Filters involving derivatives with application to reconstruction from scanned halftone images

Forchhammer, Søren; Jensen, Kim S.

Published in:
IEEE Transactions on Image Processing

Link to article, DOI:
10.1109/83.370673

Publication date:
1995

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Filters Involving Derivatives with Application to Reconstruction from Scanned Halftone Images

Søren Forchhammer and Kim S. Jensen

Abstract—This paper presents a method for designing finite impulse response (FIR) filters for samples of a 2-D signal, e.g., an image, and its gradient. The filters, which are called blended filters, are decomposable in three filters, each separable in 1-D filters on subsets of the data set.

Optimality in the minimum mean square error sense (MMSE) of blended filtering is shown for signals with separable autocorrelation function. Relations between correlation functions for signals and their gradients are derived. Blended filters may be composed from FIR Wiener filters using these relations. Simple blended filters are developed and applied to the problem of gray value image reconstruction from bilevel (scanned) clustered-dot halftone images, which is an application useful in the graphic arts. Reconstruction results are given, showing that reconstruction with higher resolution than the halftone grid is achievable with blended filters.

I. INTRODUCTION

This paper presents and treats a general method for designing finite impulse response (FIR) filters for samples of a signal and its gradient in two dimensions. The filters, which are called blended filters, are composed of three filters, each separable in 1-D filters on subsets of the samples of the signal and its gradient. This enables the use of filter techniques for 1-D data in the design of blended filters.

Blended filters are applied to the problem of reconstruction from (scanned) halftone images. Halftoning of images is used in the graphic arts to render gray value images as bilevel images. The bilevel images are basically composed of halftone dots, with (relative) areas reflecting the gray values at the specific positions. Besides this gray value information, the shape and (relative) position of the dots might give additional information correlating to the gradient of the original gray value image.

This additional information at halftone subdot level has been used in previous work on halftones. For data compression purposes, methods to determine the area coverage of the halftone (sub)dots have been devised [1], [2]. To utilize the (sub)dot information for image reconstruction, subdot areas were used to estimate gradient values. The reconstruction was thereafter achieved using simple polynomial interpolation [3].

In this paper, emphasis is put on filtering aspects of using gradient samples. Blended filters that are a more general class of filters for samples of a signal and its gradient are defined, and properties of these filters are investigated.

In Section II, blended filters are defined, and a sampling theorem for a specific bandlimitation in 2-D is given. Section III treats the problem of designing blended filters with finite impulse response (FIR). Results are given that may be used for composing blended filters from FIR Wiener filters. The case of a signal with separable autocorrelation function is treated in detail, and conditions under which blended filters are optimal are given. Application to reconstruction from halftone images is devised in Section IV, and specific reconstruction filters are developed. In Section V, numerical values of reconstruction errors for different blended filters on a test image are given, and images reconstructed from a scanned halftone image are presented.

II. BLENDED FILTERS

This section presents a fast method for designing and implementing filters for samples of a signal \( f(x,y) \), e.g., an image, and its gradient \( \nabla f(x,y) = \{f_x(x,y); f_y(x,y)\} \). The samples are organized in a regular grid with integer coordinates \((m,n)\) in the \((x,y)\)-coordinate system. The samples \( f(m,n) \) of the signal are also referred to as amplitude samples to distinguish them from the gradient samples.

A. Definition of Blended Filters

Filters in 2-D are often realized as separable filters that are decomposable in two 1-D filters to speed up the implementation. For samples of the function \( f(m,n) \) and gradient \( f_x(m,n); f_y(m,n) \), a separable filter for each of the three components may be used. A closely related possibility is to use a blended filter, which is also defined using 1-D filters.

Definition: A blended filter with samples of a signal \( f(m,n) \) and its gradient \( f_x(m,n); f_y(m,n) \) as input has the output

\[
\hat{f}(x,y) = f_h(x,y) + f_s(x,y) - f_b(x,y) \quad (1)
\]

where \( f_h \), \( f_b \), and \( f_s \) are intermediate functions given by

\[
f_h(x,y) = f_h(x,y) + f_s(x,y) - f_b(x,y) \quad (2a)
\]

\[
f_h(x,y) = f_h(x,y) + f_s(x,y) - f_b(x,y) \quad (2b)
\]

\[
f_h(x,y) = f_h(x,y) + f_s(x,y) - f_b(x,y) \quad (2c)
\]
Authorized licensed use limited to: Danmarks Tekniske Informationscenter. Downloaded on February 9, 2010 at 11:32 from IEEE Xplore. Restrictions apply.
The ideal reconstruction filters above (3) and (4) may be used as the 1-D filters of a symmetric blended filter. The
reconstruction filter for amplitude-derivative samples in 1-D
(see (4)) is used for samples of the function and the partial
derivative in the \( x/y \) direction to obtain \( f_\theta/(f_\theta) \) along
the grid lines in (2b) and (2c). The reconstruction filter for
samples of only the function in (3) is used as \( h_{\omega_0} \) and \( h_{\omega_0} \)
in (2a)-(2c). This way, exact reconstruction for a 2-D signal
with bandlimitation as shown in Fig. 2 is obtained [3].

