MOMS: Maximal-Order Interpolation of Minimal Support

Thierry Blu, *Member, IEEE*, Philippe Thévenaz, and Michael Unser, *Fellow, IEEE*

***Abstract*—We consider the problem of interpolating a signal using a linear combination of shifted versions of a compactly-supported basis function $\varphi(x)$. We first give the expression of the φ's that have minimal support for a given accuracy (also known as "approximation order"). This class of functions, which we call maximal-order-minimal-support functions (MOMS) is made of linear combinations of the B-spline of same order and of its derivatives.**

We provide the explicit form of the MOMS that maximize the approximation accuracy when the step-size is small enough. We compute the sampling gain obtained by using these optimal basis functions over the splines of same order. We show that it is already substantial for small orders and that it further increases with the approximation order L. When L is large, this sampling gain becomes linear; more specifically, its exact asymptotic expression is $2/(\pi e)L$. Since the optimal functions are continuous, but not differentiable, for even orders, and even only piecewise continuous for odd orders, our result implies that regularity has little to do with approximating performance.

These theoretical findings are corroborated by experimental evidence that involves compounded rotations of images.

I. Introduction

THIS paper deals with the problem of finding a good interpolation model for fitting the uniform samples (not necessarily perfect) of a signal. The prevalence of interpolation in digital image processing stems from a basic inconsistency between the world of natural phenomena, modeled by continuously-defined variables, and the world of computers, where discrete—and in practice, finite—data reign. This is especially true of imaging where such models may include general transformations such as rotations or translations [1], data reduction and magnification [2], Cartesian-to-polar coordinate conversions [3]–[5], data reslicing or resampling [6]–[8], image warping [9], [10], gradient estimations [11], and many others. Common to all these operations is the need to access the value of a signal in between samples.

The Classical Approach to Interpolation: The usual technique for mapping a discrete signal f_n onto a continuously-defined signal $f_{\rm int}(x)$ is to express it on a basis made of shifts of a function $\varphi_{\rm int}(x)$, according to

$$f_{\rm int}(x) = \sum_n f_n \varphi_{\rm int}\left(\frac{x}{T} - n\right). \tag{1}$$

$\varphi_{\rm int}$ is chosen so as to satisfy the *interpolation condition* $\varphi_{\rm int}(k) = \delta_k$, where δ_k is Kronecker-delta; this ensures that $f_{\rm int}(kT) = f_k$. The parameter T that appears in this formulation is a sampling step—or the inverse of a sampling frequency. As this step-size gets smaller, it is usually required that the interpolated function $f_{\rm int}(x)$ gets closer to some function $f(x)$ that should *ideally* represent the samples f_n.

The Shannon-Whittaker theory tells us that the choice $\varphi_{\rm int} = \mathrm{sinc}$ and $T \leq 1/F$ yields the exact reconstruction $f_{\rm int} = f$, provided that the bandwidth of this ideal function $f(x)$ is limited to some *finite* interval $[-F/2, F/2]$. Unfortunately, life is not ideal. Band-limited functions do not exist in practice, if only because no real signal is of infinite length. Moreover, the cardinal sinc function is not amenable to practical applications, because of its infinite support, the slow convergence of a sum of shifted sinc's, and correlatively, a strong instability in the presence of noise.

Practitioners have long considered *finite-support* interpolators $\varphi_{\rm int}$. They have designed them to be close to the sinc, with the hope that this would bring good approximation characteristics [12]–[14]. Others observed that what we will call "approximation order" improves quality [15]–[17]. An example is that of Keys' cubic kernel which has a free parameter; its optimization turns out to be the one that provides the highest approximation order [18]. In order to provide computationally efficient algorithms, these kernels were chosen to be piecewise-polynomial and of short support. Indeed, the size of the support is the most crucial single element that rules the computational complexity of an interpolation algorithm. This is because the evaluation of $f_{\rm int}$ at some point x requires the computation of S terms in (1), if S denotes the length of the support of $\varphi_{\rm int}$. Furthermore, this number increases exponentially with the dimension of the problem.

The weaknesses of this approach are twofold. Regarding the computational cost, the "small support + interpolation" property appears much too restrictive, given that infinite support can be implemented at almost no additional computational cost [19]. Regarding the approximation quality, it is not ensured that "closeness" to the sinc is truly beneficial. To be more specific, the problem lies in *how* to measure this closeness. We have been investigating these issues in some depth [20], [21] and have proposed satisfactory solutions that are summarized in the following paragraphs.

Our Approach to Interpolation [22], [23]: A fast implementation of (1) can be devised, provided that $\varphi_{\rm int}(x)$ itself can be expressed as an—arbitrary—linear combination of shifted versions of a finite-support function $\varphi(x)$. A consequence of this change of basis is that (1) becomes

$$f_{\rm int}(x) = \sum_n c_n \varphi\left(\frac{x}{T} - n\right) \tag{2}$$

Manuscript received February 29, 2000; revised March 30, 2001. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Brian L. Evans.

The authors are with the Biomedical Imaging Group, Swiss Federal Institute of Technology Lausanne, CH-1015 Lausanne EPFL, Switzerland (e-mail: thierry.blu@epfl.ch; philippe.thevenaz@epfl.ch; michael.unser@epfl.ch).

Publisher Item Identifier S 1057-7149(01)05442-2.

1057–7149/01$10.00 ©2001 IEEE

where $c_n \neq f_n$, in general. Yet, we still require consistency between the samples of the interpolated function $f_{\text{int}}(kT)$, and the measures f_k. This implies that $\sum_n c_n \varphi(k-n) = f_k$ for every integer k. Thus, the interpolation method consists of a two-step algorithm.

- The first step involves prefiltering the discrete data f_n by the low-order all-pole filter $h(z) = (\sum_n \varphi(n) z^{-n})^{-1}$, which provides $c_n = h_n * f_n$. This apparently unstable filter can be implemented very efficiently by factorization into a causal part h_c and anticausal part h_a; the f_n are first filtered by h_c in the *forward* direction, and then filtered *backward* by h_a. This results in a fast and stable algorithm.
- The second step is similar to (1), with the difference that we use (2) instead. Obviously, the cost of this step is directly given by the size of the support of $\varphi(x)$. The smaller the support, the more efficient the interpolation algorithm.

Numerous experiments [19], [23], [24] have shown that the additional cost of prefiltering is almost negligible compared to the second step. We thus claim that it is unnecessary—and detrimental to quality—to require that the support of the interpolating function φ_{int} be of finite size, or, equivalently, that the auxiliary function φ in (2) satisfy the interpolation condition. This means that the choice of the function $\varphi(x)$ is essentially *free*, except that, for computational efficiency purposes, we require that its support be as small as possible. Of course, the matching between the interpolated data and the measures is still taken into account—in the definition of the prefilter.

Approximation Issues: The interpolation problem can also be reformulated in the following way: find the function f_{int} in

$$\mathcal{V}_T = \operatorname*{span}_{n \in \mathbb{Z}} \left\{ \varphi\left(\frac{x}{T} - n\right) \right\} \tag{3}$$

such that $f_{\text{int}}(kT) = f_k$. At this stage, one must realize that interpolation is not the only method for approximating in $\mathcal{V}_T$ an arbitrary function $f(x)$; and depending on how we measure the closeness between two functions, it might not be the optimal method. For example, if we choose the $\mathbf{L}^2$ norm, then the optimal approximation of $f \in \mathbf{L}^2$ within $\mathcal{V}_T$ is $\mathcal{P}_T f$, the orthogonal projection of f onto $\mathcal{V}_T$, which is in general distinct from the solution provided by the interpolation method. It is only when the function is Nyquist-band-limited and when $\varphi = \text{sinc}$ that interpolation and orthogonal projection agree. A fair evaluation of the $\mathbf{L}^2$ approximation quality of φ, i.e., of $\mathcal{V}_T$, must thus consider the distance of f to $\mathcal{V}_T$, that is to say, the orthogonal projection error $\|f - \mathcal{P}_T f\|_{\mathbf{L}^2}$.

A rough evaluation of the approximation quality is given by the rate of decrease of the approximation error as $T \to 0$, also known as "approximation order" in approximation theory [25]. In mathematical terms, the property $\|f - \mathcal{P}_T f\|_{\mathbf{L}^2} \propto T^L$ constitutes a definition of L as the approximation order of φ. As can be expected, the approximation order of both the interpolation and the least-squares approximation methods coincide. A useful example is that of splines of degree n which have order $L = n + 1$.

A. Results and Organization of the Paper

We first give a short account of approximation theory and of our previous contributions to this matter. In particular, we introduce the Fourier approximation kernel as the fundamental quantity for studying the $\mathbf{L}^2$ approximation error (see Section II).

In wavelet theory, the order of approximation is identified—and often, hidden—as the *regularity order*; i.e., the number of regularity factors $(1 + z^{-1})$ that divide the scaling filter [26]. Since noncentered splines of degree $L-1$ are known to be the smallest scaling functions that have *regularity* order L, one may wonder whether these splines are also the smallest functions that have *approximation* order L. The answer is yes, but it turns out that the splines are not the only functions that reach this minimum. The full class of such functions, which we call maximal-order-minimal-support functions (MOMS), is characterized in this paper (see Section III).

