# Smoothed Differentiation Filters for Images ${ }^{1}$ 

Peter Meer<br>Department of Electrical and Computer Engineering, Rutgers University, Piscataway, New Jersey 08855-0909; and Center for Automation Research, University of Maryland, College Park, Maryland 20742-3411

AND

Isaac Weiss
Center for Automation Research, University of Maryland, College Park, Maryland 20742-3411
Received December 11, 1990; accepted June 11, 1991


#### Abstract

Computation of the derivatives of an image defined on a lattice structure is of paramount importance in computer vision. The solution implies least square fitting of a continuous function to a neighborhood centered on the site where the value of the derivative is sought. We present a systematic approach to the problem involving orthonormal bases spanning the vector space defined over the neighborhood. Derivatives of any order can be obtained by convolving the image with a priori known filters. We show that if orthonormal polynomial bases are employed the filters have closed form solutions. The same filter is obtained when the fitted polynomial functions have one consecutive degree. Moment preserving properties, sparse structure for some of the filters, and relationship to the Marr-Hildreth and Canny edge detectors are also proven. Expressions for the filters corresponding to fitting polynomials up to degree six and differentiation orders up to five, for the cases of unweighted data and data weighted by the discrete approximation of a Gaussian, are given in the appendices. © 1992 Academic Press, Inc.


## 1. INTRODUCTION

In computer vision an image $\mathbf{g}$ is defined as the ensemble of values $g(n, m)$ allocated to the sites of a sampling lattice. The following problem frequently appears in applications:

Given the noisy image $\mathbf{g}$, estimate at the site $(n, m)$, by computations restricted to a local neighborhood, the value of the $(p+q)$ th partial derivative ( $p$ along the $x$ axis, $q$ along the $y$ axis) of the original (uncorrupted) image.

[^0]The problem is of paramount importance in edge detection with differential operators, where first- and secondorder derivatives are employed. Higher-order derivatives are required in certain applications of 3D vision [1, 26]. In this paper we present a method through which derivatives of any order can be obtained by convolving the image $g$ with a priori known kernels. Our goal is to present a systematic approach to differentiation of discrete data preceded by least square smoothing.

The direct method, employing finite differences to compute an approximation of the derivative, is not a satisfactory solution. The difference operator amplifies the noise present in the image and the result is unreliable. Differentiation must be preceded by a smoothing operation which reduces the amount of noise at the expense of the achievable resolution for the values of the derivatives. In edge detection, given the edge profile of interest, the two operations can be combined optimally either in the image domain or in the frequency domain (see [2, pp. 55-80], for a complete review).

In this paper we take a traditional approach in which the trade-off between detection and localization [7] does not have a central role. Nevertheless, some of Canny's results will be obtained as a particular case of the new method. We focus on computational issues and present a fast solution to the proposed problem involving only convolutions with known kernels. In later sections we return in detail to the topics to be introduced in this section.
We define in the neighborhood around ( $n, m$ ) a continuous function $f(x, y)$, called by Haralick [10] the underlying function of the image. This function carries our assumptions about the local image structure. Let $f(n, m)$ be the sample at site $(n, m)$ of the function $f(x, y)$. The step size of the sampling lattice along both directions is taken equal to one. To achieve noise reduction, the number of parameters of $f(x, y)$ must be less than the number of
samples in the neighborhood. These parameters can be found by minimizing the mean square error relative to the data, the values of $\mathbf{g}$ in the neighborhood. The function $f(x, y)$ is then employed for the local analysis of $\mathbf{g}$.

The amount of computation is significantly reduced if $f(x, y)$ is defined in terms of orthogonal basis functions spanning the vector space corresponding to the neighborhood. The optimal orthogonal basis is obtained from Karhunen-Loève expansion of the data [14, 16]. A special class of basis functions, the orthogonal polynomials, are also of interest.

Orthogonal polynomial bases have often been employed in computer vision. Hueckel [15] used nine twodimensional polynomials up to the fourth degree to detect edges and lines. They correspond to Chebyshev polynomials in polar coordinates. Hartley [11] employed the first six two-dimensional Hermite polynomials to solve the same task. Paton [25] employed the first six twodimensional Legendre polynomials to describe the local structure of the image. These authors regarded the neighborhood as a continuous interval with the values $g(n, m)$ extended (zero-order interpolation) between the sampling lattice sites.

In the discrete approach, while $f(x, y)$ is still defined as a continuous function, the computations are restricted to the lattice points of the neighborhood. Haralick [10] employed this method to estimate directional derivatives in noisy images. His polynomials, up to third degree, are discrete Chebyshev polynomials, also known as Gram polynomials [13, p. 290]. Hashimoto and Sklansky [12] derived discrete approximations of Gaussian derivative estimation filters using Krawtchouk polynomials.

In this paper we too adopt the discrete approach. In Section 2 the mathematics of orthogonal bases over discrete intervals is presented. The general expression for smoothed differentiation filters is obtained in Section 3, and the filters built with orthogonal polynomial bases are discussed in Section 4. Experimental results are shown in Section 5 and the paper is concluded with a discussion section.

## 2. ORTHOGONAL BASES OVER DISCRETE INTERVALS

We now give a short review of orthogonal bases defined on discrete intervals. Without loss of generality we can employ the one-dimensional approach and define the interval as being of length $2 N+1$ centered on $n$, i.e., $k=n-N, \ldots, n, \ldots, n+N$. Let $\psi_{l}(x), l=0$, $1, \ldots, \infty$ be a set of functions and $w(x) \geq 0$ a weight function, all defined on the interval $-N \leq x \leq+N$. The functions can take only finite values and are continuous in the interval. The $l_{2}$ discrete scalar product of two functions with respect to the weight function on the discrete
set of points $n-N, \ldots, n, \ldots, n+N$ is defined as

$$
\begin{equation*}
\left(\psi_{l}, \psi_{j}\right)=\sum_{k=n-N}^{n+N} w(k-n) \psi_{l}(k-n) \psi_{j}(k-n) . \tag{1}
\end{equation*}
$$

Note that the only function values employed are the ones at the lattice sites and that the scalar product implicitly depends on $n$, the center of the interval under consideration. The set of functions $\left\{\psi_{l}\right\}$ is orthogonal if for any $l$ and $j$

$$
\left(\psi_{l}, \psi_{j}\right)=\Psi_{l}^{2} \delta_{l j}= \begin{cases}\Psi_{l}^{2}, & l=j  \tag{2}\\ 0, & \text { otherwise }\end{cases}
$$

where $\Psi_{l}^{2}=\left(\psi_{l}, \psi_{l}\right)=\left|\psi_{l}\right|^{2}$ is the $l_{2}$ norm of the function $\psi_{l}$. The set of functions $\phi_{l}(x)=\left\{\psi_{l}(x) / \Psi_{l}\right\}$, where $l=0$, $1, \ldots, \infty$, is called an orthonormal basis.
Let the samples $g(k), k=n-N, \ldots, n, \ldots$, $n+N$, be the available discrete data. We want to obtain in mean squares sense the best continuous approximation of the data on the interval by employing a continuous function $f(x)$ having $L+1$ real parameters $f_{l}, l=0$, $1, \ldots, L<2 N+1$. That is, we assume that between the available samples $g(k)$ the data can be represented by $f(x)$ built from the linear combination of the first $L+1$ basis functions

$$
\begin{equation*}
f(x) \triangleq \sum_{l=0}^{L} f_{l} \phi_{l}(x-n), \quad n-N \leq x \leq n+N \tag{3}
\end{equation*}
$$

where the expansion coefficients $f_{l}$ are to be determined. This limited set of basis functions forms the orthonormal base of an ( $L+1$ )-dimensional linear vector space. The coefficients $f_{l}$ are then determined by minimizing the $l_{2}$ error norm between the sequences $f(x), x=k$, and $g(k)$

$$
\begin{align*}
\varepsilon_{L}^{2} & =|f-g|^{2} \\
& =\sum_{k=n-N}^{n+N} w(k-n)\left[\sum_{l=0}^{L} f_{l} \phi_{l}(k-n)-g(k)\right]^{2} . \tag{4}
\end{align*}
$$

