For the PR filter bank, an FIR approximation will introduce two types of error. The first is the distortion error:

$$\varepsilon_1 = \sqrt{\sum_{k=-\infty}^{+\infty} \tfrac{1}{2}[h_1 * g_1(k) + h_2 * g_2(k) - 2\delta_0(k)]^2}, \tag{2.11}$$

which expresses the fact that the global system response is not precisely an identity. The second is the aliasing error introduced by the sampling process:

$$\varepsilon_2 = \sqrt{\sum_{k=-\infty}^{+\infty} \tfrac{1}{2}[h_1 * \tilde{g}_1(k) - h_2 * \tilde{g}_2(k)]^2}, \qquad (2.12)$$

where the notation $\tilde{g}(k) = (-1)^k g(k)$ refers to the modulation of a signal. We note that the PR constraints (2.6) are equivalent to specifying an undistorted ($\varepsilon_1 = 0$) and alias-free reconstruction ($\varepsilon_2 = 0$).

In most cases, the upper branch of the filter bank in Fig. 2 corresponds to a lowpass filter, while the lower one performs as a highpass filter. It is therefore also useful to measure the bias that affects the global system. Assuming $G_2(1) = 0$, the bias is given by

$$\text{bias}_1 = |1 - H_1(z)G_1(z)|_{z=1}|.$$

The design techniques presented next are based on the constrained minimization of these error criteria.

3. Inverse filter design

Let $\{g(k),\ k = k_1, \ldots, k_2\}$ be an FIR filter of length $M = k_2 - k_1 + 1$. The problem is to find an approximation of its inverse in terms of an FIR filter $\{h(l),\ l = l_1, \ldots, l_2\}$ of length $N = l_2 - l_1 + 1$. The approach suggested by Gülboy and Geçkinli is to minimize a mean square error criterion equivalent to ε_0^2 [9]. These authors propose a design method relying on matrix algebra for the particular case where $N = M$.

Here, we will extend their technique in two respects. First, we will allow any given value of N, except that $k_1 + l_1 \leqslant 0$ and $k_2 + l_2 \geqslant 0$ (these two bounds define the beginning and end points of the sequence $h * g \cong \delta_0$, respectively). Second, we will impose an additional constraint on the impulse response of the approximation:

$$\sum_{l=l_1}^{l_2} h(l) = \left(1 \Big/ \sum_{k=k_1}^{k_2} g(k)\right) = \alpha, \qquad (3.1)$$

which, at least, guarantees exact inversion for a constant signal (i.e. $\text{bias}_0 = 0$). Using Lagrange multipliers, we introduce the following criterion:

$$L(h(l_1), \ldots, h(l_2), \lambda) = \sum_{k=-\infty}^{+\infty} (\delta_0(k) - h * g(k))^2 + 2\lambda\left(\sum_{l=l_1}^{l_2} h(l) - \alpha\right). \qquad (3.2)$$

Using the fact that $h * g(k) = \sum_{l=l_1}^{l_2} h(l)g(k-l)$ is a sequence of length $K = M + N - 1$ which is defined (i.e. not equal to zero) for $k = k_1 + l_1, \ldots, k_2 + l_2$, this criterion is rewritten in matrix form as

$$L(\boldsymbol{h}, \lambda) = (\boldsymbol{Gh} - \boldsymbol{e})'(\boldsymbol{Gh} - \boldsymbol{e}) - 2\lambda(\boldsymbol{h}'\boldsymbol{1} - \alpha), \qquad (3.3)$$

where $\boldsymbol{G} = [g_{ij}]$ is the $K \times N$ convolution matrix

$$g_{ij} = g(k_1 + i - j), \quad i = 1, \ldots, K, j = 1, \ldots, N, \qquad (3.4)$$

$\boldsymbol{e} = [e_i]$ is the K-dimensional unit vector

$$e_i = \delta_0(k_1 + l_1 - 1 + i), \quad i = 1, \ldots, K \qquad (3.5)$$

and $\boldsymbol{h} = [h_j]$ the N-dimensional vector of unknown filter coefficients:

$$h_j = h(l_1 - 1 + j), \quad j = 1, \ldots, N. \qquad (3.6)$$

$\boldsymbol{1} = [1 \ldots 1]'$ an N-dimensional vector of ones; the symbol $'$ denotes the transpose operator. The optimum filter coefficients are obtained by setting the derivative of (3.3) with respect to $\boldsymbol{h}$ to zero, which using standard rules of matrix differentiation [8], yields

$$\frac{\partial L}{\partial \boldsymbol{h}} = 2\boldsymbol{G}'\boldsymbol{Gh} - 2\boldsymbol{G}'\boldsymbol{e} - 2\lambda\boldsymbol{1} = 0.$$