Petersen and Middleton [7] gave a detailed treatment of
reconstruction from the samples of amplitude and gradient
of an \( n \)-dimensional stochastic field. They showed that exact
reconstruction (in the zero mean square error sense) is possible
if the spectral images induced by sampling do not overlap more
than \( n+1 \) times. The blended filter obtained above using the
filters (see (3) and (4)) resolves the 3 times spectral overlap
for the bandlimitation specified by Fig. 2 in a simple manner
[3], [8].

If the ideal lowpass filter of (3) is exchanged with an ideal
bandpass filter with passband \( [\omega_0/2, \omega_0] \), the corresponding
symmetric blended filter yields exact reconstruction for the
bandpass limitation with \( \omega_0/2 < \max\{\omega_1, \omega_2\} < \omega_0 \) [8].

III. DESIGNING FIR BLENDED FILTERS

In this section, the problem of designing FIR blended filters
for practical reconstruction using samples of the 2-D signal
and its gradient is considered.

In image processing, FIR filters are often preferred because
they enable the design of linear/zero phase filters. This attribute
is important [9] because it prevents frequencies in phase at an
edge from being forced out of phase, thus distorting the edge
and blurring the image.

A number of techniques for designing 1-D and 2-D filters
[10] may be extended to the case of 2-D amplitude-gradient
filtering. Examples of this are windowing, polynomial interpola-
tion, Wiener filters, and frequency sampling. One reason for
using a blended filter is simple design and implementation.
Using a blended filter, the problem of designing the filters
may immediately be reduced to a problem of designing 1-
D (amplitude-derivative) filters. An example of this, using
the windowing technique, is to truncate the ideal impulse
responses of (3) and (4). However, generalizing a 1-D tech-
nique that is optimal for amplitude samples in some sense
does not imply that the corresponding 2-D blended filter is
optimal. Truncating the ideal impulse response(s) does not
give a min. mean square error (MMSE) approximation of the
impulse response for a 2-D blended filter as it does for 1-D
(amplitude-only) filter.

In the following, the use of polynomial interpolation for
designing blended filters is briefly described. Thereafter, the
use of FIR Wiener filters is treated, emphasizing signals with
separable autocorrelation function. The optimality of blended
Wiener filters is also examined.

A. Polynomial Interpolation

A number of techniques for designing 1-D and 2-D filters
[10] may be extended to the case of 2-D amplitude-gradient
filtering. Examples of this are windowing, polynomial interpola-
tion, Wiener filters, and frequency sampling. One reason for
using a blended filter is simple design and implementation.

The ideal reconstruction filters above (3) and (4) may be
used as the 1-D filters of a symmetric blended filter. The
reconstruction filter for amplitude-derivative samples in 1-D
(see (4)) is used for samples of the function and the partial
derivative in the \( x/y \) direction to obtain \( f_\theta/(f_\theta) \) along
the grid lines in (2b) and (2c). The reconstruction filter for
samples of only the function in (3) is used as \( h_{\omega_0} \) and \( h_{\omega_0} \)
in (2a)-(2c). This way, exact reconstruction for a 2-D signal
with bandlimitation as shown in Fig. 2 is obtained [3].

Petersen and Middleton [7] gave a detailed treatment of
reconstruction from the samples of amplitude and gradient
of an \( n \)-dimensional stochastic field. They showed that exact
reconstruction (in the zero mean square error sense) is possible
if the spectral images induced by sampling do not overlap more
than \( n+1 \) times. The blended filter obtained above using the
filters (see (3) and (4)) resolves the 3 times spectral overlap
for the bandlimitation specified by Fig. 2 in a simple manner
[3], [8].

If the ideal lowpass filter of (3) is exchanged with an ideal
bandpass filter with passband \( [\omega_0/2, \omega_0] \), the corresponding
symmetric blended filter yields exact reconstruction for the
bandpass limitation with \( \omega_0/2 < \max\{\omega_1, \omega_2\} < \omega_0 \) [8].

III. DESIGNING FIR BLENDED FILTERS

In this section, the problem of designing FIR blended filters
for practical reconstruction using samples of the 2-D signal
and its gradient is considered.

In image processing, FIR filters are often preferred because
they enable the design of linear/zero phase filters. This attribute
is important [9] because it prevents frequencies in phase at an
declaration from being forced out of phase, thus distorting the edge
and blurring the image.