A finer estimation of the approximation error takes into account the proportionality constant between $\|f - \mathcal{P}_T f\|_{\mathbf{L}^2}$ and $T^L \|f^L\|_{\mathbf{L}^2}$ as $T \to 0$. Unser [27] found a simple expression for this asymptotic approximation constant. Here, we minimize this expression within the MOMS class, yielding new functions which we call O-MOMS (see Section IV). In other words, among all the φ's that are compactly supported in $[0, L]$, the O-MOMS of order L is the one that yields the smallest asymptotic approximation error $\|f - \mathcal{P}_T f\|_{\mathbf{L}^2}$, *whatever the function* f. We also show how to design other kinds of MOMS kernels: suboptimal MOMS (SO-MOMS) and Lagrange-type interpolators "I-MOMS" (see Section V).

Our theoretical claims (influence of the approximation order, optimality of the O-MOMS) are consistently confirmed by practical experiments (see Section VI). There, we rate the interpolation behavior of several MOMS kernels.

These results point out that, belying a widespread opinion, it is definitely not necessary for the approximating function to be regular to achieve a good approximation scheme; on the contrary, we show that the *best* $\mathbf{L}^2$ approximation of a function is not more than continuous. Depending on the evenness of the approximation order, it may in fact have points of discontinuity. Even used in a suboptimal method—interpolation—the O-MOMS perform significantly better than any other kernel of the same size [23]. A nontrivial by-product of our findings is that the best kernels (used in interpolation or in orthogonal projection) are piecewise-polynomial and can be expressed using derivatives of a spline of the same degree.

B. Notation

When not otherwise stated, we often omit the range of integer values for infinite summations, as well as the range of real values for integrals. Thus, $\sum_k$ should be understood as $\sum_{k \in \mathbb{Z}}$ and $\int f(x)dx$ should be understood as $\int_{-\infty}^{\infty} f(x)dx$.

The conventional inner product $\int s_1(x) s_2(x) dx$ between two $\mathbf{L}^2$ functions s_1, s_2, is denoted $\langle s_1, s_2 \rangle$, and the associated Euclidean norm is $\|.\|_{\mathbf{L}^2}$.

The Fourier transform of $s(x)$ is $\hat{s}(\omega) = \int s(x) \mathrm{e}^{-j\omega x} dx$. Let r be a positive real number; the Sobolev space $\mathbf{W}_2^r$ is defined as the collection of functions satisfying $\int (1 + \omega^2)^r |\hat{s}(\omega)|^2 d\omega < \infty$. By analogy to this definition of regularity, we extend $\|s^{(r)}\|_{\mathbf{L}^2}$ to noninteger values of r by equating it to $\left(\int |\omega|^{2r} |\hat{s}(\omega)|^2 (d\omega/2\pi)\right)^{1/2}$. The smoothness of a function $s(x)$ can thus be characterized by the maximum r

such that $s \in \mathbf{W}_2^r$; this regularity exponent $r_{\max}$ indicates that $s(x)$ has $\lfloor r \rfloor$ derivatives in $\mathbf{L}^2$ for all $r < r_{\max}$. There is also a direct connection with *point-wise* smoothness: if $s \in \mathbf{W}_2^r$ with $r > 1/2$, then $s(x)$ has at least $\lfloor r - 1/2 \rfloor$ continuous derivatives [28].

The Riemann zeta function is defined as $\zeta(s) = \sum_{n \geq 1} n^{-s}$, which is convergent for $\Re(s) > 1$.

Most of the asymptotic expansions are presented with "$o(\cdot)$" and "$O(\cdot)$" terms: writing $f(x) = o(x^n)$ is equivalent to writing $\limsup_{x \to 0} |f(x)/x^n| = 0$. In the same spirit, writing $f(x) = O(x^n)$ is equivalent to writing $\limsup_{x \to 0} |f(x)/x^n| < \infty$ (i.e., not necessarily 0).

Noncentered B-splines of degree n, denoted $\beta^n(x)$, are piecewise-polynomial functions that are compactly supported within $[0, n+1]$. Their exact expression is given by the formula

$$\beta^n(x) = \sum_{k=0}^{n+1} (-1)^k \binom{L}{k} \frac{(x-k)_+^n}{(n)!}$$

where, by definition, x_+^n is the one-sided power function $\max(0, x)^n$. Their Fourier transform takes the simpler form

$$\hat{\beta}^n(\omega) = \left(\frac{1 - \mathrm{e}^{-j\omega}}{j\omega} \right)^{n+1}.$$

II. Approximation Using Functions Shifted by Integers

A. Approximation Order and Strang-Fix Theory

The notion of approximation order is crucial in approximation theory since it governs the rate of decrease of the approximation error as $T \to 0$. Specifically, the approximation order is defined as the exponent L such that the difference between the function $f(x)$ and its orthogonal projection $f_T(x)$ onto $\mathcal{V}_T$ tends to 0 with T^L; i.e., $\|f - f_T\|_{\mathbf{L}^2} \leq \text{const} \times T^L$. For this property to hold, it is necessary to assume that $f(x)$ and its Lth derivative belong to $\mathbf{L}^2$.

In the case where $\mathcal{V}_T$ is the space of π/T-band-limited functions, then $\varphi(x) = \text{sinc}(x)$ and it turns out that $\|f - f_T\|_{\mathbf{L}^2} \leq \text{const} \times T^r$, where r is any positive number smaller than the Sobolev regularity order of the function $f(x)$ [29]. Moreover, if $f(x)$ is indefinitely differentiable—this is the case of any band-limited function—then the convergence of $f_T(x)$ to $f(x)$ is faster than any monomial T^L.

Here instead, we assume that $\varphi(x)$ is compactly supported. Thus, the decrease rate of the approximation error cannot be infinite and is necessarily integer. More specifically, we will see (in Section III) that the support of φ must be at least of length L, for $\|f - f_T\|_{\mathbf{L}^2}$ to decrease with T^L.

To check the approximation order of $\mathcal{V}_T$, Strang and Fix established in 1974 [25] the equivalence between Lth order of approximation and the following conditions:

$$\begin{cases} \hat{\varphi}(0) = 1 \\ \forall\, n \in \mathbb{Z}^*, \quad \hat{\varphi}(2n\pi + \omega) = O(\omega^L). \end{cases} \tag{4}$$

The Strang-Fix conditions (4) can also take the following form [30, Proposition 4.4]:

$$\hat{\varphi}(0) = 1 \text{ and } \begin{cases} \forall\, n = 0 \ldots L-1, \quad \exists\, C_n \in \mathbb{R} \\ \text{such that} \\ \sum_k (x-k)^n \varphi(x-k) = C_n \text{ a.e.} \end{cases} \tag{5}$$

This condition implies that any polynomial up to degree $L-1$ belongs to $\mathcal{V}_T$. Yet another equivalent form of (4) is

$$\hat{\varphi}(0) = 1 \text{ and } \begin{cases} \forall\, P \text{ polynomial of degree} \leq L-1 \\ \exists\, C_p \in \mathbb{R} \text{ such that} \\ \sum_k P(x-k)\varphi(x-k) = C_P \text{ a.e.} \end{cases} \tag{6}$$

In the rest of this paper, we will use the latter condition as an equivalent formulation of the approximation order.

B. Approximation Kernel and Asymptotic Constant

When the sampling step T is not small, or, equivalently, when the spectrum of the function $f(x)$ has a significant content at high frequencies, the approximation error has to be characterized not only by a (integer) number, but by a function: the Fourier approximation kernel $E(\omega)$. When the approximation method is the orthogonal projection onto $\mathcal{V}_T$, the expression of this kernel is

$$E(\omega) = 1 - \hat{\overset{\circ}{\varphi}}(\omega)^* \hat{\varphi}(\omega). \tag{7}$$

We have recently shown [20], [21] that

$$\|f - f_T\|_{\mathbf{L}^2} = \sqrt{\int |\hat{f}(\omega)|^2 E(\omega T) \frac{d\omega}{2\pi}} + \rho_f(T) \tag{8}$$

where $\rho_f(T) = o(T^r)$ if $f \in \mathbf{W}_2^r$. More remarkable, this term cancels on the average over all possible shifts of the function; i.e., if we denote $f^{[\tau]}(x) = f(x - \tau)$, then

$$\lim_{N \to \infty} \frac{1}{2N} \int_{-N}^{N} \|f^{[\tau]} - (f^{[\tau]})_T\|_{\mathbf{L}^2}^2 d\tau = \int |\hat{f}(\omega)|^2 E(\omega T) \frac{d\omega}{2\pi}.$$

Finally, we also know that the correcting term $\rho_f(T)$ in (8) vanishes when $f(x)$ satifies Nyquist's band-limitation constraint.

The approximation efficiency of a function space $\mathcal{V}_T$ generated by shifts of $\varphi(x)$ is thus directly given by the closeness of $E(\omega T)$ to 0 in the frequency region where the functions to approximate are prominent; e.g., in $[-\pi/T, \pi/T]$ for functions that comply with Nyquist sampling hypothesis.

When $\varphi(x)$ satisfies Strang-Fix conditions of order L, it is easy, using (8), to find the asymptotic equivalent of the approximation error as $T \to 0$. The result takes the form

$$\|f - f_T\|_{\mathbf{L}^2} = C_\varphi^- \|f^{(L)}\|_{\mathbf{L}^2} T^L + o(T^L) \tag{9}$$

where C_φ^- is given by [27]

$$C_\varphi^- = \frac{1}{L!} \left(\sum_{n \neq 0} |\hat{\varphi}^{(L)}(2n\pi)|^2 \right)^{1/2}. \tag{10}$$

In image-processing applications, it is more frequent for the sampling step T to be fixed than to be tunable, perhaps made to decrease toward zero. However, it can easily be checked that the behavior of the approximation of smooth functions when $T \to 0$ is also that of low-pass functions of decreasing bandwidth when the step-size is kept constant. Thus, for very low-pass signals—typically, images—the asymptotic constant provides quite a reliable quality measure. This justifies the approach chosen in Section IV where we shall minimize this constant. Our choice will be further strengthened by the empirical finding that the function $\varphi(x)$ that minimizes the asymptotic constant also tends to minimize the approximation Fourier kernel $E(\omega)$ over the full Nyquist spectral range.

