

## Deblurring Gaussian Blur\*

ROBERT A. HUMMEL

*Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street,  
New York, New York 10012*

AND

B. KIMIA AND STEVEN W. ZUCKER

*Computer Vision and Robotics Lab, Department of Electrical Engineering, McGill University,  
Montreal, Quebec, Canada*

Received December 6, 1984; revised June 23, 1986

Gaussian blur, or convolution against a Gaussian kernel, is a common model for image and signal degradation. In general, the process of reversing Gaussian blur is unstable, and cannot be represented as a convolution filter in the spatial domain. If we restrict the space of allowable functions to polynomials of fixed finite degree, then a convolution inverse does exist. We give constructive formulas for the deblurring kernels in terms of Hermite polynomials, and observe that their use yields optimal approximate deblurring solutions among the space of bounded degree polynomials. The more common methods of achieving stable approximate deblurring include restrictions to band-limited functions or functions of bounded norm.

© 1987 Academic Press, Inc.

## 1. INTRODUCTION

Given an image or signal, the realization of any system for processing it must introduce some amount of degradation. Since there may be several stages each contributing to the degradation, the composition is often modeled as a Gaussian blurring operation. We consider spatially invariant Gaussian convolution defined as follows. For a bounded measurable input function  $f(x)$  defined for  $x \in \mathbb{R}^n$ , then the observed blurred output is given by

$$h(x) = \int_{\mathbb{R}^n} K(x - \xi, t) f(\xi) d\xi,$$

where

$$K(x, t) = \frac{1}{(4\pi t)^{n/2}} e^{-|x|^2/4t},$$

and  $t$  is a fixed positive value parameterizing the extent of the blur. We wish to estimate  $f(x)$  when only  $h(x)$  and the amount of blur  $t$  are known.

It would be especially nice to formulate deblurring as a convolution operation, so that

$$f = D(\cdot, t) * h.$$

\*This research was supported by NSERC Grant A4470 and MRC Grant MA6154 in Canada, and NSF Grant DCR8403300 and ONR Grant N0014-85-K-0077 in the United States.

In general, a universal deblurring kernel  $D(x, t)$  does not exist. However, if sufficient restrictions are placed on the domain of permissible functions  $f$ , then deblurring kernels can exist.

Our interest in deblurring is motivated by two concerns. First, deblurring is of significant practical importance in many image processing systems, for example in computerized tomography [1]. There are also applications in physiological optics, such as the de-focusing that automatically takes place for objects outside the depth of field of an accommodated eye.

A second motivation is provided by a desire to study stability of image representations. The mathematical analysis of an image representation must include a study of the continuity and stability of the transformation. Reconstruction methods are particularly useful for studying the stability. While blurred versions of an original signal form a classical unstable representation, many intermediate-level transformations of image data nonetheless involve some degree of filtering by blurring. For example, representations involving zero-crossings of Gaussian-filtered Laplacians of images [2], as well as many other pyramid schemes, [3–6], involve Gaussian blur. Instabilities in the representation may not be important if approximate or pseudo-inverse reconstruction methods (see, e.g., [7]) can be found that make explicit the assumptions concerning the input data. In this paper, we present an approximate-inverse method involving polynomial approximations.

## 2. THE HEAT EQUATION

### 2.1. Diffusion

There is a fundamental connection between Gaussian blurring and the heat equation. Consider a rod of infinite length onto which an impulse of heat is placed. As time evolves, the heat will diffuse and the original impulse will spread out. By elementary physics, the resulting distribution will approximate a Gaussian whose width depends on the elapsed time (see, e.g., the Feynman lectures [8]). By superposition, the model for the temperature distribution along the rod at any given time is the initial temperature distribution convolved with a Gaussian. The diffusion process effectively *convolves* the initial distribution by a Gaussian whose spread depends on how much time has evolved. This is the physically realized solution to the heat equation, which can be formulated as follows [9]. Given  $f(x)$  piecewise continuous and bounded, find  $h(x, t)$  bounded and  $C^2$  for  $t > 0$  satisfying

$$\frac{\partial h}{\partial t}(x, t) = \Delta h(x, t), \quad x \in \mathbb{R}^n, t > 0;$$

$$h(x, t) \rightarrow f(x_0) \quad \text{as} \quad (x, t) \rightarrow (x_0, 0), \quad x_0 \in \mathbb{R}^n, t > 0.$$

We denote the operator that takes  $f$  to  $h(\cdot, t)$  by  $\Omega_t$ , i.e.,  $h(x, t) = (\Omega_t f)(x)$ . The solution is given by

$$(\Omega_t f)(x) = \int_{\mathbb{R}^n} K(y - x, t) f(y) dy,$$

where  $K$  is as defined before. When restricted to a Hilbert space such as  $L^2(\mathbb{R}^n)$ ,  $\Omega_t$  becomes a symmetric bounded linear operator. We will generally interpret  $f(x)$  as an image.