A number of techniques for designing 1-D and 2-D filters
[10] may be extended to the case of 2-D amplitude-gradient
filtering. Examples of this are windowing, polynomial interpola-
tion, Wiener filters, and frequency sampling. One reason for
using a blended filter is simple design and implementation.
Using a blended filter, the problem of designing the filters
may immediately be reduced to a problem of designing 1-
D (amplitude-derivative) filters. An example of this, using
the windowing technique, is to truncate the ideal impulse
responses of (3) and (4). However, generalizing a 1-D tech-
nique that is optimal for amplitude samples in some sense
does not imply that the corresponding 2-D blended filter is
optimal. Truncating the ideal impulse response(s) does not
give a min. mean square error (MMSE) approximation of the
impulse response for a 2-D blended filter as it does for 1-D
(amplitude-only) filter.

In the following, the use of polynomial interpolation for
designing blended filters is briefly described. Thereafter, the
use of FIR Wiener filters is treated, emphasizing signals with
separable autocorrelation function. The optimality of blended
Wiener filters is also examined.

A. Polynomial Interpolation

A number of techniques for designing 1-D and 2-D filters
[10] may be extended to the case of 2-D amplitude-gradient
filtering. Examples of this are windowing, polynomial interpola-
tion, Wiener filters, and frequency sampling. One reason for
using a blended filter is simple design and implementation.
Using a blended filter, the problem of designing the filters
may immediately be reduced to a problem of designing 1-
D (amplitude-derivative) filters. An example of this, using
the windowing technique, is to truncate the ideal impulse
responses of (3) and (4). However, generalizing a 1-D tech-
The best linear estimator with respect to MMSE using $m$ observations is given by the $m$ well-known equations [15]

$$
\sum_{k=1}^{m} E[w_k y_k] h_k = E[w_k], \quad j \in \{1, m\}.
$$

(6)

If the variables are gaussian distributed, the estimator is optimal. The observations may be of any dimension, and gradients may also be incorporated. Therefore, the case of estimation from samples of a 2-D signal, and its gradient is just one special case. If the observations are from stationary process(es), the mean values of (6) may be expressed by the autocorrelations (and cross correlations). The resulting filter is called a (FIR) Wiener filter [12], [16]. The term signal will be used because the results only depend on the second moments and may be applied to deterministic signals as well as wide sense stationary processes.

To derive expressions for the relations of the correlation functions of (6), we may describe the derivative $f'(t)$ as a linear functional of $f(t)$ obtained from the impulse response $h(t)$. For $f(t)$ bandlimited, the corresponding transfer function is $H(\omega) = iw[5]$. If no noise is added in the differentiation process, relations between the spectras of the observation and its derivatives may be derived using the transfer function. The correlation functions for (6) may be obtained from the corresponding spectras. The correspondence is denoted by $\leftrightarrow$. The indices of the (cross) spectras ($S$) and correlation functions ($r$) refer to the process(es) involved.

In one dimension, we get from the theory of linear systems for the stationary process $x$ with derivative $x'$ [5]:

$$
S_r(w) = \omega^2 S_x(w) \leftrightarrow r_x(t) = -r_x'(t) \quad (7)
$$

$$
S_{xx}(w) = -i\omega S_x(w) \leftrightarrow r_{xx}(t) = -r_x'(t). \quad (8)
$$

The results from linear systems also apply in 2-D for the partial derivatives [5], giving the relations of the spectras and thereby the autocorrelation functions

$$
S_{r_{x}}(\omega_1, \omega_2) = \omega_1^2 S_{r_{x}}(\omega_1, \omega_2) \leftrightarrow r_{r_{x}}(x, y) = -\frac{\delta^2 r_{r_{x}}(x, y)}{\delta x^2}. \quad (9)
$$

$$
S_{r_{y}}(\omega_1, \omega_2) = \omega_2^2 S_{r_{y}}(\omega_1, \omega_2) \leftrightarrow r_{r_{y}}(x, y)
= \frac{\delta^2 r_{r_{y}}(x, y)}{\delta y^2}. \quad (10)
$$

$$
S_{r_{r_{x}}r_{r_{y}}}(\omega_1, \omega_2) = -i\omega_1 S_{r_{r_{x}}}r_{y} = -r_{r_{x}}r_{r_{y}}(x, y) \quad (11)
$$

$$
S_{r_{r_{x}}r_{r_{y}}}(\omega_1, \omega_2) = -i\omega_2 S_{r_{r_{y}}}r_{x} = -r_{r_{y}}r_{r_{x}}(x, y) \quad (12)
$$

Writing $r_{r_{x}}r_{r_{y}}$ as the output of a linear system with input $r_{r_{x}}r_{y}$ gives the transfer function $H(\omega_1, \omega_2) = i\omega_2/i\omega_1$, which gives

$$
S_{r_{r_{x}}r_{r_{y}}}(\omega_1, \omega_2) = (i\omega_2/i\omega_1) S_{r_{r_{x}}}r_{r_{y}}(\omega_1, \omega_2)
= -\frac{i\omega_1 i\omega_2}{i\omega_1} S_{r_{r_{x}}}r_{r_{y}}(\omega_1, \omega_2) \leftrightarrow r_{r_{x}}r_{r_{y}}(x, y)
= -\frac{\delta^2 r_{r_{x}}r_{r_{y}}(x, y)}{\delta x \delta y}. \quad (13)
$$