It follows that the general solution can be expressed as

$$\boldsymbol{h} = \boldsymbol{u} + \lambda\boldsymbol{v}, \qquad (3.7)$$

where

$$\boldsymbol{u} = (\boldsymbol{G}'\boldsymbol{G})^{-1}\boldsymbol{G}'\boldsymbol{e}, \qquad (3.8)$$

$$\boldsymbol{v} = (\boldsymbol{G}'\boldsymbol{G})^{-1}\boldsymbol{1}, \qquad (3.9)$$

and where λ is chosen to satisfy (3.1), i.e.

$$\lambda = \frac{\alpha - \boldsymbol{1}'\boldsymbol{u}}{\boldsymbol{1}'\boldsymbol{v}}. \qquad (3.10)$$

We note that $\boldsymbol{u}$ provides the unconstrained least-squares solution and is comparable to the solution found by Gülboy and Geçkinli [9]. The correction term $\boldsymbol{v}$ is used to impose the constraint of a normalized impulse response.

4. Design of FIR pseudo-inverse filter banks

Let $\{g_i(k), k = k_{1i}, \ldots, k_{2i}\}$, $i = 1, 2$ be two given FIR filters of length M_1 and M_2, respectively. The problem is to find the FIR approximations $\{h_i(l), l = l_{1i}, \ldots, l_{2i}\}$, $i = 1, 2$ of length N_1 and N_2 of the analysis filters on the left-hand side of the block diagram in Fig. 2. We also impose a constraint on the lowpass branch of the filter bank:

$$\sum_{l=l_{11}}^{l_{21}} h_1(l) = \left(2 \Big/ \sum_{k=k_{11}}^{k_{21}} g_1(k)\right) = \alpha. \tag{4.1}$$

We will solve this problem, which is a generalization of the previous one, by minimizing the total error criterion $\varepsilon_1^2 + \varepsilon_2^2$ (cf. Eqs. (2.11) and (2.12)) subject to the constraint (4.1). We note that the same procedure is also applicable when the role of the analysis and synthesis filters are interchanged. Using Lagrange multipliers, we introduce the functional

$$\begin{aligned} L(\boldsymbol{h}_1, \boldsymbol{h}_2, \lambda) &= \tfrac{1}{2}(\boldsymbol{G}_1\boldsymbol{h}_1 + \boldsymbol{G}_2\boldsymbol{h}_2 - \boldsymbol{e})'(\boldsymbol{G}_1\boldsymbol{h}_1 + \boldsymbol{G}_2\boldsymbol{h}_2 - \boldsymbol{e}) \\ &\quad + \tfrac{1}{2}(\tilde{\boldsymbol{G}}_1\boldsymbol{h}_1 - \tilde{\boldsymbol{G}}_2\boldsymbol{h}_2)'(\tilde{\boldsymbol{G}}_1\boldsymbol{h}_1 - \tilde{\boldsymbol{G}}_2\boldsymbol{h}_2) \\ &\quad - \lambda(\boldsymbol{1}'\boldsymbol{h}_1 - \alpha), \end{aligned} \tag{4.2}$$

where the first and second quadratic terms are precisely ε_1^2 and ε_2^2. In this notation, $\boldsymbol{h}_1 = [h_{j1}]$ and $\boldsymbol{h}_2 = [h_{j2}]$ are the unknown N_1- and N_2-dimensional filter coefficient vectors:

$$h_{ji} = h_i(l_{1i} - 1 + j); \quad j = 1, \ldots, N_i, \quad i = 1, 2, \tag{4.3}$$

with $N_1 = l_{21} - l_{11} - 1$ and $N_2 = l_{22} - l_{12} - 1$. $\boldsymbol{1}$ is an N_1-dimensional vector of ones and $\boldsymbol{e} = [e_i]$ is a K-dimensional vector defined by

$$e_i = 2\delta_0(i_0 - 1 + i), \quad i = 1, \ldots, K, \tag{4.4}$$

where $i_0 = \min\{k_{11} + l_{11}, k_{12} + l_{12}\}$ and where

$$\begin{aligned} K &= \max\{k_{21} + l_{21}, k_{22} + l_{22}\} \\ &\quad - \min\{k_{11} + l_{11}, k_{12} + l_{12}\} + 1. \end{aligned} \tag{4.5}$$