Minimum error is achieved if

$$
\begin{equation*}
\frac{\partial \varepsilon_{L}^{2}}{\partial f_{l}}=0, \quad l=0,1, \ldots, L \tag{5}
\end{equation*}
$$

Writing (5) explicitly and taking into account (1) we obtain

$$
\begin{equation*}
\sum_{j=0}^{L} f_{j}\left(\phi_{j}, \phi_{l}\right)=\left(g, \phi_{l}\right), \quad l=0,1, \ldots, L \tag{6}
\end{equation*}
$$

The set of equations (6) is known as the normal equations, and has a central role in mean square estimation. Since the functions $\phi_{l}$ belong to an orthonormal basis (2) the solution to the normal equations is immediate:

$$
\begin{equation*}
f_{l}=\left(g, \phi_{l}\right), \quad l=0,1, \ldots, L \tag{7}
\end{equation*}
$$

Note that $f_{l}$ implicitly depends on $n$, i.e., on the interval on which the computations are performed. Analysis of (7) shows that the expansion coefficients of the continuous approximation $f(x)$ are identical with the projections of the discrete data $g(n)$ onto the basis functions and they do not depend on the dimension of the orthonormal basis. Any other set of coefficients yields a higher error. The coefficients $f_{l}$ are computed independently of each other, and as a corollary we have that the parameters minimizing the mean square error for dimension $L_{1}$ are also the first $L_{1}$ parameters if the minimization is performed for dimension $L_{2}>L_{1}$. This important property is another reason for the popularity of orthogonal expansions for least squares problems, since whenever a better approximation is required the already available parameters do not have to be thrown away.

## 3. SMOOTHED DIFFERENTIATION FILTERS

The underlying function $f(x)$ carries all the assumed information about the image $g$ in the neighborhood under consideration. This neighborhood is centered on $n$, and thus the lattice sites of interest are $k=n-N, \ldots$, $n, \ldots, n+N$. We approximate the $p$ th derivative of the discrete noisy image $g$ at the site $n$ by

$$
\begin{equation*}
f^{(p)}(n)=\left[\frac{d^{p} f(x)}{d x^{p}}\right]_{x=n} \tag{8}
\end{equation*}
$$

that is, by the value of the continuous underlying function's $p$ th derivative at the lattice site under consideration. The expression for $f^{(p)}(x)$ can be obtained from (3) and (7) using the linearity of the differentiation operator,

$$
\begin{equation*}
f^{(p)}(n)=\sum_{l=0}^{L}\left(g, \phi_{l}\right) \phi_{l}^{(p)}(0) \tag{9}
\end{equation*}
$$

where the dependence on $n$ is through the scalar product. Writing (9) explicitly, after some manipuations we have

$$
\begin{equation*}
f^{(p)}(n)=\sum_{k=n-N}^{n+N} g(k)\left[w(k-n) \sum_{l=0}^{L} \phi_{l}(k-n) \phi_{l}^{(p)}(0)\right] \tag{10}
\end{equation*}
$$

Making a change in the summation index we obtain the solution

$$
\begin{equation*}
f^{(p)}(n)=\sum_{i=-N}^{N} g(n-i) h(i ; p)=g(n)^{*} h(n ; p) \tag{11}
\end{equation*}
$$

where

$$
h(n ; p)= \begin{cases}w(-n) \sum_{l=0}^{L} \phi_{l}(-n) \phi_{l}^{(p)}(0), & |n| \leq N  \tag{12}\\ 0, & \text { otherwise }\end{cases}
$$

with the dependence on the parameter $p$ made explicit but the dependence on $L$ and $N$ kept implicit. The expression (11) is a convolution sum between the data $g(n)$ and a finite impulse response filter $h(n ; p)$ and is denoted by the asterisk. Note that $h(n ; p)$ while completely determined by the chosen orthonormal basis does not have a closed form solution in the general case. We obtained the following important result:

To estimate the $p$ th derivative of the one-dimensional noisy image at site $n$ in the chosen neighborhood we have to convolve the image with an a priori known filter.

For the two-dimensional case we restrict ourselves to square neighborhoods. In such neighborhoods it is always possible to define a separable two-dimensional orthonormal basis built by the Cartesian product of two identical one-dimensional bases. Thus
$w(x, y)=w(x) w(y)$ and $\phi_{l_{1} l_{2}}(x, y)=\phi_{l_{1}}(x) \phi_{l_{2}}(y)$
with all the properties discussed in Section 2 holding. We approximate the $(p+q)$ th partial derivative at site ( $n$, $m$ ), $p$ along the $x$ axis and $q$ along the $y$ axis, by

$$
\begin{equation*}
f^{(p, q)}(n, m)=\left[\frac{\partial^{p+q} f(x, y)}{\partial x^{p} \partial y^{q}}\right]_{x=n, y=m} \tag{14}
\end{equation*}
$$

Similarly to the one-dimensional case we have

$$
\begin{equation*}
f^{(p, q)}(n, m)=\sum_{l_{1}=0}^{L} \sum_{l_{2}=0}^{L}\left(g, \phi_{l_{1} l_{2}}\right) \phi_{l_{1}}^{(p)}(0) \phi_{l_{2}}^{(p)}(0) \tag{15}
\end{equation*}
$$

and the separability of the orthonormal basis allows us to write

$$
\begin{equation*}
f^{(p, q)}(n, m)=\sum_{i=-N}^{N} \sum_{j=-N}^{N} g(n-i, m-j) h(i ; p) h(j ; q), \tag{16}
\end{equation*}
$$

where the expression of the filters $h(n ; \cdot)$ is given by (12). Thus, the two-dimensional problem can be reduced to convolution with the product of two filters employed in the one-dimensional case. It must be emphasized that all the results obtained so far are valid for any orthonormal basis defined on a discrete interval. The chosen basis defines the continuous approximation and therefore influences the values obtained for the derivatives. In the next section we show that by using orthonormal polynomial bases closed form solutions can be obtained for the filters.

## 4. FILTERS BUILT WITH ORTHONORMAL POLYNOMIALS

The orthonormal polynomial basis functions are defined as

$$
\begin{equation*}
\phi_{l}(x)=\sum_{k=0}^{l} a_{l, k} x^{k}, \quad l=0,1, \ldots, L \tag{17}
\end{equation*}
$$

For any $l$ and $j$ the orthonormality condition

$$
\begin{equation*}
\left(\phi_{l}, \phi_{j}\right)=\sum_{n=-N}^{N} w(n) \phi_{l}(n) \phi_{j}(n)=\delta_{l j} \tag{18}
\end{equation*}
$$

must be satisfied, and thus the orthogonality interval $[-N, N]$ and the weight function $w(x)$ uniquely determine the orthonormal polynomial base, i.e., all the coefficients $a_{l, k}$. An extensive literature on orthogonal polynomials exists. Orthogonal polynomials over discrete intervals, however, are less often discussed because they can be regarded as particular cases of polynomials defined over continuous intervals. The most complete reference on orthogonal polynomials, the treatise of Szegö [31] discusses them in only a few pages (pp. 33-37).

It is immediate from the definition (17) that the continuous function $g(x)=x^{k},-N \leq x \leq N$ and $k \leq L$, can be completely represented by polynomial basis functions. Thus, when the sequence $g(n)=n^{k}, n=-N, \ldots, N$, is applied at the input of the smoothed differentiation filter $h(n ; p)$, the sought output at $n=0$ (the center of the interval) is

$$
f^{(p)}(0)=\sum_{n=-N}^{N}(-n)^{k} h(n ; p)= \begin{cases}0, & k \neq p  \tag{19}\\ p!, & k=p\end{cases}
$$

Note that both $k$ and $p$ are less than or equal to the largest polynomial degree $L$. When $k=0$ the following properties of the filters are obtained:

$$
\sum_{n=-N}^{N} h(n ; p)= \begin{cases}1, & p=0  \tag{20}\\ 0, & p=1,2, \ldots, L\end{cases}
$$

As expected, the smoothing filter reproduces the mean level of the input, while if differentiation is also present the mean level is discarded.