The relations between the correlation functions in one dimension (see (7) and (8)) were derived in [17] differentiating the expected values without the use of the transfer functions. In the same way, the correlation results in 2-D (see (9)-(12)) may be derived without the transfer functions and, thereby, the assumption of bandlimitation. The cross correlation $r_{r_{x}}r_{r_{y}}$ of the derivatives may also be derived this way:

$$
E \left\{ f(x_1 + \epsilon, y_1) - f(x_1, y_1) \right\}
= r_{r_{x}}r_{r_{y}}(x_1 - x_2, y_1 - y_2)
= \frac{\delta^2 r_{r_{x}}r_{r_{y}}(x_1, y_1)}{\delta x_1 \delta y_1}. \quad (14)
$$

With $x_1 - x_2 = \tau_1$, $y_1 - y_2 = \tau_2$, and $\epsilon \to 0$, (14) gives

$$
\frac{\delta r_{r_{x}}r_{r_{y}}(\tau_1, \tau_2)}{\delta \tau_1} = -\frac{\delta^2 r_{r_{x}}r_{r_{y}}(\tau_1, \tau_2)}{\delta \tau_1 \delta \tau_2}. \quad (15)
$$

which is the result of the right-hand side of (13).

1) Correlation Functions for Signals with Separable Autocorrelation: For a signal with separable autocorrelation function

$$
r_{r_{x}}(x, y) = r_{r_{x}}(x) r_{r_{y}}(y) \quad (16)
$$

the 2-D process $f$ may be decomposed in 1-D processes in the $x$ and $y$ directions with derivatives $x'$ and $y'$, respectively. In this case, there are simple relations between the correlation functions.

For $r_{r_{x}}(x, y)$ separable, comparing the structure of the derivatives (1-D) and the partial derivatives (2-D) in (7)-(13), the following relations are obtained:

$$
r_{r_{x}}r_{r_{y}}(x, y) = r_{r_{x}}(x) r_{r_{y}}(y) \quad (17)
$$

$$
r_{r_{x}}r_{r_{y}}(x, y) = r_{r_{x}}(x) r_{r_{y}}(y) \quad (18)
$$

$$
r_{r_{x}}(x, y) = r_{r_{x}}(x) r_{r_{y}}(y) \quad (19)
$$

$$
r_{r_{x}}(x, y) = r_{r_{x}}(x) r_{r_{y}}(y) \quad (20)
$$

$$
r_{r_{x}}r_{r_{y}}(x, y) = -r_{r_{x}}(x) r_{r_{y}}(y) \quad (21)
$$

The relations between the correlation functions may be used when setting up (6), describing the Wiener filter. In the following section, they are used to derive results on blended filters for signals with separable autocorrelation.

2) Optimal Blended Filters for Separable Autocorrelation Functions: In the following, the optimality of blended filters and the intermediate functions $f_{n}$, $f_{x}$, and $f_{r}$ is examined.

Equation (6) may be written in matrix form. For stationary processes, the expected values are described by correlation values. The matrix of (6) is written with capital bold and the vectors with lower-case bold. The subscripts refer to the processes in the respective directions. Consider the matrices $A$ and $B$ of dimensions $p$ by $q$ and $m$ by $n$ respectively. The Kronecker product of $A$ and $B$, written $A \otimes B$, is defined as the $p \times q \times m \times n$ matrix $(a_{ij} B)$ [16]. Consider Wiener filters for samples of the signal only. If both the left-side matrix $R_{r} ( = R_{r_{x}} \otimes R_{r_{y}})$ and the right-side vector $r_{f}$
the MMSE linear interpolation of a function with separable autocorrelation function. The result applies for a rectangular region of support for each of the components given by \( w_r \). \( f_r \) and \( f_c \) are the MMSE linear solutions given by Lemma 1. \( f_r \) is the separable MMSE linear solution (22) for the samples of \( f \).

**Proof:** See Appendix A for proof.

The autocorrelation functions given by \( r_s(x, y) = e^{-(a(x^2 + y^2))} \), where \( a \) and \( c \) are constants, are an example of autocorrelation functions for which the Lemma and the Theorem may be applied. \( r_s \) describes the only rotation invariant separable functions. For bandlimited signals, the results may be applied as the derivatives of the signal and the autocorrelation exist, thereby satisfying the assumptions.

3) Composite Wiener Filtering: Images are, in general, not stationary. A composite model is a better description. In this case, the image (model) is composed of a number of statistically distinct stationary objects belonging to a set of \( K \) classes. This model leads to a simple extension of the FIR Wiener filter, which adapts to the image locally.