The convolution matrices $\boldsymbol{G}_1 = [g_{ij1}]$, $\tilde{\boldsymbol{G}}_1 = [\tilde{g}_{ij1}]$, $\boldsymbol{G}_2 = [g_{ij2}]$ and $\tilde{\boldsymbol{G}}_2 = [g_{ij2}]$ are of size $K \times N_1$, $K \times N_1$, $K \times N_2$ and $K \times N_2$, respectively. They are defined as

$$\begin{aligned} &g_{ijk} = g_k(i_k + i - j), \\ &\quad i = 1, \ldots, K; j = 1, \ldots, N_k, \quad k = 1, 2, \end{aligned} \tag{4.6}$$

$$\begin{aligned} &\tilde{g}_{ijk} = (-1)^{i_k + i - j} g_k(i_k + i - j), \\ &\quad i = 1, \ldots, K; j = 1, \ldots, N_k, \ k = 1, 2, \end{aligned} \tag{4.7}$$

where the offset constants i_1 and i_2 are given by

$$i_k = \min\{k_{11} + l_{11}, k_{12} + l_{12}\} - l_{1k}, \quad k = 1, 2.$$

The minimization of (4.2) with respect to $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ yields the system of equations:

$$[\boldsymbol{W}]\begin{bmatrix} \boldsymbol{h}_1 \\ \boldsymbol{h}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{G}_1'\boldsymbol{e} \\ \boldsymbol{G}_2'\boldsymbol{e} \end{bmatrix} + \lambda\begin{bmatrix} \boldsymbol{1} \\ \boldsymbol{0} \end{bmatrix}, \tag{4.8}$$

where the $2K \times (N_1 + N_2)$ matrix $[\boldsymbol{W}]$ is defined as

$$[\boldsymbol{W}] := \begin{bmatrix} \boldsymbol{G}_1'\boldsymbol{G}_1 + \tilde{\boldsymbol{G}}_1'\tilde{\boldsymbol{G}}_1 & \boldsymbol{G}_1'\boldsymbol{G}_2 - \tilde{\boldsymbol{G}}_1'\tilde{\boldsymbol{G}}_2 \\ \boldsymbol{G}_2'\boldsymbol{G}_1 - \tilde{\boldsymbol{G}}_2'\tilde{\boldsymbol{G}}_1 & \boldsymbol{G}_2'\boldsymbol{G}_2 + \tilde{\boldsymbol{G}}_2'\tilde{\boldsymbol{G}}_2 \end{bmatrix}. \tag{4.9}$$

This system is solved by inverting $[\boldsymbol{W}]$. The general solution can be expressed as

$$\begin{bmatrix} \boldsymbol{h}_1 \\ \boldsymbol{h}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \end{bmatrix} + \lambda\begin{bmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \end{bmatrix}, \tag{4.10}$$

where

$$\begin{bmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \end{bmatrix} := [\boldsymbol{W}]^{-1}\begin{bmatrix} \boldsymbol{G}_1'\boldsymbol{e} \\ \boldsymbol{G}_2'\boldsymbol{e} \end{bmatrix} \quad \text{and}$$

$$\begin{bmatrix} \boldsymbol{v}_1 \\ \boldsymbol{v}_2 \end{bmatrix} := [\boldsymbol{W}]^{-1}\begin{bmatrix} \boldsymbol{1} \\ \boldsymbol{0} \end{bmatrix}, \tag{4.11}$$

and where λ is chosen to satisfy the zero-bias constraint (4.1):

$$\lambda = \frac{2\alpha - \boldsymbol{1}'\boldsymbol{u}_1}{\boldsymbol{1}'\boldsymbol{v}_1}. \tag{4.12}$$

As before, we note that the solution (4.10) is expressed as the sum of two terms in which the first represents the unconstrained least-squares solution

($\boldsymbol{u}_1$ and $\boldsymbol{u}_2$), and the second is a correction term used to impose the constraint of a normalized impulse response.

5. Results

In this section, we apply the design methods described above to the problems of cubic spline interpolation [10, 21], and the decomposition of a signal using the cubic spline wavelet transform [2, 22]. We also compare the efficiency of these techniques to the standard design method that uses a truncation of the theoretical impulse responses. A tutorial overview of polynomial spline representations in general, and a discussion of their relevance to signal processing can be found in [20]. Finally, we briefly discuss a slight variation of our procedure for the design of modulated filter banks which are better suited for subband coding.

5.1. Cubic spline transform

The cubic spline interpolation of a signal $f(k)$ can be expressed as a weighted sum of cubic B-spline coefficients [18, 21]:

$$f(x) = \sum_{k=-\infty}^{+\infty} c(k)\varphi(x-k), \tag{5.1}$$

where $\varphi(x)$ is Schoenberg's cubic B-spline function

$$\varphi(x) = \beta^3(x) := \begin{cases} 2/3 - x^2 + |x|^3/2, & 0 \leqslant |x| < 1, \\ (2-|x|)^3/6, & 1 \leqslant |x| < 2, \\ 0, & 2 \leqslant |x|. \end{cases} \tag{5.2}$$