III. Minimal-Support Kernels: MOMS

We now answer a question of consequence: Which are the smallest-support functions that have a given approximation order L? Minimizing the support is indeed essential, as it ensures better localization, and hence fewer large-scale artefacts for a given approximation accuracy. In addition, as pointed out in Section I, it is computationally more efficient to have a small support, especially in higher dimensions.[1]

Theorem 1: (MOMS) For a given approximation order L, the smallest-support kernel $\varphi(x)$ is piecewise-polynomial of degree $L-1$ and its support is of size L. Moreover, the full class of these minimum-support functions is contained within an L-dimensional vector space parametrized as

$$\varphi(x) = \sum_{n=0}^{L-1} \lambda_n \frac{d^n}{dx^n} \beta^{L-1}(x-a) \tag{11}$$

where $\lambda_0 = 1$, and where a is an arbitrary shift parameter corresponding to the lower extremity of the support of $\varphi(x)$. We can also write (11) using Fourier variables

$$\hat{\varphi}(\omega) = \mathrm{e}^{-ja\omega} \Lambda(j\omega) \left(\frac{1-\mathrm{e}^{-j\omega}}{j\omega} \right)^L \tag{12}$$

where $\Lambda(x) \stackrel{\text{def}}{=} \sum_{n=0}^{L-1} \lambda_n x^n$ is a polynomial of degree $L-1$.

Some remarks on this result[2] are as follows.

- The solutions to our problem are neither trivial, nor have an infinite number of degrees of freedom. This will prove extremely convenient in the sequel, where we address some design problems.
- The most obvious, and also smoothest member of the family is the B-spline of degree $L-1$. This proves that B-splines have the smallest support for a given order.
- The minimal-support functions are piecewise polynomials with uniform knots. Moreover, due to the multiresolution property of B-splines, the MOMS can be seen to enjoy a related, though less simple, multiresolution property.
- The computational cost of Lth-order MOMS interpolation is *exactly* the same as that of Lth-order spline interpolation. This is because these functions all have the same support and the same degree.

The simple Fourier expression (12) makes it possible to express the asymptotic approximation constant C_φ^-. Using the first-order equivalent of $\hat{\varphi}$ in the neighborhood of $2n\pi$, that is $\hat{\varphi}(\omega) \approx (\Lambda(2jn\pi)/(2n\pi)^L)(\omega - 2n\pi)^L$, we find

$$C_\varphi^- = \sqrt{\sum_{n\neq 0} \left| \frac{\Lambda(2j\pi n)}{(2j\pi n)^L} \right|^2}. \tag{13}$$

From this equation, we can easily obtain the asymptotic constant for the spline of degree $L-1$, $C_{\mathrm{S}}^- = \sqrt{2\zeta(2L)}/(2\pi)^L$, which first appeared in [27].

[1]For our experiments in $D > 1$ dimensions, the kernel φ_D that we consider is obtained through a tensor product of the one-dimensional kernel $\varphi(x)$, i.e., $\varphi_D(x_1, x_2, \ldots x_D) = \varphi(x_1)\varphi(x_2)\ldots\varphi(x_D)$.

[2]After this paper was written, we became aware that Theorem 1 had already been published by Ron [31] in a more mathematically abstract and general context—the theory of exponential B-splines.

IV. Optimal Smallest-Support Functions: O-MOMS

We now design Λ in (13) in such a way as to minimize C_φ^-. This is equivalent to minimize

$$C(\Lambda)^2 = \sum_{n\neq 0} \left| \frac{\Lambda(2j\pi n)}{(2j\pi n)^L} \right|^2$$

under the constraint $\Lambda(0) = 1$. Rewriting this expression in terms of λ_n with $n = 0 \ldots L-1$, we have

$$C(\Lambda)^2 = \sum_{\substack{0\le k,l\le L-1 \\ k+l \text{ even}}} (-1)^{(k-l)/2} \frac{2\zeta(2L-k-l)}{(2\pi)^{2L-k-l}} \lambda_k \lambda_l \tag{14}$$

which involves the Riemann Zeta function defined in the notation part of this paper.

A. Solving the Minimization Problem

Numerical Method: This problem is obviously quadratic and can be rewritten in matrix form as: Minimize $\mathrm{X}^{\mathrm{t}}\mathbf{U}\mathrm{X}$ under the constraint ${\mathrm{e}_0}^{\mathrm{t}}\mathrm{X} = 1$, where

$$[\mathbf{U}]_{k,l} = \begin{cases} (-1)^{(k-l)/2} \frac{2\zeta(2L-k-l)}{(2\pi)^{2L-k-l}}, & \text{for } 0 \le k,l \le L-1 \\ \text{and } k+l \text{ even} \\ 0, & \text{otherwise} \end{cases}$$

$$\mathrm{X} = (\lambda_0, \ldots \lambda_{L-1})^{\mathrm{t}}$$

$$\mathrm{e}_0 = (1, 0, \ldots, 0)^{\mathrm{t}}.$$

The matrix involved, $\mathbf{U}$, is positive definite since $C^2(\Lambda) = 0$ implies $\Lambda(2n\pi) = 0$ for all $n \neq 0$; that is to say, $\Lambda = 0$ because Λ is a finite-degree polynomial. The solution is thus unique and takes the form $\mathrm{X}_{\min} = (\mathbf{U}^{-1}\mathrm{e}_0)/({\mathrm{e}_0}^{\mathrm{t}}\mathbf{U}^{-1}\mathrm{e}_0)$. Unfortunately, the numerical matrix method tends to be ill-conditioned for values of L that are large but still of interest.

Analytical Method: We show now that an analytical approach based on polynomials and continued fractions will enable us to find the explicit solution of our problem. Let $\varphi_{\mathrm{O}}^{L-1}(x)$ be the MOMS of order L that minimizes the asymptotic approximation constant $C(\Lambda)$. Let us also denote by $\Lambda_L(x)$ and C_L the polynomial associated to $\varphi_{\mathrm{O}}^{L-1}(x)$ according to (12), and the corresponding approximation constant, respectively. The functions $\varphi_{\mathrm{O}}^{L-1}$ are called optimal-MOMS or O-MOMS. The result is as follows.

Theorem 2: (O-MOMS) The O-MOMS are specified entirely by the induction relation

$$\Lambda_{L+1}(x) = \Lambda_L(x) + \frac{x^2}{4(4L^2-1)} \Lambda_{L-1}(x) \tag{15}$$

which is initialized by $\Lambda_1(x) = \Lambda_2(x) = 1$. Moreover, the minimal constant C_L is given by the explicit expression

$$C_L = \frac{L!}{(2L)!\sqrt{2L+1}}. \tag{16}$$

(See the proof in Appendix B.) Some remarks concerning this result are as follows.

- The induction (15) shows that the coefficients of $\Lambda_L(x)$ are positive and that this polynomial is even. We can thus

TABLE I
SAMPLING GAIN OF O-MOMS OVER B-SPLINE OF SAME DEGREE

order L	1	2	3	4	5	6	7	8	9	10
γ_L	1	1	1.223	1.463	1.707	1.951	2.195	2.437	2.680	2.922

deduce that its coefficients of even power (up to $L-1$) are all strictly positive; i.e., they do not vanish.

- When L is even, $\deg \Lambda_L = L-2$, which implies that the optimal basis function, $\varphi_L(x)$, is continuous and that its first derivative is discontinuous. When L is odd, $\deg \Lambda_L = L-1$, which implies that $\varphi_L(x)$ is discontinuous. This has an instructive and somewhat counterintuitive consequence: As far as approximation is concerned, there is *no link* between *regularity* and *quality*! In particular, this tends to invalidate interpolator design based on regularity; such an approach has moreover proved unsuccessful in a recent publication [32]. This does not mean that regularity is a useless parameter. For some applications it is indeed necessary to have a differentiable model, but this requirement should not be misinterpreted as a condition related to the accuracy of the approximation.

B. Asymptotical Approximation Gain

By construction, the optimal kernels $\varphi_{\mathrm{O}}^{L-1}(x)$ minimize the asymptotic constant C_φ^- among all MOMS. It is interesting to compare the minimal constant with that of splines in such a way that it reflects the gain in sampling density brought by using O-MOMS instead of splines.