Lebedev and Mirkin [19] introduced a filter weighting the FIR Wiener filter \( f_s(x, y) \) of each class \( \theta \). The weighting functions \( p(\theta | m(x, y)) \) are the conditional probability for \( f(x, y) \) that belong to the class \( \theta \), given the data within the window of the filter \( m(x, y) \). The composite estimate \( f(x, y) \), given by weighting the Wiener estimates of each class is

\[
\hat{f}(x, y) = \sum_{\alpha=1}^{K} p(\theta | m(x, y)) f_{\theta}(x, y). \tag{24}
\]

Under the assumptions of Gaussian distributions of the image and additive and independent noise within each class, over the filter window, Lebedev and Mirkin showed that this estimate is optimal (MMSE) for smoothing an image. In the next section, the estimate (24) is modified by letting \( m(x, y) \) be a vector function that operates on the data within a window in the neighborhood of \( f(x, y) \). This will be called a composite Wiener filter.

**IV. APPLICATION TO RECONSTRUCTION FROM HALFTONES**

In this section, application of amplitude-gradient filtering to the problem of reconstruction from scanned halftone images is described. A scanner delivers a (high-resolution) bilevel image representing the halftone. This is to be converted into a (digital) gray value image at a (lower) resolution, which preferentially is higher than that of the halftone grid. One interesting point is the acquisition of gradient samples with information not contained by the amplitude samples.

Basically, halftone images are bilevel images where dot areas represent gray values, but there may be more information available than just the area of the dot halftone dot. That is, the halftone image may convey more information than a sampling at the halftone screen resolution. To design high-resolution reconstruction schemes for halftone images, it is therefore necessary to take a closer look at the screening method used.

In conventional halftoning, the screen function is added to the gray value image and thresholded in a photomechanical process. Threshold screening (or electronic halftoning [20]) is a digital equivalent of conventional screening. The gray
value image function \( f(x, y) \) is thresholded with the periodic threshold (or screen) function to create the resulting bilevel image. Considering clustered dots, the transitions in the bilevel halftone image represent the crossings of the image function and a threshold (or screen) function (Fig. 4). This way, the halftone dots are shifted relative to the grid point according to the gradient of the image function across the dot. Clustered dot halftoning, as opposed to split dot dithering, is used in offset printing in the graphic arts industry. Reconstruction from halftone images may be used 1) to input (old) halftone material, 2) for data compression, and 3) for offset to gravure conversion.

The problem of reconstruction from halftones may be described as a reversal of the nonlinear screening process. Using a (nonlinear) estimate of the gradient (25) transforms the problem into a problem of reconstruction from (amplitude-gradient) samples in a regular grid. This makes it easy to apply local methods. Operating directly on the gradient may also have advantages as it relates to the edges (or high frequencies), which are important in image processing.

Other approaches to the reconstruction problem include filtering in the frequency domain, reconstruction from nonuniform samples [21], reconstruction from level crossings [22], and projections onto convex sets [23]. These techniques are more global methods, and the last three are relatively complex and vulnerable to the inaccuracies in the estimation of the phase of the halftone screen frequency. However, they have the potential of yielding reconstructions with even higher resolution.

A. Acquisition of Gradient Samples

In previous work on scanned halftone images [1], [2] directed toward data compression, the halftone dots have been located and divided in four triangular regions. The areas of these subdots are determined. The region connected with a whole dot is called a cell (Fig. 5). For data compression purposes, the (sub)dot areas may be coded to represent the image. Here, as in [3], the subdot areas are used to obtain estimates of the gradient, and the cell dot areas are converted into amplitude samples. The triangle and cell values are first corrected to reduce quantization effects [14].

The derivatives \( f_x \) and \( f_y \) at a cell sample point can be expressed as a function of the sample values in the four triangles forming the cell (white or black), assuming that the derivatives are constant across the dot area. Let \( s_1, s_2, s_3, \) and \( s_4 \) be the values of the four triangles of a cell (Fig. 5). For clustered halftone dots with a diamond dot shape, geometric calculations give the following estimate of the partial derivatives in the cell center [3]:

\[
\begin{align*}
    f_x &= \frac{2\sqrt{2}(s_2 - s_4)}{\sqrt{(s_1 + s_2 + s_3 + s_4)}}, \\
    f_y &= \frac{2\sqrt{2}(s_4 - s_1)}{\sqrt{(s_1 + s_2 + s_3 + s_4)}},
\end{align*}
\]

The idea of acquiring amplitude and gradient samples may be used for other types of halftone dots (including dispersed dots) as long as the gradient of the original image correlates to the bilevel halftone image. The gradient expression should ideally reflect the halftoning method used.