Both $f(x)$ and $\varphi(x)$ are continuous piecewise cubic functions with continuous first- and second-order derivatives (polynomial splines of order $n = 3$). In [21], it was shown that the cubic spline interpolation problem could be solved by direct B-spline filtering:

$$c(k) = (b_1^3)^{-1} * f(k), \tag{5.3}$$

where $(b_1^3)^{-1}$ is the convolution inverse of a sampled cubic spline kernel (indirect B-spline filter):

$$b_1^3(k) := \beta^3(x)|_{x=k} \quad \xleftrightarrow{z} \quad B_1^3(z) = \frac{z + 4 + z^{-1}}{6}. \tag{5.4}$$

We have approximated the IIR filter $(b_1^3)^{-1}$ by a series of symmetrical FIR filters of increasing length using the different design techniques. The corresponding reconstruction errors and biases in percent ($\{100 \times \varepsilon_0, 100 \times \text{bias}_0\}$) are given in Table 1. The standard technique is to simply truncate the impulse response; the impulse response is computed numerically by sampling the theoretical frequency response ($1/B_1^3(e^{j2\pi f})$) and using a 64-point inverse FFT. As expected, the unconstrained least-squares solution ($\lambda = 0$) gives rise to the smallest reconstruction error. The bias as well is significantly reduced when compared to the truncated impulse response design (TIRD). The constrained

Table 1
Performance comparison (reconstruction error and bias in percent) of different design techniques for the symmetrical FIR approximation of the inverse cubic spline filter as a function of N

N	Truncated impulse response (TIRD)	Least squares (LSD)	Constrained least squares (CLSD)
3	{11.32, 19.62}	{9.667, 10.28}	{10.99, 0.0}
5	{3.03, 5.26}	{2.62, 2.81}	{2.86, 0.0}
7	{0.813, 1.41}	{0.702, 0.754}	{0.752, 0.0}
9	{0.218, 0.377}	{0.188, 0.202}	{0.199, 0.0}
11	{0.058, 0.101}	{0.050, 0.054}	{0.053, 0.0}
13	{0.016, 0.027}	{0.014, 0.015}	{0.014, 0.0}

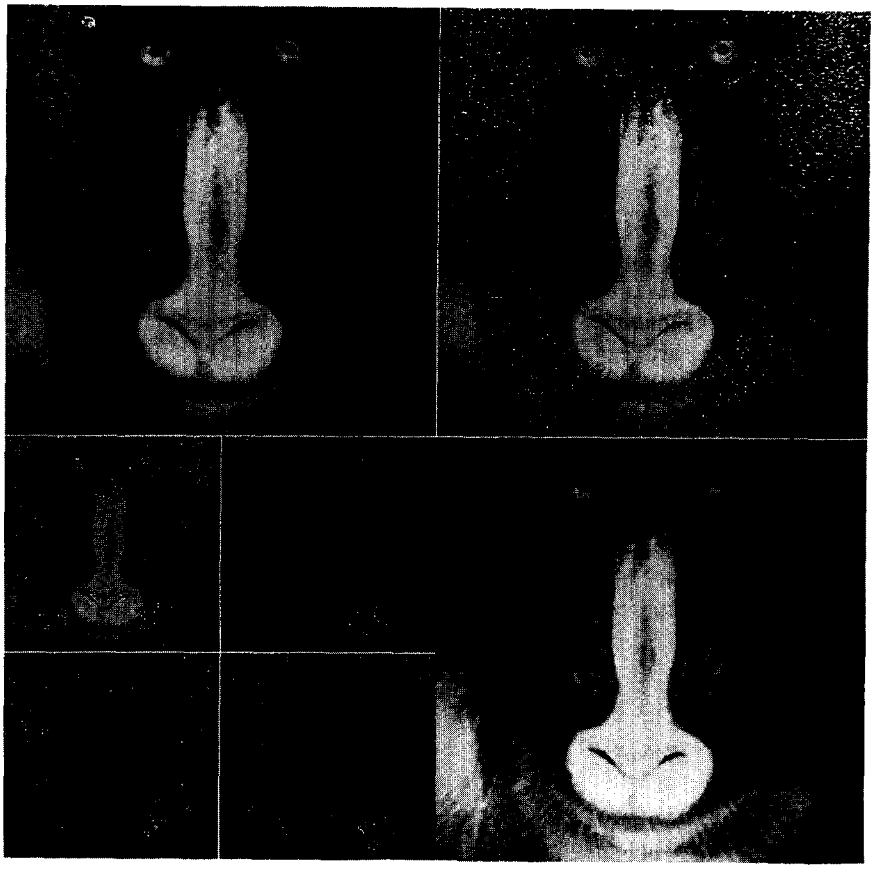

Fig. 3. Experimental example: (a) 256 × 256 test image, (b) cubic spline transform, (c) first iteration of the cubic B-spline wavelet transform, (d) reconstructed image by indirect B-spline wavelet transform and indirect cubic spline transform.