Assume that we wish to reconstruct a function $f(x)$ using either B-splines of order L (i.e., of degree $L-1$) with sampling step-size T_{s}, or O-MOMS of the same order with sampling step-size T_{o}, such that the $\mathbf{L}^2$ approximation error is the same in both cases. We want to find the relation between T_{s} and T_{o}. Assuming that in both cases the oversampling of $f(x)$ is large enough, we get the approximation error from (9) which provides $C_{\mathrm{o}}^- T_{\mathrm{o}}^L \|f^{(L)}\|_{\mathbf{L}^2} = C_{\mathrm{s}}^- T_{\mathrm{s}}^L \|f^{(L)}\|_{\mathbf{L}^2}$ with obvious notations for the spline and O-MOMS asymptotic constants C_{s}^- and C_{o}^-. We finally obtain the sampling gain γ_L

$$\gamma_L \stackrel{\mathrm{def}}{=} \frac{T_{\mathrm{o}}}{T_{\mathrm{s}}} = \left(\frac{C_{\mathrm{o}}^-}{C_{\mathrm{s}}^-}\right)^{-1/L} = \left(\frac{(2L)!}{L!}\right)^{1/L} \frac{(2(2L+1)\zeta(2L))^{1/2L}}{2\pi}. \quad (17)$$

Theorem 3: (Asymptotic gain) When the order L is large, the sampling gain of O-MOMS over splines that have the same approximation order, *increases linearly* with L. Specifically

$$\gamma_L = \frac{2}{\pi e} L + O(\log L). \quad (18)$$

Proof: A basic application of Stirling's formula, $\log n! = \log(\sqrt{2\pi n}\mathrm{e}^{-n} n^n) + O(n^{-1})$, provides $1/L \log((2L)!/L!) = \log(4L/e) + O(L^{-1})$. Taking the logarithm of γ_L in (17), we readily obtain (18) after some manipulations which involve $\log \zeta(2L) = o(\log L)$ and $\log(2L+1) = O(\log L)$.

TABLE II
O-MOMS OF ORDERS 1 TO 6

L	O-MOMS	$\frac{C_{\mathrm{O}}^-}{C_{\mathrm{S}}^-}$
1	$\varphi_{\mathrm{O}}^0(x) = \beta^0(x)$	1
2	$\varphi_{\mathrm{O}}^1(x) = \beta^1(x)$	1
3	$\varphi_{\mathrm{O}}^2(x) = \beta^2(x) + \frac{1}{60}\frac{d^2}{dx^2}\beta^2(x)$	$\frac{1}{1.8}$
4	$\varphi_{\mathrm{O}}^3(x) = \beta^3(x) + \frac{1}{42}\frac{d^2}{dx^2}\beta^3(x)$	$\frac{1}{4.6}$
5	$\varphi_{\mathrm{O}}^4(x) = \beta^4(x) + \frac{1}{36}\frac{d^2}{dx^2}\beta^4(x) + \frac{1}{15120}\frac{d^4}{dx^4}\beta^4(x)$	$\frac{1}{14.5}$
6	$\varphi_{\mathrm{O}}^5(x) = \beta^5(x) + \frac{1}{33}\frac{d^2}{dx^2}\beta^5(x) + \frac{1}{7920}\frac{d^4}{dx^4}\beta^5(x)$	$\frac{1}{55.1}$

This theorem shows that the gain is quite substantial, and that it even tends to infinity as $L \to \infty$, which we did not expect *a priori*. This linear behavior is to be compared with the constant—asymptotic—gain, namely π, of spline versus Daubechies approximation [21].

C. Examples

The first six O-MOMS are shown in Table II with their asymptotic constant relative to that of the spline of same order. The asymptotic gain of O-MOMS over splines of same order is given in Table I for orders $1 \ldots 10$; the figures clearly confirm the asymptotic linear behavior that we have predicted in Theorem 3.

The optimal fourth-order function is plotted in Fig. 1 where it is compared to the cubic B-spline. Note the cusp at $x = 0$ which indicates that $\varphi_{\mathrm{O}}^3(x)$ is not continuously differentiable. Moreover, the plot of the ratio of the approximation kernels of the cubic O-MOMS and the cubic spline in Fig. 2 shows that the approximation using $\varphi_{\mathrm{O}}^3(x)$ is always better than that using $\beta^3(x)$. The gain in apprroximation error even exceeds 6 dB over half of the sampling bandwidth. According to (8), using the optimal function instead of a cubic spline for reconstruction (here $T = 1$), we expect at least a 6 dB SNR gain for signals whose frequency content lays essentially in the first half of the frequency domain. Obviously, this gain increases dramatically when the signal is more low-pass.

V. OTHER DESIGNS

A. Suboptimal MOMS (SO-MOMS)

Because the O-MOMS are discontinuous for even degrees (i.e., odd orders), it might be useful to design suboptimal interpolators that have a continuous derivative, while minimizing the asymptotic approximation constant among their class. Those suboptimal-MOMS or SO-MOMS, thus satisfy the same minimization problem as the O-MOMS in Section IV, with the restriction that $\Lambda(x)$ is of degree at most $L-3$ in (14).

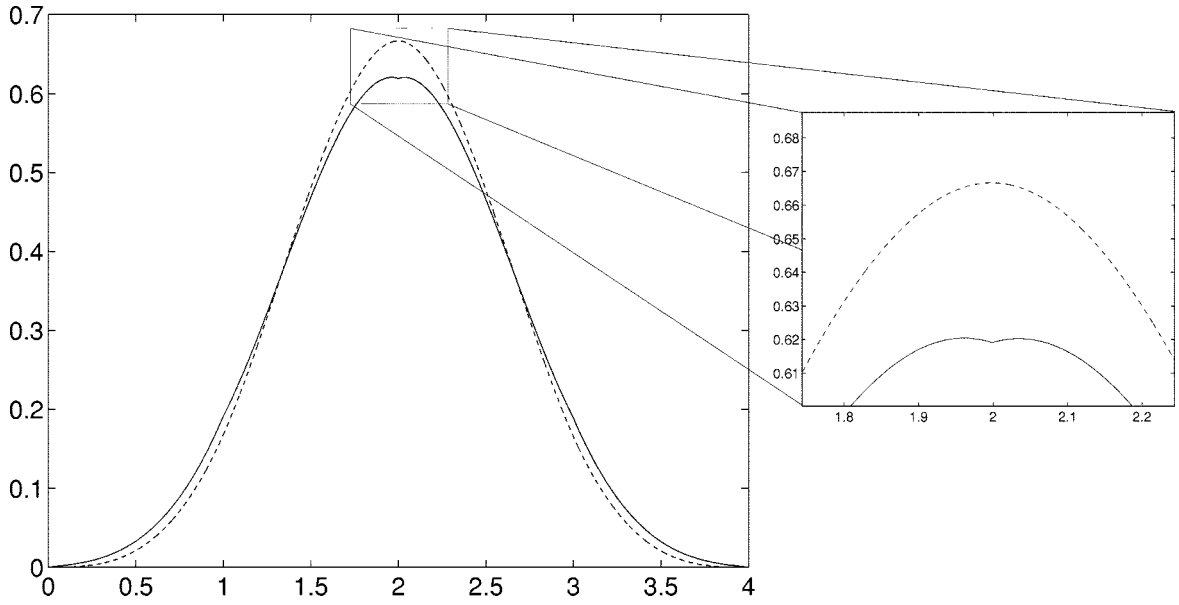

Fig. 1. Optimal function $\varphi_{\rm opt}$ (solid) and B-spline function (dashed) for $L = 4$.

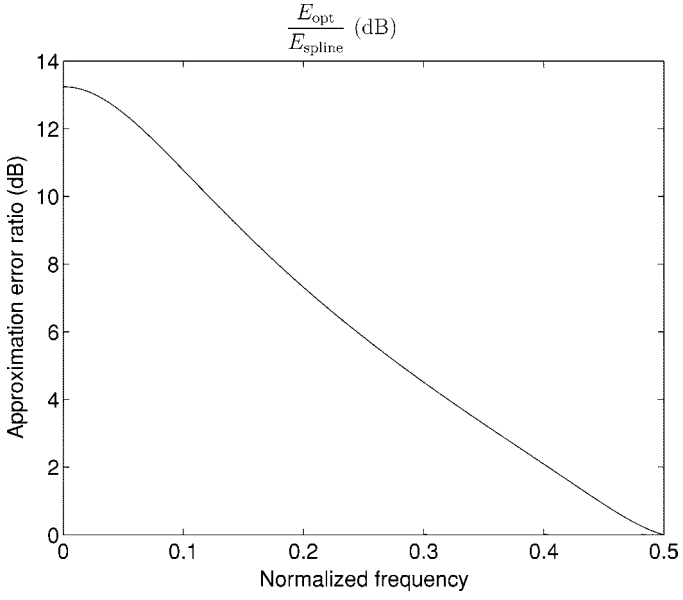

Fig. 2. Ratio between the optimal approximation kernel and the corresponding B-spline kernel for $L = 4$.

While a systematic analysis of those SO-MOMS is possible, we prefer to restrict ourselves to the case $L = 5, 6$ which are of most interest to us. The corresponding interpolators are still short and their degree is not too large, so that the interpolation algorithm is reasonably fast. Note that for $L \leq 4$, the SO-MOMS are the B-splines of same order.

Case $L = 5$*:* If we let $\Lambda(x) = 1 + ax + bx^2$ in (12), then the asymptotic constant $C(\Lambda)^2$ takes the form

$$C(\Lambda)^2 = a^2\frac{2\zeta(8)}{(2\pi)^8} + b^2\frac{2\zeta(6)}{(2\pi)^6} - 2b\frac{2\zeta(8)}{(2\pi)^8} + \frac{2\zeta(10)}{(2\pi)^{10}}$$

which is minimal for $a = 0$ and $b = (\zeta(8))/(4\pi^2\zeta(6)) = 1/40$. We thus have

$$\begin{aligned}\varphi^4_{\rm SO}(x) =& \beta^4(x) + \frac{1}{40}\frac{d^2}{dx^2}\beta^4(x)\\ C^-_{\varphi^4_{\rm SO}} =& \frac{C^-_{\beta^4}}{10}\end{aligned} \tag{19}$$

The loss $C^-_{\varphi^4_{\rm So}}/C^-_{\varphi^4_{\rm o}} \approx 1.45$ over the discontinuous O-MOMS approximation constant is the price to pay for having a continuously differentiable interpolator.