The $k$ th moment of the sequence $h(n ; p)$ has the definition

$$
\begin{equation*}
\mu_{k}[h(n ; p)] \triangleq \sum_{n=-N}^{N} h(n ; p) n^{k} \tag{21}
\end{equation*}
$$

and (19) yields

$$
\mu_{k}[h(n ; p)]= \begin{cases}0, & k \neq p  \tag{22}\\ (-1)^{p} p!, & k=p\end{cases}
$$

In the case of smoothing ( $p=0$ ) (22) becomes

$$
\mu_{k}[h(n ; 0)]= \begin{cases}1, & k=0  \tag{23}\\ 0, & k=1,2, \ldots, L\end{cases}
$$

We can proceed now to compute the moments of the output $f^{(p)}(n)$ as functions of the moments of the finite input sequence $g(n)$. The input is defined in the interval $[-M, M]$ and thus the output is nonzero in the interval $[-M-N, M+N]$,

$$
\begin{align*}
\mu_{k} & {\left[f^{(p)}\right](n) } \\
& =\sum_{n=-M-N}^{M+N} f^{(p)}(n) n^{k} \\
& =\sum_{n=-M-N}^{M+N}(i+n-i)^{k} \sum_{i=-N}^{N} g(n-i) h(i ; p)  \tag{24}\\
& =\sum_{j=0}^{k}\binom{k}{j} \sum_{i=N}^{N} i^{j} h(i ; p) \sum_{n=-M-N}^{M+N}(n-i)^{k-j} g(n-i) \\
& =\sum_{j=0}^{k}\binom{k}{j} \mu_{j}[h(n ; p)] \mu_{k-j}[g(n)],
\end{align*}
$$

where $\binom{k}{j}$ are the binomial coefficients. Taking into account (22) we can reduce (24) to

$$
\begin{equation*}
\mu_{k}\left[f^{(p)}(n)\right]=(-1)^{p} p!\binom{k}{p} \mu_{k-p}[g(n)] . \tag{25}
\end{equation*}
$$

The following relation between the moments of the input and of the output is obtained:

$$
\mu_{k}\left[f^{(p)}(n)\right]= \begin{cases}0, & k<p  \tag{26}\\ (-1)^{p} \frac{k!}{(k-p)!} \mu_{k-p}[g(n)], & k \geq p\end{cases}
$$

All the moments of the output of degree less than the order of differentiation $p$ are zero. Higher moments are proportional with the first moments of the input.

The case of smoothing $p=0$ is of particular interest:

$$
\begin{equation*}
\mu_{k}[f(n)]=\mu_{k}[g(n)], \quad k=0,1, \ldots, L \tag{27}
\end{equation*}
$$

The smoothing filters $h(n, 0)$ preserve all the moments of the input signal $g(n)$ up to $L$, the degree of the polynomial fitted to the discrete data. It must be emphasized that the input $g(n)$ is not necessarily a signal which can be completely represented by polynomials; for example, it can be a noisy step edge. Thus the smoothing filters maximally reduce the corrupting noise and yield minimum distortion. These properties justify Burt's [6] hierarchical surface interpolation method in which only planar fits were employed. Since by our method the parameters of the smoothing polynomials are not computed explicitly additional improvements and/or higher-order local fits might be achieved in Burt's algorithm.

### 4.1. Chebyshev and Krawtchouk Polynomial Bases

Let the dimension of the space spanned by the orthonormal polynomial bases be $L+1$; i.e., the degree of the fitted polynomials is $L$. Given the orthogonality interval $-N \leq x \leq N$, different bases correspond to different weight functions $w(x)$. Two weight functions are of interest for us.

The Chebyshev polynomials $t_{l}(x)$ are obtained when $w(x)=1 ;-N \leq x \leq N$. They are also known as the Gram polynomials to distinguish them from polynomials with the same name but different weight function defined over a continuous interval.

The Krawtchouk polynomials $k_{l}(x)$ are generated when

$$
\begin{align*}
w(x)= & \binom{2 N}{N-x} a^{N-x} b^{N+x} \\
= & \frac{(2 N)!}{(N-x)!(N+x)!} a^{N-x} b^{N+x} \\
& a, b>0, a+b=1 ;-N \leq x \leq N \tag{28}
\end{align*}
$$

Since only the samples of $w(x), x=n=-N, \ldots, N$, are used in the scalar product (1) the binomial coefficients can be written as the ratio of factorials and not as of gamma functions. For our purpose it suffices to consider the Krawtchouk polynomials corresponding to $a=b=\frac{1}{2}$.

Several methods for generating the polynomials belonging to an orthogonal family are available. The recurrence method through a formula connecting three orthogonal polynomials of consecutive degrees is relevant here. The orthonormal polynomials are then obtained by division by their norm.

The Chebyshev polynomials can be generated from the relation (adapted from Steffen [29])

$$
\begin{align*}
t_{l+1}(x)= & 2 \frac{2 l+1}{l+1} x t_{l}(x) \\
& -\frac{l}{l+1}(2 N+1-l)(2 N+1+l) t_{l-1}(x) \tag{29}
\end{align*}
$$

The Krawtchouk polynomials can be generated from the relation (adapted from Greenleaf [8])

$$
\begin{equation*}
k_{l+1}(x)=\frac{x}{l+1} k_{l}(x)-\frac{1}{4(l+1)}(2 N+1-l) k_{l-1}(x) \tag{30}
\end{equation*}
$$

where $l=1,2, \ldots, L-1$; and the initial polynomials are $t_{0}(x)=k_{0}(x)=1$ and $t_{1}(x)=2 x$ and $k_{1}(x)=x$. Because only the term $x$ multiplies the $l$ th polynomial in both recurrence relations it is immediate to prove that the polynomials of even degree are even functions (only even powers of $x$ ) and the polynomials of odd degree are odd functions (only odd powers of $x$ and no free term). Thus

$$
\begin{array}{ll}
t_{2 j}^{(2 q+1)}(0)=k_{2 j}^{(2 q+1)}(0)=0, & j, q=0,1, \ldots  \tag{31}\\
t_{2 j+1}^{(2 q)}(0)=k_{2 j+1}^{(2 q}(0)=0, & j, q=0,1, \ldots
\end{array}
$$

Odd-order derivatives of the even-degree polynomials and even-order derivatives of the odd-degree polynomials are zero in the center of the neighborhood.

The importance of this property becomes clear when we examine the expression for the smoothed differentiation filters (12). For any two consecutive values $L$ and $L$ +1 , one of them (say $L+1$ ) yields $\phi_{L+1}^{(p)}(0)=0$ for polynomial bases. Thus, the same filter is obtained for the computation of the $p$ th derivative in the center of the neighborhood if the underlying function is assumed to be a polynomial of degree $L$ or $L+1$.

For example, assume that smoothing is performed by fitting at every site a line segment (first-order fit) within a neighborhood centered on that site. The value of the fit at the center is the free term in the line segment's equation, which as is well known is equal to the mean value of the samples (zero-order fit). Similarly, there is no need to use a second-order fit when computing the smoothed first derivative; the first-order fit gives identical results.