least-squares design (CLSD) leads to a reconstruction error that is between that of the TIRD and LSD but has the advantage of a zero bias. As an example, we give the computed transfer function of an 11-point CLSD approximation of the inverse cubic spline filter:

$$\begin{aligned} H(z) \cong\ & 1.73209 - 0.46405[z + z^{-1}] \\ & + 0.124384[z^{-2} + z^{-2}] \\ & - 0.0332243[z^{3} + z^{-3}] \\ & + 0.00883099[z^{4} + z^{-4}] \\ & - 0.0019876[z^{5} + z^{-5}]. \end{aligned}$$

This approximation has no bias and an error that is less than 1/1000. For two-dimensional signals, this filter is applied successively along the rows and columns to provide an image of the bicubic spline coefficients (direct cubic spline transform). This process is illustrated with the example in Fig. 3. The filtered image, which represents the cubic spline coefficients, is displayed in Fig. 3(b). The initial image (Fig. 3(a)) can be recovered with no noticeable loss by convolving with the operator defined by (5.4) (indirect cubic spline transform). These algorithms were coded in FORTRAN on a low-end workstation (Macintosh II fx) for biomedical image processing applications. The FIR implementation is computationally quite efficient. Typically, the direct cubic spline filtering of a 256 × 256 image is performed in fewer than 5 s and the indirect filtering in fewer than 3 s.

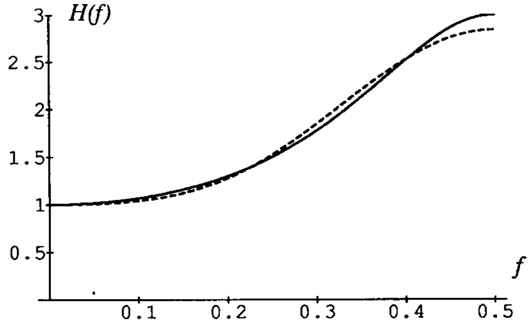

Fig. 4. Frequency responses of the direct cubic spline filter (solid line) and its 5-points CLSD approximation (dashed line).

The frequency responses of the direct cubic spline filter and its 5-point FIR CLSD approximations are given in Fig. 4. The graph associated with the 11-point approximation given above is virtually undistinguishable from the theoretical curve. The direct cubic spline filter $(b_1^3)^{-1}$ tends to accentuate higher frequencies which explains why the cubic spline coefficient image (Fig. 3(b)) is a sharpened version of the original.

5.2. Cubic B-spline wavelet transform

As shown in [22], the cubic spline function $f(x)$ given by (5.1) can be decomposed further as

$$f(x) = \sum_{k=-\infty}^{+\infty} c_{(1)}(k)\varphi(x/2-k) + \sum_{k=-\infty}^{+\infty} d_{(1)}(k)\psi(x/2-k), \qquad (5.5)$$

in which the first term of the sum corresponds to a coarser resolution cubic spline approximation and the second is a wavelet representation of the residual error, as defined in [12, 13]. The $d_{(1)}$'s are the cubic spline wavelet coefficients at resolution level (1). The basis function $\varphi(x/2)$ is the cubic B-spline expanded by a factor of two. It can be represented as a weighted sum of basic B-spline functions:

$$\varphi(x/2) = \sum_{k=-\infty}^{+\infty} g_1(k)\varphi(x-k), \qquad (5.6)$$

where $g_1(k) \overset{z}{\longleftrightarrow} \frac{1}{8}(z^2+4z+6+4z^{-1}+z^{-2})$. The function $\psi(x/2)$ is the cubic spline wavelet of compact support and is defined as in [22]:

$$\psi(x/2) = \sum_{k=-\infty}^{+\infty} g_2(k+1)\varphi(x-k), \qquad (5.7)$$

where

$$g_2(k) \overset{z}{\longleftrightarrow} \left(\frac{6-4[z+z^{-1}]+[z^2+z^{-2}]}{8}\right) \times \left(\frac{2416-1191[z+z^{-1}]+120[z^2+z^{-2}]-[z^3+z^{-3}]}{5040}\right).$$