Case $L = 6$*:* If we let $\Lambda(x) = 1 + ax + bx^2 + cx^3$ in (12), then the asymptotic constant $C(\Lambda)^2$ takes the form

$$\begin{aligned}C(\Lambda)^2 =& a^2\frac{2\zeta(10)}{(2\pi)^{10}} + b^2\frac{2\zeta(8)}{(2\pi)^8} + c^2\frac{2\zeta(6)}{(2\pi)^6}\\ &- ac\frac{2\zeta(8)}{(2\pi)^8} - 2b\frac{2\zeta(10)}{(2\pi)^{10}} + \frac{2\zeta(12)}{(2\pi)^{12}}\end{aligned}$$

which is minimal for $a = c = 0$ and $b = \zeta(10)/(4\pi^2\zeta(8)) = 5/198$. We thus have

$$\begin{aligned}\varphi^5_{\rm so}(x) =& \beta^5(x) + \frac{5}{198}\frac{d^2}{dx^2}\beta^5(x)\\ C^-_{\varphi^5_{\rm so}} \approx& \frac{C^-_{\beta^5}}{20.7}\end{aligned} \tag{20}$$

The loss $C^-_{\varphi^5_{\rm So}}/C^-_{\varphi^5_{\rm o}} \approx 2.66$ over the continuous O-MOMS approximation constant is the price to pay for having a twice-continuously-differentiable interpolator. As a consequence, we recommend SO-MOMS only in operations that require some depth of differentiation. Moreover, notice that, for orders smaller than five, SO-MOMS coincide with B-splines of same order.

TABLE III
SO-MOMS OF ORDERS 1 TO 6

L	SO-MOMS	$\frac{C_{\mathrm{SO}}^-}{C_{\mathrm{S}}^-}$
1	$\varphi_{\mathrm{SO}}^0(x) = \beta^0(x)$	1
2	$\varphi_{\mathrm{SO}}^1(x) = \beta^1(x)$	1
3	$\varphi_{\mathrm{SO}}^2(x) = \beta^2(x)$	1
4	$\varphi_{\mathrm{SO}}^3(x) = \beta^3(x)$	1
5	$\varphi_{\mathrm{SO}}^4(x) = \beta^4(x) + \frac{1}{40}\frac{d^2}{dx^2}\beta^4(x)$	$\frac{1}{10}$
6	$\varphi_{\mathrm{SO}}^5(x) = \beta^5(x) + \frac{5}{198}\frac{d^2}{dx^2}\beta^5(x)$	$\frac{1}{20.7}$

We have summarized the values of the SO-MOMS in Table III, with their asymptotic constant, expressed in function of that of the spline of same order.

B. Interpolating MOMS (I-MOMS)

It is also possible to constrain the MOMS $\varphi(x)$ to be interpolating. Note however that the interpolation condition is unnecessarily restrictive and that it leads to marquedly suboptimal interpolation. This is confirmed by our experiments (see Section VI). Accordingly, we are generally not in favor of this condition, except for applications where memory is tight and the image array needs to be stored in byte format.

Due to the size of the support of φ, the interpolation conditions $\varphi(n+\alpha) = \delta_n$ add up to $L-1$ constraints if α is integer, and to L constraints otherwise; these equations match the L degrees of freedom of the MOMS. In fact, these interpolators derive from Lagrange's interpolation formula [17], [33], [34]: If we denote $P(x) = x(x+1)\ldots(x+L-1)$, then

$$\varphi(x) = \begin{cases} \frac{P([x]-\alpha)}{P'(-[x])(x-\alpha)} & \text{if } 0 \le x \le L \\ 0 & \text{otherwise,} \end{cases}$$

which can be shown to belong to the class of MOMS of order L.

We can obtain the polynomial $\Lambda(x)$ defining these I-MOMS directly, by expressing the interpolation condition in the Fourier Domain as

$$\begin{gathered} \forall\, n \in \mathbb{Z},\ \varphi(n+\alpha) = \delta_n \\ \Updownarrow \\ \forall\, \omega \in \mathbb{R}, \sum_n \hat{\varphi}(\omega + 2n\pi)\mathrm{e}^{j\alpha(\omega+2n\pi)} = 1. \end{gathered}$$

Since $\varphi(x)$ is of order L, we have that $\hat{\varphi}(\omega+2n\pi) = O(\omega^L)$ when $n \neq 0$. This precisely means that

$$\Lambda(j\omega) = \mathrm{e}^{-j\alpha\omega}\left(\frac{j\omega}{1-\mathrm{e}^{-j\omega}}\right)^L + O(\omega^L). \tag{21}$$

Our polynomial is thus given by the first L coefficients of the development of $\mathrm{e}^{-\alpha z}(z/(1-\mathrm{e}^{-z}))^L$ in Taylor series around the origin. For example, if $\alpha = L/2$, the I-MOMS are centered: The first six I-MOMS are given in Table IV with their asymptotic constant, expressed in function of that of the spline of same order. Note that these functions are discontinuous for odd orders.

TABLE IV
I-MOMS OF ORDERS 1 TO 6

L	I-MOMS	$\frac{C_{\mathrm{I}}^-}{C_{\mathrm{S}}^-}$
1	$\varphi_{\mathrm{I}}^0(x) = \beta^0(x)$	1
2	$\varphi_{\mathrm{I}}^1(x) = \beta^1(x)$	1
3	$\varphi_{\mathrm{I}}^2(x) = \beta^2(x) - \frac{1}{8}\frac{d^2}{dx^2}\beta^2(x)$	7.1
4	$\varphi_{\mathrm{I}}^3(x) = \beta^3(x) - \frac{1}{6}\frac{d^2}{dx^2}\beta^3(x)$	7.8
5	$\varphi_{\mathrm{I}}^4(x) = \beta^4(x) - \frac{5}{24}\frac{d^2}{dx^2}\beta^4(x) + \frac{3}{128}\frac{d^4}{dx^4}\beta^4(x)$	54.8
6	$\varphi_{\mathrm{I}}^5(x) = \beta^5(x) - \frac{1}{4}\frac{d^2}{dx^2}\beta^5(x) + \frac{1}{30}\frac{d^4}{dx^4}\beta^5(x)$	64.7

VI. EXPERIMENTS

The theoretical results that we have obtained in the previous sections are of special relevance to image processing. As observed in [34] and [35], the state-of-the-art interpolation method that performs best uses splines as basis functions. We will show that we can obtain visible improvement by using cubic O-MOMS instead of cubic splines in an interpolation experiment; the results for O-MOMS of other degrees also show the same trend [23], [24]. This is a very practical issue since, as indicated in introduction, there is no penalty from the point of view of computational cost [19]. The algorithms are essentially equivalent: The functions are cubic polynomials, have the same support, and the recursive prefilters have the same degree.

Description of the Experiments: To demonstrate experimentally the superiority of the new interpolators, we rotated images using the most direct method, i.e., we implemented the following steps: interpolation of the discrete image $I(k,l)$ into $I(x,y)$; rotation by the angle θ, giving $I_\theta(x,y) = I(x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta)$; and finally, resampling at the integers to get the discrete image $I_\theta(k,l)$.

We have used four sets of 512×512 images: 1) a circular symmetric test image that exhibits increasingly high frequencies near the center, which provides a direct insight of the behavior of the interpolation method with respect to spatial frequencies; 2) "Lena," which is significantly more low-pass than the other images, and thus is more robust to interpolation; 3) "Barbara," which has a strong texture content; and 4) a boat image "La Cornouaille," particularly interesting for its thin straight lines.

To amplify the differences, we applied 15 successive rotations by an angle of $2\pi/15$, and the final result was finally compared to the original. This was repeated for each interpolators. The final cumulative results are shown in Figs. 3 and 4, and the obtained SNR are computed. By increasing the order, we can saturate the SNR at a limit value that depends on the high-frequency content of the image. For reference, the measured saturation levels are 6 dB for the concentric circles, 37 dB for "Lena," 34 dB for "Barbara," and 40 dB for "La Cornouaille."

Influence of the Approximation Order: The rotated images in Fig. 3 confirm earlier practical findings [34]–[36] that a higher

Fig. 3. Comparison between three interpolation methods of increasing order: 15 rotations of the top images by an angle of $2\pi/15$. The higher the approximation order, the better the quality.

order is beneficial to quality. Note that cubic splines $(\text{order} = 4)$ win over the linear $(\text{order} = 2)$ and keys $(\text{order} = 3)$ methods, not only because they are of higher order but also because they are not constrained to satisfy the interpolation condition.

We made our experiments on a Power Macintosh G4/450 MHz. The computation time for the rotation experiment corresponding to a nonoptimized implementation of the linear, Keys, and cubic spline kernels was 5 s, 12 s, and 14 s, respectively.[3]

[3]A specialized implementation developed in our group at EPFL is actually faster than Keys' interpolation [24]

Influence of the Approximation Constant: We then set the approximation order to fourand compared the interpolation results using kernels that have the smallest size, i.e., the cubic MOMS. Once again, the theoretical prediction is confirmed by the experiment: among interpolators that have the same order, the smaller the asymptotic constant (see Tables II and III), the better the quality.