Another consequence of (31) is that all the smoothed differentiation filters are either odd or even sequences. The derivatives at zero always sift out all the polynomials of the same nature (odd or even), leaving only the other homogeneous group inside the sum. We have

$$
\begin{gather*}
h(-n ; 2 q)=h(n ; 2 q), \quad h(-n ; 2 q+1)=-h(n ; 2 q+1) \\
|n| \leq N, \quad q=0,1, \ldots \tag{32}
\end{gather*}
$$

The expressions of the first five orthogonal Chebyshev polynomials were obtained from (29) by the help of the Macsyma program package. They are proportional to the polynomials found by Hildebrand [13, p. 290]:

$$
\begin{aligned}
t_{0}(x)= & 1 \\
t_{1}(x)= & 2 x \\
t_{2}(x)= & 6 x^{2}-2 N(N+1) \\
t_{3}(x)= & 20 x^{3}-4\left(3 N^{2}+3 N-1\right) x \\
t_{4}(x)= & 70 x^{4}-10\left(6 N^{2}+6 N-5\right) x^{2} \\
& +6 N\left(N^{2}-1\right)(N+2) \\
t_{5}(x)= & 252 x^{2}-140\left(2 N^{2}+2 N-3\right) x^{3} \\
& +4\left(15 N^{4}+30 N^{3}-35 N^{2}-50 N+12\right) x .
\end{aligned}
$$

The general expression for the squared-norm of these Chebyshev polynomials is

$$
\begin{equation*}
T_{l}^{2}=\frac{2 N+1}{2 l+1} \prod_{j=1}^{l}\left[(2 N+1)^{2}-j^{2}\right] . \tag{34}
\end{equation*}
$$

The orthonormal polynomials are obtained by dividing the expressions in (33) by $T_{l}$.
The filters based on Chebyshev polynomials were computed from (12) with the help of Macsyma and are given in Appendix A. Convolution with the filters involves the fitting of an $L$ th degree polynomial to the data. The fitted polynomial has $L+1$ coefficients which are determined implicitly by the filtering and thus the dimension of the interval $2 N+1$ must be at least equal to $L+1$, i.e., $N \geq$ L/2.

The filters built with Chebyshev polynomial base are called Savitzky-Golay filters after the names of the researchers who proposed them first in 1964. The filters are frequently used in chemistry for processing spectrometric measurements. Their popularity is shown by the fact that according to the Science Citation Index the paper was cited in 1987 alone 124 times. Savitzky and Golay [27] tabulated the filters up to the fifth degree. The tables of Savitzky and Golay were corrected by Steinier et al. [30], who also discussed the least squares problem in a more general context. Madden [18] was the first to give closed form general expressions for the Savitzky-Golay filters. He did not elaborate on the method by which the filters were obtained. Madden's results differ from ours given in Appendix A by a minus sign for the filters which are odd sequences. This is due to Madden's correlation type definition in contrast with ours (11) which is of convolution type. We also found an error in his expression for the filter (A10). Bromba and Ziegler [3, 5] discussed only the Savitzky-Golay smoothing filters. Different
properties are presented and a recursive filter implementation (more adequate for on-line scanning of the input data) was proposed. Steffen [29] and Schüssler and Steffen [28] also restricted themselves to smoothing filters. They gave two different methods of obtaining the Savitzky-Golay filters, proved their properties, and showed the shape of the filters and of the corresponding frequency responses for $N=11, L=0,2, \ldots, 22$.

The expressions of the first five orthogonal Krawtchouk polynomials corresponding to $a=b=\frac{1}{2}$ in the weight function (28) were obtained from (30) by the help of the Macsyma program package. They are proportional to the polynomials found by Greenleaf [8]:

$$
\begin{align*}
& k_{0}(x)=1 \\
& k_{1}(x)=x \\
& k_{2}(x)=\frac{1}{4}\left[2 x^{2}-N\right]  \tag{35}\\
& k_{3}(x)=\frac{1}{12}\left[2 x^{3}-(3 N-1) x\right] \\
& k_{4}(x)=\frac{1}{96}\left[4 x^{4}-4(3 N-2) x^{2}+3 N(N-1)\right] \\
& k_{5}(x)=\frac{1}{48}\left[4 x^{5}-20(N-1) x^{3}+\left(15 N^{2}-25 N+6\right) x\right] .
\end{align*}
$$

The general expression for the squared-norm of these Krawtchouk polynomials is

$$
\begin{equation*}
K_{l}^{2}=2^{-2 l}\binom{2 N}{2 N-l} \tag{36}
\end{equation*}
$$

The expressions of the filters built with the Krawtchouk orthonormal base are given in Appendix B. They are less frequently employed than the Savitzky-Golay filters; only Bromba and Ziegler [4] discussed their use for the case of pure smoothing.

The behavior of the Chebyshev and Krawtchouk polynomial bases at the limit is also of interest. If we let $x=$ $N z$ and make $N$ go to infinity, the Chebyshev polynomials defined over the discrete interval $[-N, N]$ become the Legendre polynomials in the variable $z$, defined over the continuous interval $[-1,1]$ with unit weight function [13, p. 290]. Similarly, the Krawtchouk polynomials at the limit yield Hermite polynomials defined on the continuous interval $[-\infty, \infty]$ with Gaussian weights [8].

In Section 1 we mentioned that both Legendre and Hermite polynomials were employed in computer vision for feature detection. The relative small neighborhood sizes, however, raise questions about the validity of using continuous orthogonal polynomial bases. As was discussed above, these continuous polynomials become correct approximations for data defined on a discrete lattice only at the limit. It is of interest to mention that Hashimoto and Sklansky [12] used Krawtchouk polynomials to approximate Gaussian derivative estimation filters. Our
results, using the discrete approach from the beginning, shows that there is no need to consider these filters as approximations.

### 4.2. Properties of Filters Built with Krawtchouk Polynomials

In this section we discuss some interesting properties of the smoothed differentiation filters built with Krawtchouk polynomials.

Examination of the expressions for the filters (Appendix B) shows that they all contain the sampled values of the weight function (28). To show the relationship between some of these filters and well-known edge detectors developed in the continuous domain, we should regard the weight as the the discrete approximation of a Gaussian. By the DeMoivre-Laplace theorem [23, p. 66] if $N$ is relative large and $n>\sqrt{N / 2}$, the weight of the Krawtchouk polynomial base becomes a good approximation of a Gaussian with mean 0 and variance $N / 2$.

The expression of the approximated Gaussian is

$$
\begin{equation*}
G(x)=\frac{1}{\sqrt{2 \pi N}} e^{-\left(x^{2} / N\right)} \tag{37}
\end{equation*}
$$

where we took into account the unit step size of the sampling lattice. The first and second derivatives are

$$
\begin{align*}
& G^{(1)}(x)=-\frac{2 x}{N} G(x)  \tag{38a}\\
& G^{(2)}(x)=\frac{2}{N^{2}}\left(2 x^{2}-N\right) G(x) \tag{38b}
\end{align*}
$$

Given that the sampled Gaussian $G(n)$ is a good approximation of the sampled weight function $w(n)$, comparison of (38a) with (B5) and (38b) and (B8) reveals the following equivalences:

$$
\begin{array}{ll}
h_{K}(n ; 1)=G^{(1)}(n) & \text { for } L=1 \text { or } 2 \\
h_{K}(n ; 2)=\frac{2 N}{2 N-1} G^{(2)}(n) & \text { for } L=2 \text { or } 3 \tag{39b}
\end{array}
$$

Thus the two smoothed differentiation filters built with Krawtchouk polynomials are close approximations of the sampled derivatives of a Gaussian with variance $N / 2$. The approximation improves as the support of the filters increases.

Canny [7] proposed the first derivative of a Gaussian as an efficient approximation for his optimal step edge operator. Canny defined optimality in terms of localization accuracy and minimum false alarms and the optimization was performed in the continuous domain. In our discrete approach we employed a different, least mean square
criterion to determine the expressions for the smoothed differentiation filters. Interestingly, when the local structure of the image is assumed to be planar or quadratic the obtained filter is very similar to Canny's edge detector. Note that higher-order underlying functions yield different first-order differentiation filters ((B6) and (B7)) and in our paradigm the first derivative of the Gaussian is no longer the desirable operator.