B. Blended Filters for Reconstruction from Halftones

Having acquired both amplitude and gradient samples, the methods for designing blended filters may be applied to these data. As mentioned previously, designing symmetric blended filters (see (1) and (2)) for 2-D data only requires the design of two 1-D filters. The reconstructions rendered at the end of this section have four gray value samples per halftone dot (before being screened again for reproduction). Using four gray value samples per halftone dot is a rule of thumb in the graphic arts. It is also theoretically sufficient for maintaining the information of a signal bandlimited as depicted on Fig. 2. A halftone image has two interlaced grids: one with white dots and one with black dots at the grid points (Fig. 5). At four samples per dot, using the samples of both the black and the white grid, only one 1-D amplitude-derivative filter is required to interpolate the intermediate samples (Fig. 5). This means a discrete 1-D filter that doubles the resolution along the black and/or white grid lines may be used. To keep the filter simple, the intermediate samples are found from the intermediate functions \( f_0 \) and \( f_1 \) of the blended filter along the lines of the black grid, whereas only the amplitude samples are used from the white grid. The design is hereafter limited to the 1-D amplitude-derivative filters doubling the resolution along the grid lines of this setup.
The polynomial (Hermite) filter (see Section III) and the filters obtained by truncating the ideal impulse responses of the filters of Section II (see (3) and (4)) are data independent and straightforward to calculate.

The Wiener filters and the composite filters of Section III require a model and/or statistical measurement of the data. The stationarity of image data may be improved by subtracting the local mean and normalizing by the local variance [16], [24]. A model with a (slowly) varying mean value added to a zero-mean signal is used. The model imposed also assumes symmetry giving symmetric filters. The local mean is subtracted, and here, the local variance may be used to control the weighting functions of a composite filter. The local mean is calculated within the filter window to simplify the filter. The Wiener filters used are based on measured mean values of the products of the observations (6) (with mean value subtracted) within the filter window. The resulting filter output used is then the output of the Wiener filter for the zero-mean data plus the mean value within the filter window. The resulting filter will be referred to as a (constrained) Wiener filter. This implies that the mean value of the original amplitude values is maintained, giving a nonbiased estimate for the nonzero mean image data. This also implies that the sum of the coefficients for the amplitude samples is one.

To design the composite filter, four classes (or 'training sets') of the test image (Fig. 10) are chosen. These are parts of the mug (■), the flower leaves (▲), the flower centers (●), and the calm background region (⋆). The (local) variance of the amplitude samples within the filter window and a Beaudet operator used is a first-order approximation to the (4,2) Wiener filter of (26) may be described by

\[
H'(\omega) = 0.5 + h_1\cos(\pi - \omega) + h_3\cos(3(\pi - \omega)) + l_1\omega\sin(\pi - \omega), 0 \leq \omega < \pi
\]

which expresses the attenuation of the aliasing coming from frequencies originally located at \(\omega\). Fig. 8 shows this aliasing attenuation for the (4,2) Wiener filter of the mug.

V. RESULTS

In this section, the results of applying blended filters to amplitude-gradient samples acquired from halftone images are given. The object is reconstruction of gray value images from bilevel (scanned) halftone images. The reconstruction results are evaluated both objectively (MSE) and subjectively.

Two test images have been used for development and evaluation of the reconstruction filters.

1) The first image is referred to as the scanned test image (Fig. 10). It is screened using conventional contact screening...
Fig. 7. Transfer function up to the screen frequency (π) of filter (solid line) doubling the sampling rate along a grid line. The (4,2) 1-D filter derived from the data of the mug part of the test image is used to interpolate the intermediate points. The transfer function (dotted line) with the derivative samples left out is shown for comparison.

Fig. 8. Attenuation $H'(\omega)$ of aliasing frequencies originally located at $\omega$ from 0 to the screen frequency (π) corresponding to half the new rate.

and thereafter scanned as a bilevel image at 640 dots/cm on a SCITEX Raystar giving a 4864 by 6144 pixel bilevel image. The halftone screen is angled at 63° with a distance of 12.3 bilevel pixels between halftone dots. A line code was added after the screening but before the scanning. The scanned test image is the primary test image for visual evaluation of reconstruction results.

2) The second image, which is referred to as the threshold screened test image, is derived from the same photograph as the first but scanned as a gray value image on a RASTEC Pixact scanner at 100 dots/cm. This digital image has been used to create a (continuous) reference image function by using threshold screening together with a 2-D separable Lagrange interpolation of third order (in 1-D). The screen angle is 63° and the dot spacing the same as above, i.e., 12.3 pixels between the halftone dots. The data of this image has been used to develop the filters and measure errors in the reconstructions. The derivative estimate (in the y direction by (25)) from the scanned test image is shown (Fig 9). Besides showing noise at the screen frequency, this plot illustrates low-frequency patterns due to small variations in the actual halftone grid together with quantization of the grid estimate. It is important to avoid these artifact patterns in the reconstructions. As seen in the images presented below, the reconstruction filters succeed in this.

A. Error Measurements

As mentioned, the threshold screened test image was generated digitally from a continuous description of a gray value image. This enables measurements of errors as the image function may be calculated at any point. This is used to measure the errors of the cell and the triangle samples that are the input to the derivative estimate (25) and the succeeding reconstruction filter. The errors of the reconstructed images are also measured.