This function satisfies the orthogonality condition: $\psi(x/2) \perp \varphi(x/2-k), k\in Z$. We can replace the basis functions in (5.5) by their explicit formulas (5.6) and (5.7), respectively. If we then equate the corresponding expression and the B-spline expansion (5.1), we obtain the reconstruction algorithm:

$$c_{(0)}(k) := c(k) = g_1 * [c_{(1)}]_{\uparrow 2}(k) + \delta_{-1} * g_2 * [d_{(1)}]_{\uparrow 2}(k), \qquad (5.8)$$

where $\delta_i(k)$ is the convolution shift operator, i.e., $\delta_i * a(k) = a(k-i)$, and where the up-sampling operator is defined by (1.4). It is the operation performed by the right side of block diagram in Fig. 2. The problem is to determine FIR approximations of the wavelet filters h_1 and h_2 that perform the decomposition

$$c_{(1)}(k) = [h_1 * c_{(0)}]_{\downarrow 2}(k),$$
$$d_{(1)}(k) = [\delta_1 * h_2 * c_{(0)}]_{\downarrow 2}(k), \qquad (5.9)$$

according to the procedure illustrated by the left side of the diagram in Fig. 2. This decomposition can be applied iteratively on the coarser resolution part of the approximation to yield a complete wavelet decomposition. The transfer functions of these filters can be determined by substituting the z-transforms of $g_1(k)$ and $g_2(k)$ into (2.7) and (2.8); more details can be found in [23].

We have approximated these operators by a series of symmetrical FIR filters of increasing length ($N = N_1 = N_2$) using the different design techniques. The corresponding reconstruction errors and biases in percent ($\{100\times\varepsilon_1, 100\times\varepsilon_2, 100\times \mathrm{bias}_1\}$) are given in Table 2. Although the reduction of the different errors as a function of N is

Table 2
Performance comparison (distortion, aliasing error, and bias in percent) of different design techniques for the FIR approximations of the analysis filters for the cubic B-spline wavelet as a function of N

N	Truncated impulse response (TIRD)	Least squares (LSD)	Constrained least squares (CLSD)
3	{65.14, 31.97, 69.45}	{45.03, 10.54, 34.10}	{47.87, 10.42, 0.0}
7	{33.01, 15.29, 33.57}	{22.85, 10.06, 16.78}	{23.51, 10.18, 0.0}
11	{17.58, 8.16, 17.53}	{11.88, 6.56, 8.82}	{12.11, 6.63, 0.0}
15	{9.4, 4.39, 9.32}	{6.28, 3.77, 4.71}	{6.36, 3.81, 0.0}
19	{5.02, 2.4, 4.96}	{3.34, 2.07, 2.52}	{3.38, 2.09, 0.0}
23	{2.66, 1.38, 2.62}	{1.79, 1.12, 1.35}	{1.80, 1.113, 0.0}
27	{1.39, 0.93, 1.33}	{0.95, 0.60, 0.72}	{0.96, 0.60, 0.0}
31	{0.69, 0.87, 0.58}	{0.51, 0.32, 0.39}	{0.51, 0.32, 0.0}

Table 3
FIR filter coefficients for the cubic B-spline wavelet transform (the kernels g_1 and g_2 are exact and h_1 and h_2 were obtained via CLSD with $N = 27$)

k	h_1	h_2	g_1	g_2
0	+ 0.892995	+ 1.47401	+ 0.75	+ 0.601786
− 1, 1	+ 0.400474	− 0.468232	+ 0.5	− 0.458383
− 2, 2	− 0.282547	− 0.740512	+ 0.125	+ 0.196032
− 3, 3	− 0.233318	+ 0.345154		− 0.0415923
− 4, 4	+ 0.128883	+ 0.387516		+ 0.0030754
− 5, 5	+ 0.12641	− 0.195611		− 0.000024802
− 6, 6	− 0.0666382	− 0.204225		
− 7, 7	− 0.0683554	+ 0.104778		
− 8, 8	+ 0.0346756	+ 0.105509		
− 9, 9	+ 0.0360809	− 0.0541513		
− 10, 10	− 0.0181137	− 0.050628		
− 11, 11	− 0.0189224	+ 0.0258674		
− 12, 12	+ 0.0072949	+ 0.0185111		
− 13, 13	+ 0.00757909	− 0.0094291		

not as rapid as in the previous example, the same qualitative observations can be made. The performance of CLSD is only very slightly suboptimal. Since it has no bias, this approach should definitely be recommended for image processing applications. The detailed results of this constrained optimization for $N = 27$ are given in Table 3.