Marr and Hildreth [19] employed the second-order derivative of a Gaussian as edge detector. They gave an empirical justification in the continuous domain for the proposed operator. The equivalence relation (39b) suggests that the discrete mean square criterion for Gaus-sian-weighted data yields a similar filter only when the local structure of the image is assumed to be a polynomial of degree two or three.

Implementational aspects of the smoothed differentiation filters are also of importance. We now show that some of the filters built with Krawtchouk polynomials have sparse structure which reduces fourfold the amount of computation needed in two dimensions. The SavitzkyGolay filters built with Chebyshev polynomials do not have this property.

Let a filter $h(n)$ be defined on the interval $n=$ $-N, \ldots, N$. The filter has spare structure if close to half of its coefficients are zero in an alternating pattern. That is,

$$
\begin{gather*}
h(n)=0, \quad|n|=2 q \\
q=c,(c+1),(c+1), \ldots,[N / 2] \tag{40a}
\end{gather*}
$$

where $c$ is either 0 or 1 . Another pattern is

$$
\begin{gather*}
h(n)=0, \quad|n|=2 q+1 \\
q=0,1, \ldots,\lfloor(N-1) / 2\rfloor \tag{40b}
\end{gather*}
$$

Mutiplications with these zero-valued coefficients do not have to be performed and in convolutions a significant decrease in the amount of computation is achieved. Some of these filters with sparse structure are known as halfband filters. Another application of half-band filters in computer vision is described in Meer et al. [20].

The smoothed differentiation filters described in this paper have three parameters in addition to the discrete variable $n$. The size of the filter is related to $N$, the degree of the fitted smoothing polynomials is related to $L$, and the order of the differentiation is $p$. We have found that the filters have sparse structure if either

$$
\begin{equation*}
N=L^{+} \quad \text { and } \quad p=0 \tag{41}
\end{equation*}
$$

or

$$
\begin{equation*}
N=L^{-}=p \tag{42}
\end{equation*}
$$

where the superscript of the parameter $L$ stands for the smaller or larger polynomial degree yielding the same filter.

The first two sparse-structured smoothing filters defined by condition (41) are listed in Table 1. The smallest filter ( $N=1$ ) is not included because it does not have zero-valued coefficients.

Condition (42) yields the sparse-structured smoothed differentiation filters shown in Table 2. The first filter ( $N=1$ ) is given only for completeness.

We prove now that the spare structure conditions (41) and (42) are a consequence of a property of the roots of Krawtchouk polynomials (35). The roots of the algebraic equation $k_{N}(x)=0$ are integers having alternating pattern. While the roots of any orthogonal polynomial are real and distinct and lie inside the orthogonality interval [31, p. 44], the above property of Krawtchouk polynomials is not shared with the Chebyshev polynomials.

Proof of Condition (41). With the help of the Chris-toffel-Darboux summation formula Bromba and Ziegler [4] have shown that in the case of pure smoothing ( $p=0$ ) the expression (12) can be further simplified for the Krawtchouk polynomial base to

$$
\begin{equation*}
h(n ; 0,2 q)=C_{2 q} \frac{k_{2 q+1}(n)}{n} w(n), \quad q=0,1, \ldots, \tag{43}
\end{equation*}
$$

where $C_{2 q}$ is a constant, $k_{2 q+1}(n)$ is the $(2 q+1)$ th Krawtchouk polynomial, the weight function $w(n)$ is defined in (28), and the dependence on the degree of fitting polynomial $L=2 q$ was made explicit. Recall that if $L=$ $2 q+1$ the same smoothing filter is obtained. The polynomials $k_{2 q+1}(n)$ are odd functions and thus the division with $n$ is always defined. The binomial coefficients $w(n)$ are always nonzero, the sparse structure of the smoothing filters (41) is then generated by the roots of the even polynomial $k_{2 q+1}(n) / n$. For example, if $q=1$ we are in the $L^{-}=2, N=L^{+}=3$ case and the equation of interest is

$$
\begin{equation*}
\frac{k_{3}(n)}{n}=\frac{1}{12}\left[2 n^{2}-8\right]=0 \tag{44}
\end{equation*}
$$

having the solutions $n= \pm 2$ which coincide with the zeros of the filter in Table 1.

TABLE 1
Smoothing Filters with Sparse Structure

| Filter size: | Degree of <br> $2 N+1$ | Nature of <br> polynomials: $L^{-}, L^{+}$ | $h(n)=0$ <br> sequence |
| :---: | :---: | :---: | :---: |
| 7 | 2 or 3 | Even | 2 |
| 11 | 4 or 5 | Even | 2,4 |

TABLE 2
Smoothed Differentiation Filters with Sparse Structure

| Filter <br> size: <br> $2 N+1$ | Degree of <br> polynomials: <br> $L^{-}, L^{+}$ | Order of <br> differentiation: <br> $p$ | Nature of <br> sequence | $h(n)=0$ <br> $\|n\|$ |
| :---: | :---: | :---: | :---: | ---: |
| 3 | 1 or 2 | 1 | Odd | 0 |
| 5 | 2 or 3 | 2 | Even | 1 |
| 7 | 3 or 4 | 3 | Odd | 0,2 |
| 9 | 4 or 5 | 4 | Even | 1,3 |
| 11 | 5 or 6 | 5 | Odd | $0,2,4$ |

Proof of Condition (42). Since the $N$ th-order derivative of an $l<N$ th degree polynomial is zero, condition (42) reduces the expression (12) of the filter to

$$
\begin{equation*}
h(n ; N, N)=N!a_{N, N} w(-n) k_{N}(-n) \tag{45}
\end{equation*}
$$

where $a_{N, N}$ is the leading coefficient of the $k_{N}(x)$ Krawtchouk polynomial. Note that $w(-n)=w(n)$ and $k_{N}(-n)$ $= \pm k_{N}(n)$ depending on whether $N$ is an even or an odd number. Thus, again a sparse filter structure is generated by the roots of the Krawtchouk polynomials. For example, if $N=4$ the roots are given by

$$
\begin{equation*}
4 n^{4}-40 n^{2}+36=0 \tag{46}
\end{equation*}
$$

with the solutions $n= \pm 1, \pm 3$ which coincide with the zeros of the smoothed fourth-order differentiation filter in Table 2.

## 5. EXPERIMENTAL RESULTS

We have applied several smoothed differentiation filters to the $128 \times 128$ pebbles image (upper left, Fig. 1) and $158 \times 158$ vase image (upper left, Fig. 2). All the images presented in the figures are magnified to $256 \times 256$ and have $\mathbf{2 5 5}$ gray levels. These two images were chosen as examples since most of the information conveyed is carried by nonplanar patches. The results are obtained with filters built employing the Krawtchouk polynomial base. Results obtained with the equivalent SavitzkyGolay filters are undistinguishable by the eye. As was already mentioned in Section 1, the goal of this paper is only to present a mathematically complete approach to least square processing of images. Therefore we have restricted our experiments to the generation of raw data, i.e., the output of the filter with an additional linear scaling (gray levels stretched between 0 and 255) in the cases of differentiated images.
All the filters were of size $9 \times 9$. The same derivative is taken along both directions and thus, for example, the first derivative filter will not respond to horizontal or ver-


FIG. 1. Smoothing of the pebbles image. Upper left: Original. Upper right: Zero-degree polynomials. Lower left: Second-degree polynomials. Lower right: Fourth-degree polynomials. The pebbles images are $128 \times 128$, magnified to $256 \times 256$. The filter size is $9 \times 9$ in all these and subsequent images.
tical edges since one of its components is always zero. We did not investigate the effect of "nonisotropic" filters, in which the supports of the filter and the degrees of the fitted polynomials differ along the $n$ and $m$ coordinates of the lattice. These are the oriented edge detectors (among them the discrete versions of the Marr-Hildreth and the Canny detectors) and are obtained by combining pure smoothing components with smoothed differentia-


FIG. 2. Smoothing of the vase image. Upper left: Original. Upper right: Zero-degree polynomials. Lower left: Second-degree polynomials. Lower right: Fourth-degree polynomials. The vase images are 158 $\times 158$, magnified to $256 \times 256$.
tion ones. The separability along the two lattice coordinate axes (13) makes building of such edge detectors immediate.
A serious amount of blurring of the images appears when zero-order polynomials (we will refer from now on only to $L^{-}$) are fitted to the data through convolution with the smoothing filter (upper right, Figs. 1 and 2). However, when the degree of the fitted polynomials is increased to two (lower left, Figs. 1 and 2) most of the details remain sharp. An additional increase of the polynomial degree to four does not seem to yield further improvements (lower right, Figs. 1 and 2).