As expected, the cubic I-MOMS suffers from the interpolation condition, which makes it as fast as Keys' kernel (12 s in our experiment) but hinders approximation accuracy. This is

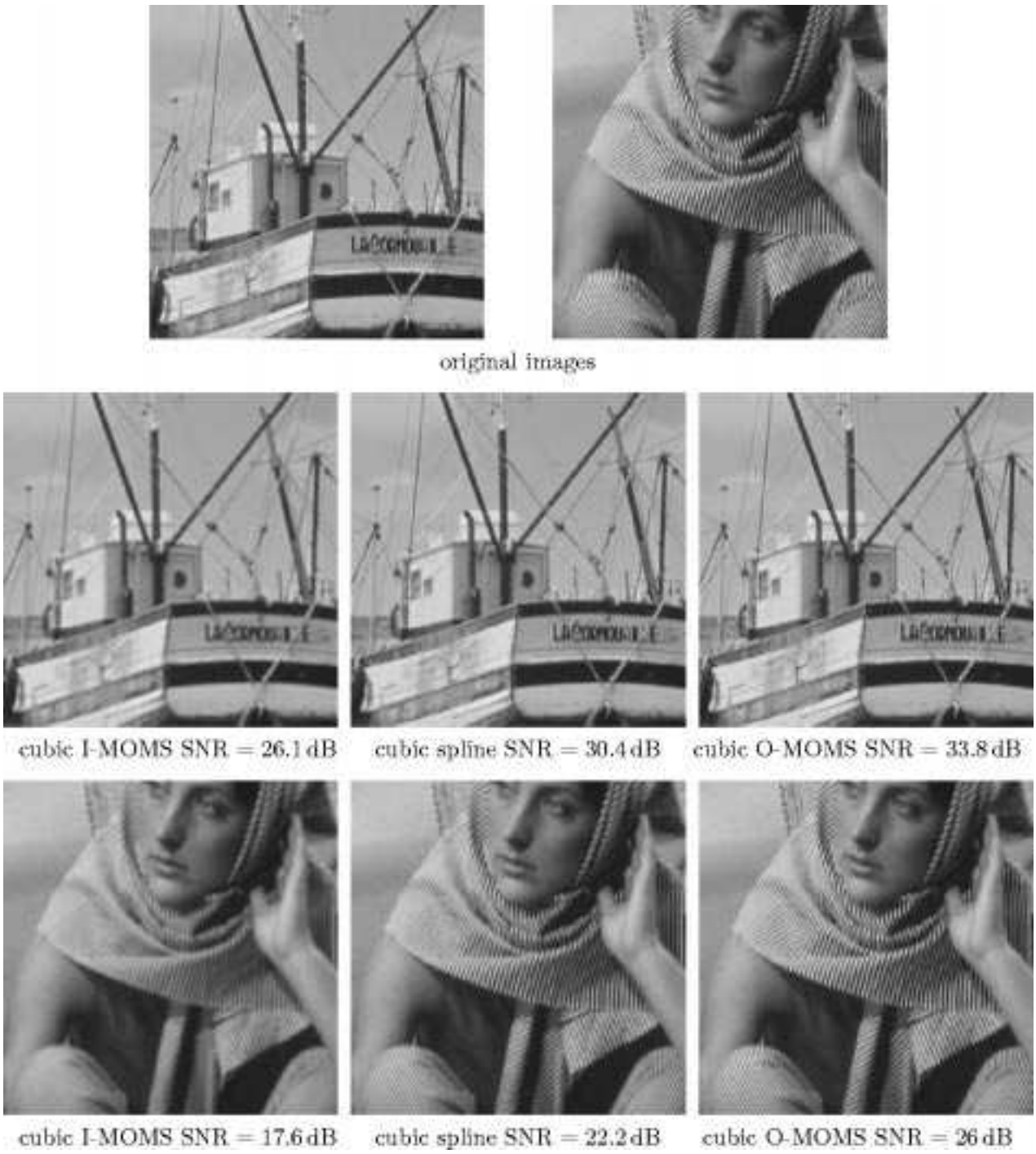

Fig. 4. Comparison between three interpolation methods of same order ($L = 4$), same support, and of decreasing asymptotic constant: 15 rotations of the top images by an angle of $2\pi/15$. The smaller the asymptotic constant, the better the quality. Remarks: The interpolation condition of the cubic I-MOMS is detrimental to accuracy; the cubic spline is the smoothest cubic MOMS, but does not provide the highest quality; the cubic O-MOMS minimizes the asymptotic constant among cubic MOMS and gives the best results, even though it is not smooth.

a constraint from which cubic B-splines and cubic O-MOMS are free. As it appears, the higher regularity of the spline does not seem to be at play since the cubic O-MOMS has a better performance, despite being barely continuous. Moreover, it requires the same computation time. Note that the gain in SNR is quite large, which shows that a careful design can really bring a dramatic improvement. A visual inspection reveals that, compared to cubic spline, the cubic O-MOMS interpolation is more faithful to the thin lines of the boat image, and to the striped pattern of Barbara's trousers.

VII. Conclusion

We have shown in this paper how to design basis functions so that they have the largest approximation order for a given support. This defines the class of MOMS which are optimal in this

respect and also have the same fixed computational cost. The MOMS are intimately related to the B-splines of same order. We were able to further minimize the asymptotic approximation constant [27] among this class of functions, which provided us with the O-MOMS. A practical image rotation experiment has shown that, used as interpolators, these new basis functions are superior to the B-splines of same order which, until now, were considered the state-of-the-art in interpolation. The only disadvantage of the O-MOMS is that, unlike B-splines, they are not very regular. This is why we also designed SO-MOMS to address this regularity issue.

The set of new basis functions that we have presented in this paper should be especially attractive for biomedical image processing, where quality is a key concern.

APPENDIX A
PROOF OF THEOREM 1 ("MOMS")

Let $\varphi_L(x)$ be an Lth-order kernel that is compactly supported within $[a,b]$. That is, it satisfies (6). Let us define the function $\psi(x)$ by

$$\psi(x)=\sum_{n\geq 0}\varphi_L'(x-n) \tag{22}$$

where the differentiation is taken in the sense of distributions. Thus, we have

$$\varphi_L'(x)=\psi(x)-\psi(x-1). \tag{23}$$

Then, we have the following properties:

i) *$\psi(x)$ is compactly supported within $[a,b-1]$.*
This is a consequence of $\sum_n \varphi_L(x-n)=1$ (Strang-Fix condition of order 1) which is equivalent to $\sum_n \varphi_L'(x-n)=0$ in the distributional sense. Because of this identity, we also have the following alternative definition of $\psi(x)$:

$$\psi(x)=-\sum_{n\leq -1}\varphi_L'(x-n).$$

According to this expression, the support of $\psi(x)$ is contained within $]-\infty,b-1]$. But, according to the initial definition (22), we also have support $(\psi)\subset[a,+\infty[$. Hence, we conclude support $(\psi)\subset[a,b-1]$.

ii) *$\psi(x)$ satisfies the Strang-Fix conditions of order $L-1$.*
To prove this, we differentiate (6). This yields $\sum_k P'(x-k)\varphi_L(x-k)+\sum_k P(x-k)\varphi_L'(x-k)=0$. We replace the first term of this equation by $C_{P'}$ (since P' is a polynomial of degree $<L-1$) and $\varphi_L'(x)$ in the second term by its expression (23). Since ψ is compactly supported, we easily obtain

$$\sum_k \underbrace{(P(x-k)-P(x-k+1))}_{Q(x-k)}\psi(x-k)=-C_{P'}.$$

Now, if $P(x)$ spans the entire set of polynomials of degree $\leq L-1$, then $Q(x)$ spans the entire set of polynomials of degree $\leq L-2$. This also means that for all polynomial $Q(x)$ of degree $\leq L-2$, there exists a constant C_Q such that $\sum_k Q(x-k)\psi(x-k)=C_Q$. In addition, we see that if $P(x)=x$, then $Q(x)=-1$; thus, $-\sum_k \psi(x-k)=-\sum_k \varphi_L(x-k)$. Integrating over $[0,1]$ yields to $\int\psi=\int\varphi_L=1$. Thus, $\psi(x)$ satisfies the Strang-Fix conditions of order $L-1$.

Thanks to these properties, we can reason by induction on the approximation order, setting $\varphi_{L-1}(x)=\psi(x)$. This induction process yields a set of $L+1$ distributions $\{\varphi_{L-n}(x)\}_{n=0\ldots L}$ that enjoy the following properties:

i) $\varphi_{L-n}(x)$ is compactly supported within $[a,b-n]$.
ii) $\varphi_{L-n}(x)$ satisfies Strang-Fix conditions of order $L-n$.
iii) $\varphi_{L-n}(x)$ is linked to $\varphi_L(x)$ through the Fourier relationship

$$\hat{\varphi}_L(\omega)=\hat{\beta}^{n-1}(\omega)\hat{\varphi}_{L-n}(\omega).$$

This expression is obtained directly by taking the Fourier transform of (23). Note that it can also be written as the convolution $\varphi_L(x)=\beta^{n-1}*\varphi_{L-n}(x)$.

Conversely, if we take a distribution $\psi(x)=\varphi_0(x)$ that has zeroth-order approximation and that is supported within $[a,b']$, then $\varphi_L(x)$ defined by $\varphi_L(x)=\beta^{L-1}*\psi(x)$ is compactly supported within $[a,b'+L]$ and has approximation order L.