Table I illustrates the quality of the cell and triangle samples after the transformation of dot areas into amplitude values (see Section IV). The mean square error (MSE) and peak signal-
Table I gives the errors for the designed Wiener and composite interpolation filters of sixth order (in 1-D). An improvement when going from (6,0) to (4,2) Wiener filters is observed, showing the value of using the derivative samples for a fixed filter order. A further improvement is achieved using the (4,2) composite filter. Increasing the filter order to (4,4) gave only 0.2% improvement in MSE for both the Wiener and the composite filters. This indicates that two samples of the derivative is adequate in the 1-D filters for practical use here. Using a (5,0) Wiener filter for reducing the noise of the amplitude samples gave an improvement from 19.9 to 17.9 in MSE values.

The (composite) Wiener filters designed in this paper are obtained from the digitally screened test image. The filters will therefore reflect the screening method and the masking method used in the processing.

B. Image Results

To test the methods on real data and for subjective evaluation, the scanned test image has been used for reconstruction using the filters designed for the threshold screened test image. The reconstruction result using a (4,2) Wiener filter is presented in Fig. 11.

Comparing the Wiener filter reconstruction (Fig. 11) with the scanned bilevel image (Fig. 10), the reconstruction is not quite as sharp but is still a good reconstruction. The reduced sharpness is mainly seen in the line code, the letters
Fig. 12. Enlargements of part of reconstructed images. From top to bottom: (a) Lagrange polynomial interpolated image \((\hat{f}_0)\); (b) Hermite polynomial interpolated image \((\hat{f}_0)\) using the samples of the derivative in the \(x\) direction; (c) image reconstructed using a polynomial blending filter \((\hat{f})\). The improvement is mainly seen in the background of the mug.

The enlarged part of the image in Fig. 12 illustrates the effect of using gradient data and blended filters. The difference is mainly seen at edges and in detailed areas, e.g., the textured background of the mug.

Fig. 13 illustrates that a reconstructed image may be used as input to a system for further processing, e.g., screening with a different halftone screen in the graphic arts. A very close look at the scanned halftone image (Fig. 10) and the reconstruction (Fig. 11) will also reveal the same difference in the angle of the halftone grid.

VI. Conclusion

Filters for interpolation from samples of an image function and its gradient have been treated. A blended filter technique that composes the 2-D filter of 1-D filters has been introduced. For signals with separable autocorrelation function, blended filters have been shown to be (MMSE) optimal for linear interpolation over a rectangular region of support. The
filter is decomposed in 1-D MMSE linear filters. Other filter techniques in 1-D may also be used to design blended filters.

Blended filters have been applied to the problem of image reconstruction from (scanned) bilevel halftone images. An application useful in the graphic arts industry. Local mean values were subtracted and 1-D MMSE Wiener filters were designed. Local variance and derivative estimates were used to control an adaptive composite Wiener filter based on Wiener, and polynomial blended filters all gave good results ranked in the given order. The derivative data are used in 1-D filters of small sizes with four amplitude samples and two derivative samples. The resulting reconstructions based on local data have high resolution (higher than that of the halftone grid).

APPENDIX A

PROOF OF THEOREM

To prove the lemma and the theorem, some matrix results and notation are used. The following properties [16] of the Kronecker product are used:

\[
(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}
\]  
(A)\hspace{2cm}

\[
(A \otimes B) \cdot (C \otimes D) = A \cdot C \otimes B \cdot D
\]  
(B)\hspace{2cm}

\[
(A \otimes B)^{-T} = A^T \otimes B^T
\]  
(C)

where \(T\) denotes transposition.

Equation (6) may be written in matrix form. Let subscripts \(x'\) and \(y'\) of vectors and matrices refer to the derivatives of the 1-D processes in the respective directions. The subscripts \(a, b,\) and \(c\) refer to the subsets of data involved in the respective intermediate functions \((f_a, f_b,\) or \(f_c)\) of the blended filter (see (1) and (2)). An additional subscripting of \(a, b,\) and \(c\) with \(x\) and \(y\) denotes decomposition in the given direction.

The matrices are organized by letting the indices \(j\) and \(k\) of (6) run through the samples in the order \(f, f_x,\) and \(f_y.\) Within each set of data the order is row by row over the filter window. This organization of data is equivalent to forming a 1-D set for each output of the filter by concatenating the data in the order \(f, f_x,\) and \(f_y\) and within each component using row ordering.

For a separable autocorrelation function \(r_f, (16)-(21)\) are valid for the matrices and vectors of the equations substituting the Kronecker product. The decomposition of the 2-D Wiener filter \(h(f)\) of (22) in the two 1-D Wiener filters \(h_a\) and \(h_y\) uses (A1) and (A2).