The images were also corrupted by zero mean, white Gaussian noise-the pebbles image with noise having standard deviation 50 (upper left, Fig. 3), and the vase image with noise having standard deviation 30 (upper left, Fig. 4). As expected the increase in the degree of the smoothing polynomials also increased the amount of noise retained. The second-degree polynomials (lower left, Figs. 3 and 4) appear to optimize the trade-off between the smoothness of the background and the sharpness of the details.
Different derivatives of the pebbles image are shown in Fig. 5. At the upper left, the first derivative obtained with first-degree polynomials is given. As was discussed above the horizontal and vertical components of the filter are approximations of the first derivative of a Gaussian with standard deviation $\sigma=\sqrt{2}$. The gray level changes are, however, much more sharply defined if third-order polynomials are employed for the smoothing part of the filter (upper right, Fig. 5). Similar improvement can be observed for the smoothed second-derivative filters. In


FIG. 3. Smoothing of the noisy pebbles image. Upper left: Original corrupted by zero mean, white, Gaussian noise with standard deviation 50. Upper right: Zero-degree polynomials. Lower left: Second-degree polynomials. Lower right: Fourth-degree polynomials.


FIG. 4. Smoothing of the noisy vase image. Upper left: Original corrupted by zero mean, white, Gaussian noise with standard deviation 30. Upper right: Zero-degree polynomials. Lower left: Second-degree polynomials. Lower right: Fourth-degree polynomials.
the lower left of Fig. 5 the second-order derivative is shown computed with second-degree polynomials. In this case the horizontal and vertical components of the filter are approximations of the second derivative of a Gaussian with standard deviation $\sigma=\sqrt{2}$. In the lower right the same derivative is computed with fourth-degree smoothing polynomials.


FIG. 5. Derivatives of the pebbles image. Upper left: First-derivative, first-degree polynomials. Upper right: First-derivative, third-degree polynomials. Lower left: Second-derivative, second-degree polynomials. Lower right: Second-derivative, fourth-degree polynomials.

The upper left of Fig. 6 shows the first derivative of the vase image computed after smoothing with third-order polynomials. The edges delineating the vase can be recognized. The second derivative based on fourth-degree polynomials (upper right), the third derivative based on third-degree polynomials (lower left), and the fourth derivative based on fourth-degree polynomials (lower right) all retained the highlight feature at the center of the vase showing the step-like nature of this feature.

## 6. DISCUSSION

We have presented a systematic approach to least square approximation of images and of their derivatives. We have shown that by choosing orthonormal polynomial bases for the description of the neighborhood within which the computation is performed, the result can be achieved by convolution with filters expressed in closed form. We have proved that only every second underlying surface degree is meaningful in processing. A spline polynomial decomposition-based technique was also employed to obtain the solution in the more general case of minimal-curvature filters where a regularization term is also present in the minimization [32, 33].
One application of the low-order smoothed differential filters is in edge detection. The numerical differentiation along one coordinate direction combined with smoothing along the other direction yields oriented edge detection masks. Directional derivatives can also be obtained when


FIG. 6. Derivatives of the vase image. Upper left: First-derivative, third-degree polynomials. Upper right: Second-derivative, fourth-degree polynomials. Lower left: Third-derivative, third-degree polynomials. Lower right: Fourth-derivative, fourth-degree polynomials.
the values of the partial derivatives are available [10]. We have shown that the use of Gaussian smoothed derivatives (in the Canny and the Marr-Hildreth edge detectors) implies additional assumptions about the degree of the underlying polynomial surface, at least when an $l_{2}$ norm is employed.

The degree of the underlying polynomial is the most important parameter of the filters. In addition to describing the underlying polynomial surface it also controls the amount of achieved smoothing, as the experimental results have also shown. Given the data and a family of models (e.g., polynomial surfaces), the model order problem deals with the selection of the model best fitting the data. There is no unique solution for the problem and it is an active research area both in the context of classic [22, Chap. 7] and robust [9, pp. 366-367] statistical techniques. The final decision is dictated by the desired tradeoff between noise reduction and sharpness of details. Note also that for a given neighborhood size all the filters require the same amount of computation since none of the parameters of the fitted polynomials are computed explicitly.

Even higher-order polynomial surfaces cannot model discontinuities (edges) in an image. The image data are piecewise; that is, they are composed of several surfaces separated by model discontinuities (edges). Application of least-squares-based surface estimation techniques to piecewise data violates the basic assumptions of least squares and distortions are always introduced. Discontinuity preserving smoothing procedures must use either adaptive least squares methods (e.g., [24]) and/or high breakdown point robust estimators [21]. However, if the
data are homogeneous (consist of only one surface) least squares often yields the optimum solution.

In our approach the discrete nature of the image has a central role and all the computations are restricted to the sites of the sampling lattice. We believe that the discreteness of the data must be taken into account when developing algorithms for computer vision. Transition from a continuous formulation of the problem to discrete algorithms may not always yield correct results. For example, recently Lindeberg [17] has shown that scale-space properties in the continuous domain can be violated after sampling. The filters proposed in this paper allow treatment of least square problems directly in the discrete realm.

## APPENDIX A

## Smoothed Differentiation Filters from Chebyshev Polynomials

In this appendix the expressions for smoothed differentiation filters derived from Chebyshev polynomials up to degree five are given. The filters were obtained from (12) with the help of the Macsyma program package. The unit weights of the Chebyshev polynomials yield unweighted smoothing of the input. Convolution with a filter supplies the value of the derivative in the center of the neighborhood.

## Notation

Support of the filter: $-N, \ldots, 0, \ldots, N$.
Degree of the fitted polynomial: $L$.
Order of differentiation: $p$.
The corresponding filter is $h_{C}(n)$.

Unweighted Smoothing Filters
$L=0$ or $1, p=0$.

$$
\begin{equation*}
h_{C}(n)=\frac{1}{2 N+1} . \tag{A1}
\end{equation*}
$$

$L=2$ or $3, p=0$.

$$
\begin{equation*}
h_{C}(n)=-\frac{3\left[5 n^{2}-\left(3 N^{2}+3 N-1\right)\right]}{(2 N-1)(2 N+1)(2 N+3)} . \tag{A2}
\end{equation*}
$$

$L=4$ or $5, p=0$.

$$
\begin{equation*}
h_{C}(n)=\frac{15\left[63 n^{4}-35\left(2 N^{2}+2 N-3\right) n^{2}+\left(15 N^{4}+30 N^{3}-35 N^{2}-50 N+12\right)\right]}{4(2 N-3)(2 N-1)(2 N+1)(2 N+3)(2 N+5)} . \tag{A3}
\end{equation*}
$$

Unweighted Smoothing, First-Order Differentiation Filters
$L=1$ or $2, p=1$.

$$
\begin{equation*}
h_{C}(n)=-\frac{3 n}{N(N+1)(2 N+1)} . \tag{A4}
\end{equation*}
$$

$L=3$ or $4, p=1$.

$$
\begin{equation*}
h_{C}(n)=\frac{5\left[7\left(3 N^{2}+3 N-1\right) n^{3}-5\left(3 N^{4}+6 N^{3}-3 N+1\right) n\right]}{(N-1) N(N+1)(N+2)(2 N-1)(2 N+1)(2 N+3)} \tag{A5}
\end{equation*}
$$