Now, minimizing the support of $\varphi_L(x)$ means finding the smallest b such that $\psi(x)$ exists. Of course, this is possible only if $b'=b-L\geq a$, which yields $\psi(x)$ as a single-point distribution. This shows that the minimum size of the support of $\varphi_L(x)$ is $b-a=L$ as claimed in the first part of the theorem.

Finally, we know from distribution theory [37, Th. XXXV] that the only distributions that have a support of zero-measure are finite linear combinations of the Dirac distribution and of its derivatives at this point. Thus, for the minimal-support function $\varphi_L(x)$, there exist constants λ_n such that $\psi(x)=\sum_{n\geq 0}\lambda_n\delta^{(n)}(x-a)$, where the sum is finite and where $\lambda_0=1$ because $\int\psi=1$. This means that $\varphi_L(x)=\sum_{n\geq 0}\lambda_n(d^n)/(dx^n)\beta^{L-1}(x-a)$. Since we must restrict ourselves to $\mathbf{L}^2$ functions, the summation has to run from 0 to $L-1$. This provides the final result (11) of our theorem.

APPENDIX B
PROOF OF THEOREM 2 ("OPTIMAL-MOMS")

Fourier Version of the Minimization Problem: Instead of the Matrix $\mathbf{U}$, we use the $\mathcal{U}$ function defined by the entire series

$$\mathcal{U}(z)=\sum_{n\geq 1}(-1)^{n+1}\frac{2\zeta(2n)}{\pi^{2n}}z^{2n}. \tag{24}$$

This series converges in the complex plane for $|z|<\pi$: within the convergence disc, a classical relation provides $\mathcal{U}(z)=z\coth Z-1$ [38]. A straightforward consequence of (24) is that

$$\int_0^{2\pi}\mathcal{U}(\mathrm{e}^{j\theta})\mathrm{e}^{-nj\theta}\frac{d\theta}{2\pi}=\begin{cases}(-1)^{\frac{n}{2}+1}\frac{2\zeta(n)}{\pi^n} & n \text{ even}>0\\ 0 & \text{otherwise.}\end{cases} \tag{25}$$

Replacing n by $2L-k-l$ in (25) allows to rewrite (14) as

$$C(\Lambda)^2 = \sum_{k,l=0}^{L-1} \lambda_k \lambda_l \frac{(-1)^{L+l+1}}{2^{2L-k-l}} \int_0^{2\pi} \mathcal{U}(\mathrm{e}^{j\theta}) \mathrm{e}^{-(2L-k-l)j\theta} \frac{d\theta}{2\pi}$$

which becomes

$$C(\Lambda)^2 = \frac{(-1)^{L+1}}{2^{2L+1}\pi} \int_0^{2\pi} \mathcal{U}(\mathrm{e}^{j\theta}) \Lambda(2\mathrm{e}^{j\theta}) \lambda(-2\mathrm{e}^{j\theta}) \mathrm{e}^{-2jL\theta} d\theta$$

after some rearrangements.

Finding the polynomial $\Lambda_L(z)$ of degree $L-1$ that minimizes the quadratic functional $C(\Lambda)^2$ under the constraint $\Lambda(0)=1$ yields the following equivalent conditions:

$$\int_0^{2\pi} \mathrm{e}^{nj\theta} \mathcal{U}(\mathrm{e}^{j\theta}) \Lambda_L(2\mathrm{e}^{j\theta}) \mathrm{e}^{-2jL\theta}\, d\theta = 0, \quad \forall\, n = 1 \ldots L-1$$

and $\Lambda_L(0)=1$, which are valid for $L \geq 1$. The $L-1$ conditions above are clearly equivalent to requiring that the polynomial series $\Lambda_L(2z)\mathcal{U}(z)$ is lacunary in the powers $z^{L+1}, z^{L+2}, \ldots z^{2L-1}$. We can thus express our solution under the following form: There exist a polynomial of degree $\leq L, P_L(z) = \sum_{0 \leq n \leq L} p_{L,n} z^n$, and an entire series, $Q_L(z) = \sum_{n \geq 0} q_{L,n} z^n$, such that

$$\mathcal{U}\left(\frac{z}{2}\right) \Lambda_L(z) = P_L(z) + z^{2L} Q_L(z) \tag{26}$$

and $\Lambda_L(0)=1$. Note that if another polynomial $M(z)$ of degree $\leq L-1$ satisfies (26), then, because of the *unicity* of the solution of our minimization problem, we have $M(z) = M(0)\Lambda_L(z)$. We will see next that the continued-fraction of $\mathcal{U}(z)$ involves precisely such a polynomial $M(z)$.

Also note that, as a consequence of the minimization of the quadratic form, we find a simplified expression for $C_L^2 = C(\Lambda_L)^2$. Specifically, we have $C_L^2 = (1)/(2\pi) \int_0^{2\pi} \mathcal{U}(\mathrm{e}^{j\theta}) \Lambda_L(-j\mathrm{e}^{j\theta}) \mathrm{e}^{-2jL\theta} d\theta$, that is to say, $C_L^2 = (-1)^L Q_L(0)$.

The Continued Fraction of $\mathcal{U}(z)$ and its Link with $\lambda_L(z)$: Decomposing the entire function $\mathcal{U}(z)$ in a continued fraction consists of iterating the following process:$\mathcal{U}(z) = \mathcal{U}(0) + z^2/(\mathcal{U}_{1(z)}), \mathcal{U}_1(z) = \mathcal{U}_1(0) + z^2/(\mathcal{U}_2(z))$ and so forth. In the end, $f(z)$ is given by the continued fraction

$$\mathcal{U}(z) = a_0 + \cfrac{z^2}{a_1 + \cfrac{z^2}{a_2 + \cfrac{z^2}{a_3 + \cdots}}} \tag{27}$$

the continued fraction of $\tanh(z)$ is known in [38], so is that of $z \coth z - 1$. Thus, the coefficients a_n in (27) are given by $a_n = 2n+1$ for $n \geq 1$ and $a_0 = 0$.

If we keep the first n terms only we get the "nth convergent" of $f(z)$, which can be written as the rational fraction $A_n(z)/B_n(z)$. The continued fraction theory tells us that these polynomials satisfy the following constitutive properties:

- Induction equations [38]

$$\begin{aligned} A_n(z) &= a_n A_{n-1}(z) + z^2 A_{n-2}(z) \\ B_n(z) &= a_n B_{n-1}(z) + z^2 B_{n-2}(z) \end{aligned} \tag{28}$$

  initialized by $A_0(z)=0$ and $B_0(z)=1$.
- $\deg A_n = 2\lfloor n+1/2 \rfloor$ and $\deg B_n = 2\lfloor n/2 \rfloor$, which can easily be shown using the induction relations (28) satisfied by A_n and B_n.
- By construction, the difference between $\mathcal{U}(z)$ and its nth convergent $A_n(z)/B_n(z)$ is $O(z^{2n+2})$, that is to say,

$$\mathcal{U}(z) - \frac{A_n(z)}{B_n(z)} = z^{2n+2} R_n(z) \tag{29}$$

  where $R_n(z)$ is entire.

If we replace z by $z/2$ in (29) and let $n = L-1$, we observe that we obtain two polynomials $A_{L-1}(z/2)$ of degree $\leq L$, and $B_{L-1}(z/2)$ of degree $\leq L-1$ such that $\mathcal{U}(z/2)B_{L-1}(z/2) = A_{L-1}(z/2) + z^{2L} S_L(z)$, where $S_L(z)$ is entire. This is exactly the form of (26) which we know to have a unique solution such that $\Lambda_L(0)=1$. Consequently, we have

$$\Lambda_L(z) = \frac{B_{L-1}(\frac{z}{2})}{B_{L-1}(0)}.$$

Moreover, using the induction relation (28) and the known value for a_n, we get $B_n(0) = (2n+1)B_{n-1}(0)$. Thus, replacing B_n by the corresponding value of Λ_{n+1} for $n = L-2, L-1, L$, we easily get the induction (15) for Λ_L.

Asymptotic Constant: Finally, in order to evaluate the asymptotic constant C_L, we need to compute $Q_L(0)$, where $Q_L(z)$ is the entire series defined by (26). Because of the induction (15), we have $z^2 Q_{L+1}(z) = Q_L(z) + 1/(4(4L^2-1))Q_{L-1}(z)$. This means that $Q_L(0) = -1/(4(4L^2-1))Q_{L-1}(0)$. Due to the link $C_L^2 = (-1)^L Q_L(0)$, we thus have

$$C_L^2 = \frac{C_{L-1}^2}{4(4L^2-1)}.$$

By induction on L, and by using $C_1^2 = 1/12$, we now easily get [16].

ACKNOWLEDGMENT

The authors wish to thank the anonymous reviewers and the Associate Editor Prof. B. Evans for their comments which helped improve the readability of this paper.

REFERENCES

[1] E. V. R. Di Bella, A. B. Barclay, R. L. Eisner, and R. W. Schafer, "A comparison of rotation-based methods for iterative reconstruction algorithms," *IEEE Trans. Nucl. Sci.*, vol. 43, pp. 3370–3376, Dec. 1996.

[2] M. R. Smith and S. T. Nichols, "Efficient algorithms for generating interpolated (zoomed) MR images," *Magn. Reson. Med.*, vol. 7, pp. 156–171, 1988.

[3] A. P. Berkhoff, H. J. Huisman, J. M. Thijssen, E. M. G. P. Jacobs, and R. J. F. Homan, "Fast scan conversion algorithms for displaying ultrasound sector images," *Ultrason. Imag.*, vol. 16, pp. 87–108, Apr. 1994.