The corresponding symmetrical FIR convolution operators permit the implementation of a cubic wavelet transform with a reconstruction error that is of the order of 1%, a level of precision that should be sufficient for most applications. In two dimensions, these operators should be applied successively along the rows and columns of the data. In this case, there are four types of basis functions corresponding to the different cross-products between the scaling function φ and the wavelet ψ in the spatial variables x and y. This decomposition procedure is illustrated in Fig. 3(c), which was obtained from Fig. 3(b) by filtering and decimation using the FIR kernels in Table 3. The corresponding CPU time was of the order of 13 s. The upper left quadrant in Fig. 3(c) represents the cubic B-spline coefficients at the coarser resolution level, while the three other quadrants contain the three different types of wavelet coefficients (vertical, horizontal, and diagonal). This procedure can be re-applied iteratively on the lowpass component of the image (upper left corner). The B-spline coefficients at the finer resolution level were then recovered by indirect transformation using the operators g_1 and g_2. The initial image was finally reconstructed by indirect B-spline transform (convolution with b_1^3); it is displayed in Fig. 3(d). The final reconstruction error in this example was less than 0.3% (SNR = 45.43 dB), which was not anticipated given the fact that FIR approximations have been used for both the direct cubic spline and B-spline wavelet transforms and that all intermediate filtering results were truncated and stored in small integer format. The fact that this experimental error is much smaller than the error specifications (ε_1 and ε_2) in Table 2 can be explained by the fact that the filter design is based on a worst-case scenario in which the input of the system is an impulse; this latter situation is also equivalent to a white-noise excitation.

The main feature of the example considered immediately above is that the underlying scaling functions and wavelets are very well localized in time and frequency [22]. These properties together with the fast algorithm also described above should make this decomposition a useful tool for the analysis of a variety of non-stationary signals.

5.3. Design of modulated filter banks

In the context of subband coding, the bandpass characteristics of the filter bank is one of the most determinant factors [11, 28]. For this reason, we do not recommend the use of the previous decomposition in such applications. Rather, we would like to have a system that performs a near perfect subband decomposition. One approach to obtaining such a filter bank is to start by designing a good half-band FIR filter $g_1(k)$ using any standard procedure [17]. The corresponding highpass filter is then obtained by simple modulation

$$g_2(k) = (-1)^k g_1(k). \tag{5.10}$$

We can then use our least-squares design procedure to determine the corresponding analysis (or synthesis) filters $h_1(k)$ and $h_2(k)$. In this case, it would be judicious to use a slightly modified constraint:

$$\mathbf{1}'\boldsymbol{h}_1 + \mathbf{1}'\boldsymbol{h}_2 = \alpha, \tag{5.11}$$

where α is chosen in order to guarantee perfect reconstruction of the DC component. This simple modification of the algorithm ensures that the least-squares filters $h_1(k)$ and $h_2(k)$ are also modulated versions of each other. Moreover, by using the fact that $\boldsymbol{G}_2 = \tilde{\boldsymbol{G}}_1$, we can make use of symmetries and reduce the total complexity of the procedure by a factor of two.

6. Conclusion

In this paper, we have studied a class of FIR approximation techniques for the inversion of FIR convolution operators. The search for an exact solution would require the implementation of infinite impulse response operators. Although recursive algorithms have been developed for this purpose (see, for example, [21]), the IIR approach may not always be practical. The main advantage of FIR filters is their simplicity and the fact that they do not require floating point storage; they are also immune to the propagation of roundoff errors. In addition, FIR algorithms are better suited for parallel processing.Their principal disadvantage is that they give rise to truncation errors. However, an error of the order of 1% is usually acceptable for most images.

Our design technique is based on the constrained minimization of the inversion (or reconstruction) error associated with an impulse sequence. The corresponding algorithms involve standard matrix algebra and are simple to implement. An important built-in feature is the precise recovery of the DC signal component (zero-bias constraint). This last property may result in a substantial performance improvement, especially when the DC component accounts for a large proportion of the total signal energy, as is typically the case for digital images.

The main use of the one-channel algorithm is for the design of FIR inverse (or deconvolution) filters suitable for signal restoration. The two-channel algorithm is in fact an algebraic least-squares formulation of the perfect reconstruction filter bank problem. These procedures are also directly applicable to the design of fast algorithms for continuous signal transforms using shifted basis functions of compact support (e.g. B-spline transform and wavelet transform), as illustrated by our examples. Finally, we note that the present computational techniques can be extended for the design of multirate filter banks with more than two channels. In fact, it should be possible to derive a direct algebraic solution for the general N channel case considered in [15].

7. References

[1] E.H. Adelson, E. Simoncelli and R. Hingorani, "Orthogonal pyramid transforms for image coding", *Proc. SPIE Conf. Visual Communication and Image Processing*, Cambridge, MA, October 1987, pp. 50–58.

[2] C.K. Chui and J.Z. Wang, "On compactly supported spline wavelets and a duality principle", *Trans. Amer. Math. Soc.*, Vol. 330, No. 2, 1992, pp. 903–915.

[3] A. Croisier, D. Esteban and C. Galand, "Perfect channel splitting by use of interpolation, decimation, tree decomposition techniques", *Proc. Internat. Conf. on Information Sciences/Systems*, Patras, August 1976, pp. 443–446.

[4] I. Daubechies, "Orthogonal bases of compactly supported wavelets", *Comm. Pure Appl. Math.*, Vol. 41, 1988, pp. 909–996.