$L=5$ or $6, p=1$.

$$
\begin{align*}
h_{C}(n)= & \left(-\frac{21}{4}\right) \frac{33\left(15 N^{4}+30 N^{3}-35 N^{2}-50 N+12\right) n^{5}}{(N-2)(N-1) N(N+1)} \\
& \frac{-105\left(6 N^{6}+18 N^{5}-15 N^{4}-60 N^{3}+17 N^{2}+50 N-12\right) n^{3}}{(N+2)(N+3)(2 N-3)(2 N-1)}  \tag{A6}\\
& \frac{+7\left(25 N^{8}+100 N^{7}-50 N^{6}-500 N^{5}-95 N^{4}+760 N^{3}+180 N^{2}-300 N+72\right) n}{(2 N+1)(2 N+3)(2 N+5)}
\end{align*}
$$

Unweighted Smoothing, Second-Order Differentiation Filters
$L=2$ or $3, p=2$.

$$
\begin{equation*}
h_{C}(n)=\frac{30\left[3 n^{2}-N(N+1)\right]}{N(N+1)(2 N-1)(2 N+1)(2 N+3)} \tag{A7}
\end{equation*}
$$

$L=4$ or $5, p=2$.

$$
\begin{align*}
h_{C}(n)= & \left(-\frac{105}{2}\right) \frac{15\left(6 N^{2}+6 N-5\right) n^{4}-21\left(4 N^{4}+8 N^{3}-4 N^{2}-8 N+5\right) n^{2}}{(N-1) N(N+1)(N+2)(2 N-3)} \\
& \frac{+5(N-1) N(N+1)(N+2)\left(2 N^{2}+2 N-3\right)}{(2 N-1)(2 N+1)(2 N+3)(2 N+5)} \tag{A8}
\end{align*}
$$

Unweighted Smoothing, Third-Order Differentiation Filters
$L=3$ or $4, p=3$.

$$
\begin{equation*}
h_{C}(n)=-\frac{210\left[5 n^{3}-\left(3 N^{2}+3 N-1\right) n\right]}{(N-1) N(N+1)(N+2)(2 N-1)(2 N+1)(2 N+3)} \tag{A9}
\end{equation*}
$$

$L=5$ or $6, p=3$.

$$
\begin{align*}
h_{C}(n)= & \left(\frac{945}{2}\right) \frac{77\left(2 N^{2}+2 N-3\right) n^{5}-15\left(12 N^{4}+24 N^{3}-28 N^{2}-40 N+39\right) n^{3}}{(N-2)(N-1) N(N+1)(N+2)(N+3)} \\
& +\frac{7\left(6 N^{6}+18 N^{5}-15 N^{4}-60 N^{3}+17 N^{2}+50 N-12\right) n}{(2 N-3)(2 N-1)(2 N+1)(2 N+3)(2 N+5)} \tag{A10}
\end{align*}
$$

Unweighted Smoothing, Fourth-Order Differentiation Filter
$L=4$ or $5, p=4$.

$$
\begin{equation*}
h_{C}(n)=\frac{1890\left[35 n^{4}-5\left(6 N^{2}+6 N-5\right) n^{2}+3(N-1) N(N+1)(N+2)\right]}{(N-1) N(N+1)(N+2)(2 N-3)(2 N-1)(2 N+1)(2 N+3)(2 N+5)} \tag{A11}
\end{equation*}
$$

## Unweighted Smoothing, Fifth-Order Differentiation Filter

$L=5$ or $6, p=5$.

$$
\begin{align*}
h_{C}(n)= & -\frac{20790\left[63 n^{5}-35\left(2 N^{2}+2 N-3\right) n^{3}\right.}{(N-2)(N-1) N(N+1)(N+2)(N+3)}  \tag{A12}\\
& \frac{\left.+\left(15 N^{4}+30 N^{3}-35 N^{2}-50 N+12\right) n\right]}{(2 N-3)(2 N-1)(2 N+1)(2 N+3)(2 N+5)} .
\end{align*}
$$

## APPENDIX B

## Smoothed Differentiation Filters from Krawtchouk Polynomials

In this appendix the expressions for smoothed differentiation filters derived from Krawtchouk polynomials up to degree five are given. The filters were obtained from (12) with the help of the Macsyma program package. The expression for the sampled weight function

$$
\begin{equation*}
w(n)=\frac{1}{2^{2 N}}\binom{2 N}{N-n}=\frac{1}{2^{2 N}} \frac{(2 N)!}{(N-n)!(N+n)!} \tag{B1}
\end{equation*}
$$

appears in all the filters and is not given in its explicit form. Because of the presence of $w(n)$ the data can be regarded as weighted by the discrete approximation of a Gaussian with variance $N / 2$. Convolution with a filter returns the value in the center of the neighborhood.

## Notation

Support of the filter: $-N, \ldots, 0, \ldots, N$.
Degree of the fitted polynomial: $L$.
Order of differentiation: $p$.
The corresponding filter is $h_{K}(n)$.

## Gaussian-Weighted Smoothing Filters

$L=0$ or $1, p=0$.

$$
\begin{equation*}
h_{K}(n)=w(n) . \tag{B2}
\end{equation*}
$$

$L=2$ or $3, p=0$.

$$
\begin{equation*}
h_{K}(n)=-\frac{2 n^{2}-(3 N-1)}{2 N-1} w(n) . \tag{B3}
\end{equation*}
$$

$L=4$ or $5, p=0$.

$$
\begin{equation*}
h_{K}(n)=\frac{4 n^{4}-20(N-1) n^{2}+\left(15 N^{2}-25 N+6\right)}{2(2 N-3)(2 N-1)} w(n) . \tag{B4}
\end{equation*}
$$

Gaussian-Weighted Smoothing, First-Order Differentiation Filters
$L=1$ or $2, p=1$.

$$
\begin{equation*}
h_{K}(n)=-\frac{2 n}{N} w(n) . \tag{B5}
\end{equation*}
$$

$L=3$ or $4, p=1$.

$$
\begin{equation*}
h_{K}(n)=\frac{2\left[2(3 N-1) n^{3}-\left(15 N^{2}-15 N+4\right) n\right]}{3(N-1) N(2 N-1)} w(n) . \tag{B6}
\end{equation*}
$$

$L=5$ or $6, p=1$.

$$
\begin{align*}
h_{K}(n)= & \left(-\frac{1}{15}\right) \frac{4\left(15 N^{2}-25 N+6\right) n^{5}-20\left(21 N^{3}-63 N^{2}+56 N-12\right) n^{3}}{(N-2)(N-1) N} \\
& \frac{+\left(525 N^{4}-2100 N^{3}+2835 N^{2}-1480 N+276\right) n}{(2 N-3)(2 N-1)} w(n) . \tag{B7}
\end{align*}
$$

Gaussian-Weighted Smoothing, Second-Order Differentiation Filters
$L=2$ or $3, p=2$.

$$
\begin{equation*}
h_{K}(n)=\frac{4\left[2 n^{2}-N\right]}{N(2 N-1)} w(n) . \tag{B8}
\end{equation*}
$$

$L=4$ or $5, p=2$.

$$
\begin{equation*}
h_{K}(n)=\left(-\frac{4}{3}\right) \frac{4(3 N-2) n^{4}-2\left(24 N^{2}-39 N+17\right) n^{2}+15 N(N-1)^{2}}{(N-1) N(2 N-3)(2 N-1)} w(n) \tag{B9}
\end{equation*}
$$

Gaussian-Weighted Smoothing, Third-Order Differentiation Filters
$L=3$ or $4, p=3$.

$$
\begin{equation*}
h_{K}(n)=-\frac{8\left[2 n^{3}-(3 N-1) n\right]}{(N-1) N(2 N-1)} w(n) \tag{B10}
\end{equation*}
$$