Proof of Lemma 1: The 2-D solution to be decomposed is given by

\[
h = (R_h)^{-1} r_y.
\]  
(A4)

The Wiener filter in 1-D involving samples of \(f\) and \(f_x\) is given by

\[
h_a = (R_{a})^{-1} r_y.
\]  
(A5)

\[
R_a = \begin{bmatrix} R_x & R_{x\prime} \\ R_{x\prime} & R_y \end{bmatrix}
\]  
(A6a)

\[
h_y = [h_x f_h y \ f_a f_x]
\]  
(A6b)

where \([\cdot]\) denotes concatenation. The Wiener filter in the \(y\) direction is given by

\[
h_y = (R_y)^{-1} r_y
\]  
({A7})

Combining \(h_a\) and \(h_y\) by direct matrix multiplication, using (16), (17), (19), (A2), (A3), and \(R_a = R_x^y\), we get

\[
h = h_a \otimes h_y = ((R_h)^{-1} r_y) \otimes ((R_y)^{-1} r_y)
\]  
\[
= ((R_x)^{-1} \otimes (R_y)^{-1})(r_a \otimes r_y)
\]  
\[
= \left[ \begin{bmatrix} R_x & R_{x\prime} \\ R_{x\prime} & R_y \end{bmatrix} \right]^{-1} (r_a \otimes r_y)
\]  
\[
= \begin{bmatrix} R_y & R_{y\prime} \\ R_{y\prime} & R_y \end{bmatrix}^{-1} r_y
\]  
(A8)

The last line is the Wiener filter (A4). Given samples of the function \(f\) and the other derivative \(f_y,\) the results are found by symmetry (in \(x\) and \(y\)).

Proof of Theorem 1: The Wiener reconstruction filter \((h)\) may be written in the form

\[
h^T = (h_f + h_f - h_a)^T h_b h_f^T f_f
\]  
(A9)

where \(h_b^T = [h_f h_f^T f_f]\) and \(h_c^T = [h_f h_f^T f_f,\) if it is a blended filter. Furthermore, \(h_b\) should be the solution to (6), which may be written as

\[
\begin{bmatrix} R_f & R_{f\prime} f_f & R_f & R_{f\prime} \\ R_{f\prime} & R_f & R_{f\prime} & R_f \end{bmatrix} \begin{bmatrix} h_f + h_f - h_a \end{bmatrix} = \begin{bmatrix} r_f \\ r_f^T \\ r_f^T f_f \\ r_f^T f_f \end{bmatrix}
\]  
(A10)

The equations involving \(r_f\) on the right-hand side are satisfied by inserting each of the solutions \(h_a\) given by (22) and \(h_y\) and \(h_y\) from Lemma 1. These equations are, therefore, also satisfied by \(h_b + h_a - h_c.\) For the equations involving \(r_f, \) on the right-hand side and subtracting the solution of Lemma 7.1, \(h_b, i.e.

\[
R_{f_f}^T h_f + R_f h_f f_f = r_f
\]  
(A11)

gives

\[
R_{f_f}^T (h_f - h_a) + R_f h_f f_f = 0.
\]  
(A12)

To prove that this holds, the following identities valid over the specified region of support of the filters are used with \(h_f^T = [h_f^T h_f^T f_f]:\)

\[
h_a = h_a \otimes h_a
\]  
(from 22)

\[
h_f = h_a \otimes h_a f_f (\text{using Lemma 1 in } y)
\]

\[
h_f f_f = h_a \otimes h_a f_f (\text{using Lemma 1 in } y)
\]

\[
R_{f_f} f_f = R_{f_f} f_f (\text{by 17})
\]

\[
R_{f_f} f_f = -R_{f_f} f_f (\text{by 21})
\]

Using (A1) and (A2), we get

\[
(R_{f_f} f_f) (h_a \otimes (h_a - h_{c_f}))
\]  
\[
= (-R_{f_f} f_f) (h_a \otimes h_{c_f})
\]  
(A13)
As \( r_y(x) \) is even, (8) gives \( r_y(x) = -r_{2x'}(x) \) and \( R_y^{x'} \) = 
\(-R_{x'}^{y}\). Therefore, using (A2)

\[
R_{x_d}^{y} h_{x} \otimes R_{y}^{x} (h_{x} - h_{f}) = R_{x_d}^{y} h_{x} \otimes R_{y}^{x} h_{c,f},
\]

which is satisfied for

\[
R_y (h_x - h_{f}) = R_{y}^{x} h_{c,f},
\]

which is true because

\[
R_y h_x = r_y = R_{y}^{x} h_{c,f} + R_{y}^{x} h_{c,f},
\]

where the right equality is by the symmetric version of (A5)

and (A6) (replacing \( x \) by \( y \)).

By symmetry (replacing \( x \) with \( y \)), the equations involving \( r_{f,f} \) on the right-hand side of (A10) are also satisfied.

REFERENCES