[5] J.G. Daugman, "Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression", *IEEE Trans. Acoust. Speech Signal Process.*, Vol. ASSP-36, No. 7, July 1988, pp. 1169–1179.

[6] T. Ebrahimi, T.R. Reed and M. Kunt, "Video coding using a pyramidal Gabor expansion", *Proc. Visual Communications and Image Processing '90*, Lausanne, Switzerland, 1990, pp. 489–502.

[7] D. Gabor, "Theory of communication", *J. Inst. Elec. Eng.*, Vol. 93, No. III, 1946, pp. 429–457.

[8] F.A. Graybill, *Matrices with Application in Statistics*, Wadsworth, Belmont, CA, 1983.

[9] Z. Gülboy and N.C. Geçkinli, "On FIR filters having approximate FIR inverses with a specified LMS error", *Signal Processing*, Vol. 19, No. 1, January 1990, pp. 9–15.

[10] H.S. Hou and H.C. Andrews, "Cubic splines for image interpolation and digital filtering", *IEEE Trans. Acoust. Speech Signal Process.*, Vol. ASSP-26, No. 6, 1978, pp. 508–517.

[11] V.K. Jain and R.E. Crochiere, "Quadrature mirror filter design in the time domain", *IEEE Trans. Acoust. Speech Signal Process.*, Vol. ASSP-32, No. 2, April 1984, pp. 353–361.

[12] S.G. Mallat, "Multiresolution approximations and wavelet orthogonal bases of $L^2(R)$", *Trans. Amer. Math. Soc.*, Vol. 315, No. 1, 1989, pp. 69–87.

[13] S.G. Mallat, "A theory of multiresolution signal decomposition: The wavelet representation", *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. PAMI-11, No. 7, 1989, pp. 674–693.

[14] Y. Meyer, *Ondelettes et Opérateurs I: Ondelettes*, Hermann, Paris, 1990.

[15] B. Paillard, J. Soumagne, P. Mabilleau and S. Morisette, "Subband decomposition: An LMS-based algorithm to approximate the perfect reconstruction bank in the general case", *IEEE Trans. Signal Process.*, Vol. 39, No. 1, January 1991, pp. 233–238.

[16] J.A. Parker, R.V. Kenyon and D.E. Troxel, "Comparison of interpolating methods for image resampling", *IEEE Trans. Med. Imaging*, Vol. MI-2, No. 1, 1983, pp. 31–39.

[17] J.G. Proakis and D.G. Manolakis, *Introduction to Digital Signal Processing*, Macmillan, New York, 1990.

[18] I.J. Schoenberg, *Cardinal Spline Interpolation*, SIAM, Philadelphia, PA, 1973.

[19] E.P. Simoncelli and E.H. Adelson, "Non-separable extensions of quadrature mirror filters to multiple dimensions", *Proc. IEEE*, Vol. 78, No. 4, 1990, pp. 652–664.

[20] M. Unser and A. Aldroubi, "Polynomial splines and wavelets – A signal processing perspective", in: C.K. Chui, ed., *Wavelets – A Tutorial in Theory and Applications*, Academic Press, San Diego, 1992, pp. 91–122.

[21] M. Unser, A. Aldroubi and M. Eden, "Fast B-spline transforms for continuous image representation and interpolation", *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 13, No. 3, March 1991, pp. 277–285.

[22] M. Unser, A. Aldroubi and M. Eden, "On the asymptotic convergence of B-spline wavelets to Gabor functions", *IEEE Trans. Inform. Theory*, Vol. 38, No. 2, 1992, pp. 864–872.

[23] M. Unser, A. Aldroubi and M. Eden, "A family of polynomial spline wavelet transforms", *Signal Processing*, Vol. 30, No. 2, January 1993, pp. 141–162.

[24] P.P. Vaidyanathan, "Quadrature mirror filter banks, M-band extensions and perfect-reconstruction technique", *IEEE ASSP Mag.*, Vol. 4, July 1987, pp. 4–20.

[25] M. Vetterli, "Multi-dimensional sub-band coding: some theory and algorithms", *Signal Processing*, Vol.6, No. 2, April 1984, pp. 97–112.

[26] M. Vetterli, "A theory of multirate filter banks", *IEEE Trans. Acoust. Speech Signal Process.*, Vol. ASSP-35, No. 3, March 1987, pp. 356–372.

[27] M. Vetterli and C. Herley, "Wavelets and Filter Banks: Theory and Design", *IEEE Trans. Signal Process.*, Vol. SP-40, No. 9, September 1992, pp. 2207–2232.

[28] J.W. Woods and S.D. O'Neil, "Sub-band coding of images", *IEEE Trans. Acoust. Speech Signal Process.*, Vol. ASSP-34, October 1986, pp. 1278–1288.