$L=5$ or $6, p=3$.

$$
\begin{align*}
h_{K}(n)= & \frac{8\left[4(N-1) n^{5}-2\left(12 N^{2}-27 N+16\right) n^{3}\right.}{(N-2)(N-1) N} \\
& \frac{\left.+\left(21 N^{3}-63 N^{2}+56 N-12\right) n\right]}{(2 N-3)(2 N-1)} w(n) \tag{B11}
\end{align*}
$$

## Gaussian-Weighted Smoothing, Fourth-Order Differentiation Filter

$L=4$ or $5, p=4$.

$$
\begin{equation*}
h_{K}(n)=\frac{16\left[4 n^{4}-4(3 N-2) n^{2}+3(N-1) N\right]}{(N-1) N(2 N-3)(2 N-1)} w(n) \tag{B12}
\end{equation*}
$$

## Gaussian-Weighted Smoothing, Fifth-Order Differentiation Filter

$L=5$ or $6, p=5$.

$$
\begin{equation*}
h_{K}(n)=-\frac{32\left[4 n^{5}-20(N-1) n^{3}+\left(15 N^{2}-25 N+6\right) n\right]}{(N-2)(N-1)(2 N-3)(2 N-1)} w(n) \tag{B13}
\end{equation*}
$$

## REFERENCES

1. H. H. Barrow and J. M. Tenenbaum, Interpreting line drawings as three dimensional surfaces, Artificial Intelligence 17, 1981, 75-117.
2. A. P. Blicher, Edge Detection and Geometric Methods in Computer Vision, Technical Report AIM-352, Department of Computer Science, Stanford University, 1985.
3. M. U. A. Bromba and H. Ziegler, Efficient computation of polynomial smoothing digital filters, Anal. Chem. 51, 1979, 1760-1762.
4. M. U. A. Bromba and H. Ziegler, Explicit formula for filter function of maximally flat nonrecursive digital filters, Electron. Lett. 16, 1980, 905-906.
5. M. U. A. Bromba and H. Ziegler, Application hints for SavitzkyGolay digital smoothing filters, Anal. Chem. 53, 1981, 1583-1586.
6. P. Burt, Moment images, polynomial fit filters, and the problem of surface interpolation, in Proceedings, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Ann Arbor, MI, June 5-9, 1988, pp. 144-152.
7. J. Canny, A computational approach to edge detection, IEEE Trans. Pattern Anal. Mach. Intell. PAMI-8, 1986, 679-698.
8. H. E. H. Greenleaf, Curve approximation by means of functions analogous to the Hermite polynomials, Ann. Math. Statist. 3, 1932, 204-256.
9. F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel, Robust Statistics: An Approach Based on Influence Functions, Wiley, New York, 1986.
10. R. M. Haralick, Digital step edges from zero crossing of second directional derivatives, IEEE Trans. Pattern Anal. Mach. Intell. PAMI-6, 1984, 58-68.
11. R. Hartley, A Gaussian-weighted multiresolution edge detector, Computer. Vision Graphics Image Process. 30, 1985, 70-83.
12. M. Hashimoto and J. Sklansky, Multiple-order derivatives for detecting local image characteristics, Comput. Vision Graphics Image Process. 39, 1987, 28-55.
13. F. B. Hildebrand, Introduction to Numerical Analysis, McGrawHill, New York, 1956.
14. R. A. Hummel, Feature detection using basis functions, Comput. Graphics Image Process. 9, 1979, 40-55.
15. M. F. Hueckel, A local visual operator which recognizes edges and lines, J. Assoc. Comput. Mach. 20, 1973, 634-647.
16. R. Lenz, Optimal filters for the detection of linear patterns in 2-D and higher dimensional images, Pattern Recognition 20, 1987, 163172.
17. T. Lindeberg, Scale-space for discrete signals, IEEE Trans. Pattern Anal. Mach. Intell. PAMI-12, 1990, 234-254.
18. H. H. Madden, Comments on the Savitzky-Golay convolution method for least-squares fit smoothing and differentiation of digital data, Anal. Chem. 50, 1978, 1383-1386.
19. D. Marr and E. Hildreth, Theory of edge detection, Proc. Roy. Soc. London B 207, 1980, 187-217.
20. P. Meer, E. S. Baugher, and A. Rosenfeld, Frequency domain
analysis and synthesis of image pyramid generating kernels, IEEE Trans. Pattern Anal. Mach. Intell. PAMI-9, 1987, 512-522.
21. P. Meer, D. Mintz, and A. Rosenfeld, Least median of squares based robust analysis of image structure, in Proceedings, DARPA Image Understanding Workshop, Pittsburgh, PA, September 1990, pp. 231-254.
22. D. C. Montgomery and E. A. Peck, Introduction to Linear Regression Analysis, Wiley, New York, 1982.
23. A. Papoulis, Probability, Random Variables, and Stochastic Processes, McGraw-Hill, New York, 1972.
24. R.-H. Park and P. Meer, Multiresolution Adaptive Least Squares Smoothing of Images, CAR-TR-513, Computer Vision Laboratory, University of Maryland, College Park, 1990.
25. K. Paton, Picture description using Legendre polynomials, Comput. Graphics Image Process. 4, 1975, 40-54.
26. J. Ponce and M. Brady, Toward a surface primal sketch, A. I. Memo 824, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 1985.
27. A. Savitzky and M. J. E. Golay, Smoothing and differentiation of data by simplified least squares procedures, Anal. Chem. 36, 1964, 1627-1639.
28. H. W. Schüssler and P. Steffen, Some advanced topics in filter design, in Advanced Topics in Signal Processing (J. S. Lim and A. V. Oppenheim, Eds.), pp. 416-491, Prentice-Hall, Englewood Cliffs, NJ, 1988.
29. P. Steffen, On digital smoothing filters: A brief review of closed form solutions and two new filter approaches, Circuits Systems Signal Process. 5, 1986, 187-210.
30. J. Steinier, Y. Termonia, and J. Deltour, Comments on smoothing and differentiation of data by simplified least square procedure, Anal. Chem. 44, 1972, 1906-1909.
31. G. Szegö, Orthogonal Polynomials, American Mathematical Society Colloquium Publications, 4th ed., Vol. XXIII, Providence, RI, 1975.
32. I. Weiss, Image smoothing and differentiation with minimal curvature filters, CAR-TR-470, Computer Vision Laboratory. University of Maryland, College Park, 1989.
33. I. Weiss, High order differentiation filters that work, CAR-TR-545, Computer Vision Laboratory. University of Maryland, College Park, 1991.


PETER MEER was born in Oradea, Romania, on February 14, 1949. He received the Dipl. Engn. degree from the Bucharest Polytechnic Institute, Bucharest, Romania, in 1971, and the D.Sc. degree from the Technion, Israel Institute of Technology, Haifa, Israel, in 1986, both in electrical engineering. From 1971 to 1979 he was with the Computer Research Institute, Cluj, Romania, working on R\&D of digital hardware. Between 1986 and 1990 he was an assistant research scientist at the Center for Automation Research, University of Maryland at College Park. In 1991 he joined the Department of Electrical and Computer Engineering, Rutgers University, Piscataway, New Jersey, as an assistant professor. His research interests include application of estimation techniques and probabilistic algorithms to machine vision problems.

I. WEISS received his Ph.D. in physics from the Tel-Aviv University, Israel. Before joining the Center for Automation Research of the University of Maryland, he was a research scientist at the Courant Institute of Mathematics of New York University and at the Massachusetts Institute of Technology. His current research interests are computer vision, image processing, pattern recognition, and robotics.


[^0]:    ${ }^{1}$ The support of the National Science Foundation and the U.S. Air Force of Scientific Research under NSF Grant DCR-86-03723, and of the Office of Naval Research under Grant N00014-88-K-0348, is gratefully acknowledged.