[4] W. D. Richard and R. M. Arthur, "Real-time ultrasonic scan conversion via linear interpolation of oversampled vectors," *Ultrason. Imag.*, vol. 16, pp. 109–123, Apr. 1994.
[5] S. D. Fuller, S. J. Butcher, R. H. Cheng, and T. S. Baker, "Three-dimensional reconstruction of icosahedral particles—The uncommon line," *J. Struct. Biol.*, vol. 116, pp. 48–55, Jan.-Feb. 1996.
[6] J. L. Ostuni, A. K. S. Santha, V. S. Mattay, D. R. Weinberger, R. L. Levin, and J. A. Frank, "Analysis of interpolation effects in the reslicing of functional MRI images," *J. Comput. Assisted Tomography*, vol. 21, pp. 803–810, 1997.
[7] B. Migeon and P. Marche, "In vito 3-D reconstruction of long bones using B-scan image processing," *Med. Biol. Eng. Comput.*, vol. 35, pp. 369–372, July 1997.
[8] M. Haddad and G. Porenta, "Impact of reorientation algorithms on quantitative myocardial SPECT perfusion imaging," *J. Nucl. Med.*, vol. 39, pp. 1864–1869, 1998.
[9] F. M. Weinhaus and V. Devarajan, "Texture mapping 3-D models of real-world scenes," *ACM Comput. Surv.*, vol. 29, pp. 325–365, Dec. 1997.
[10] C. Frederick and E. L. Schwartz, "Conformal image warping," *IEEE Comput. Graph. Appl.*, vol. 10, pp. 54–61, Mar. 1990.
[11] T. Möller, R. Machiraju, K. Mueller, and R. Yagel, "Evaluation and design of filters using a Taylor series expansion," *IEEE Trans. Vis. Comput. Graphics*, vol. 3, pp. 184–199, Apr.-June 1997.
[12] F. J. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform," *Proc. IEEE*, vol. 66, pp. 51–83, 1978.
[13] C. R. Appledorn, "A new approach to the interpolation of sampled data," *IEEE Trans. Med. Imag.*, vol. 15, pp. 369–376, June 1996.
[14] N. A. Dodgson, "Quadratic interpolation for image resampling," *IEEE Trans. Image Processing*, vol. 6, pp. 1322–1326, Sept. 1997.
[15] R. G. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, no. 6, pp. 1153–1160, 1981.
[16] I. German, "Short kernel fifth-order interpolation," *IEEE Trans. Signal Processing*, vol. 45, pp. 1355–1359, May 1997.
[17] A. Schaum, "Theory and design of local interpolators," *Comput. Vis. Graph. Image Process.*, vol. 55, pp. 464–481, Nov. 1993.
[18] S. K. Park and R. A. Schowengerdt, "Image reconstruction by parametric cubic convolution," *Comput. Vis. Graph. Image Process.*, vol. 23, pp. 258–272, 1983.
[19] T. Blu, P. Thévenaz, and M. Unser, "Generalized interpolation: higher quality at no additional cost," in *Proc. Int. Conf. Image Process.*, vol. III, Kobe, Japan, Oct. 1999, pp. 667–671.
[20] T. Blu and M. Unser, "Quantitative Fourier analysis of approximation techniques: Part I—Interpolators and projectors," *IEEE Trans. Signal Processing*, vol. 47, pp. 2783–2795, Oct. 1999.
[21] ——, "Quantitative Fourier analysis of approximation techniques: Part II—Wavelets," *IEEE Trans. Signal Processing*, vol. 47, pp. 2796–2806, Oct. 1999.
[22] M. Unser, "Splines: a perfect fit for signal and image processing," *IEEE Signal Processing Mag.*, vol. 16, pp. 22–38, Nov. 1999.
[23] P. Thévenaz, T. Blu, and M. Unser, "Interpolation revisited," *IEEE Trans. Med. Imag.*, vol. 19, pp. 739–758, July 2000.
[24] ——, "Image interpolation and resampling," in *Handbook of Medical Imaging, Processing and Analysis*, I. N. Bankman, Ed. New York: Academic, 2000, pp. 393–420.
[25] G. Strang and G. Fix, "A Fourier analysis of the finite element variational method" (in Cremonese), in *Constructive Aspect of Functional Analysis*, G. Geymonat, Ed. Rome, Italy: Edizioni Cremonese, 1971, pp. 796–830.
[26] G. Strang and T. Q. Nguyen, *Wavelets and Filter Banks*. Cambridge, MA: Wellesley-Cambridge, 1996.
[27] M. Unser, "Approximation power of biorthogonal wavelet expansions," *IEEE Trans. Signal Processing*, vol. 44, pp. 519–527, Mar. 1996.
[28] O. Rioul, "Simple regularity criteria for subdivision schemes," *SIAM J. Math. Anal.*, vol. 23, no. 6, pp. 1544–1576, Nov. 1992.
[29] T. Blu and M. Unser, "Approximation error for quasiinterpolators and (multi-) wavelet expansions," *Appl. Comput. Harmon. Anal.*, vol. 6, no. 2, pp. 219–251, Mar. 1999.
[30] C. de Boor, "Quasiinterpolation and approximation power of multivariate splines," in *Computation of Curves and Surfaces*, W. Dahmen, Ed. Norwell, MA: Kluwer, 1990, pp. 313–345.
[31] A. Ron, "Factorization theorems for univariate splines on regular grids," *Isr. J. Math.*, vol. 70, no. 1, pp. 48–68, 1990.
[32] E. H. W. Meijering, K. J. Zuiderveld, and M. A. Viergever, "Image reconstruction by convolution with symmetrical piecewise nth-order polynomial kernels," *IEEE Trans. Image Processing*, vol. 8, pp. 192–201, Feb. 1999.
[33] R. W. Schafer and L. R. Rabiner, "A digital signal processing approach to interpolation," *Proc. IEEE*, vol. 61, pp. 692–702, June 1973.
[34] E. H. W. Meijering, W. J. Niessen, and M. A. Viergever, "The sinc approximating kernels of classical polynomial interpolation," in *Proc. Int. Conf. Image Process.*, vol. III, Kobe, Japan, Oct. 1999, pp. 652–656.
[35] M. Unser, P. Thévenaz, and L. Yaroslavsky, "Convolution-based interpolation for fast, high-quality rotation of images," *IEEE Trans. Image Processing*, vol. 4, pp. 1371–1381, Oct. 1995.
[36] M. Unser and I. Daubechies, "On the approximation power of convolution-based least-squares versus interpolation," *IEEE Trans. Signal Processing*, vol. 45, pp. 1697–1711, July 1997.
[37] L. Schwartz, *Théorie des Distributions* (in French). Paris, France: Hermann, 1966.
[38] M. Abramovitz and I. A. Stegun, *Handbook of Mathematical Functions*. Washington, DC: Nat. Bur. Stand., 1972.

Thierry Blu (M'96) was born in Orléans, France, in 1964. He received a degree from École Polytechnique, France, in 1986, and a degree from Télécom Paris (ENST), France, in 1988. In 1996, he received the Ph.D. degree in electrical engineering from ENST for a study on interated rational filter banks, applied to wide-band audio coding.

He is currently with the Biomedical Imaging Group at Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, on leave from France Télécom National Center for Telecommunications Studies (CNET), Issy-les-Moulineaux, France. His research interests include (multi-) wavelets, multire-solution analysis, multirate filter banks, approximation and sampling theory, and psychoacoustics.

Philippe Thévenaz was born in Lausanne, Switzerland, in 1962. He received the degree in microengineering from the Swiss Federal Institute of Technology (EPFL), Lausanne, in January 1986, and the Ph.D. degree from the Institute of Microtechnology (IMT) of the University of Neuchâtel, Switzerland, in June 1993, with a thesis on the use of the linear prediction residue for text-independent speaker recognition. While at IMT, he worked in the domain of image processing (optical flow) and in the domain of speech processing (speech coding and speaker recognition).

He was a Visiting Fellow with the Biomedical Engineering and Instrumentation Program, National Institutes of Health (NIH), Bethesda, MD, where he developed research interests that include splines and multi-resolution signal representations, geometric image transformations, and biomedical image registration. Since 1998, he has been with the EPFL as First Assistant.

Michael Unser (M'89–SM'94–F'99) received the M.S. (summa cum laude) and Ph.D. degrees in electrical engineering from the Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, in 1981 and 1984, respectively.

From 1985 to 1997, he was with the Biomedical Engineering and Instrumentation Program, National Institutes of Health, Bethesda, MD. He is now Professor and Head of the Biomedical Imaging Group at the EPFL. His main research area is biomedical image processing. He has a strong interest in sampling theories, multi-resolution algorithms, wavelets, and the use of splines for image processing. He is the author of 90 published journal papers in these areas.

Dr. Unser is an Associate Editor for the IEEE TRANSACTIONS ON MEDICAL IMAGING. He is on the editorial boards of several other journals, including *Signal Processing*, IEEE TRANSACTIONS ON IMAGE PROCESSING (1992–1995), and IEEE SIGNAL PROCESSING LETTERS (1994–1998). He serves as a regular chair for the SPIE Conference on Wavelets, held annually since 1993. He received the 1995 Best Paper Award and the 2000 Magazine Award from the IEEE Signal Processing Society. In January 1999, he was elected Fellow of the IEEE with the citation "for contributions to the theory and practice of splines in signal processing."