The space of functions that can be blurred in this way is very large. Indeed, the condition that  $f(x)$  is bounded can be weakened. The solution is still given by convolution against the “source kernel”  $K(x, t)$  [10],  $h(\cdot, t) = K(\cdot, t) * f$ . The source kernel is the *fundamental solution* to the heat equation on the unbounded domain  $\mathbb{R}^n$ . We also note that the blurring operator satisfies a semigroup property — that if a function  $f(x)$  has been blurred for some time  $t_1$  by  $\Omega_{t_1}$ , and the resultant function is blurred by  $\Omega_{t_2}$ , the end result is the same as blurring  $f(x)$  for a time  $t_1 + t_2$ . That is,  $\Omega_{t_1} \circ \Omega_{t_2} = \Omega_{t_1+t_2}$ . The two Gaussian blurring operators, each of which may have its own physical justification, results in one composite Gaussian operator. Indeed, by the central limit theorem, other blurring operators will also compose into approximate Gaussians when iterated.

## 2.2. Deblurring

Deblurring is the inverse problem to blurring, and can be modeled as a diffusion process running backwards in time. Formally, the problem of reconstructing  $f(x)$  given a blurred function  $h(x)$  and a blurring amount  $t > 0$  is the inverse heat equation problem, and poses technical difficulties not present in the forward heat equation problem.

First, finding an inverse to  $\Omega_t$  presupposes that  $\Omega_t$  is one-to-one. In fact, the blurring operator *is* one-to-one providing minor restrictions are placed on the domain of  $\Omega_t$ . However, without certain growth restrictions, it is possible to find distinct functions  $f$  and  $\tilde{f}$  satisfying  $\Omega_t f = \Omega_t \tilde{f}$ , (see [11]). Second, a solution  $f$  to the problem  $\Omega_t f = h$ , given  $h$ , exists only if  $h$  is sufficiently smooth. In general, an inverse cannot be found, and even if  $h$  is sufficiently smooth, an arbitrarily small change can destroy the smoothness. John [12] discusses the technical conditions needed for the existence of an inverse. Finally, in a general function space the deblurring problem is horribly ill-conditioned. This means that there can exist pairs of functions  $f$  and  $\tilde{f}$  that are arbitrarily far apart whose images under  $\Omega_t$  are arbitrarily close. The prototypic example is  $f(x) = A \sin(\omega x)$  and  $\tilde{f}(x) = 0$  in one space dimension. Then  $(\Omega_t f)(x) = A e^{-\omega^2 t} \sin(\omega x)$ , which for  $\omega$  large can be very close to  $\Omega_t \tilde{f} = 0$ .

Deblurring can be understood somewhat better in terms of the Fourier transform. If we denote the Fourier transform of a function  $g(x)$  by  $\hat{g}(\omega)$ , then the blurring operator  $\Omega_t$  is a multiplier operator given by

$$\widehat{(\Omega_t f)}(\omega) = e^{-\omega^2 t} \hat{f}(\omega).$$

By means of this formula,  $\Omega_t$  can be extended to operate on the class of temperate distributions  $\mathcal{S}$  of Fourier transformable distribution [13]. In particular,  $\Omega_t f$  is defined for any polynomial  $f$ . Further, the formula shows that  $\Omega_t$  is one-to-one on any class of Fourier transformable functions. Moreover, our earlier observation that deblurring can not generally be represented by a convolution kernel can be observed from the formula, since although

$$\hat{f}(\omega) = e^{\omega^2 t} \cdot \hat{h}(\omega),$$

a general convolution formula is not possible since  $e^{\omega^2 t}$  is not the Fourier transform of any tempered distribution.

These difficulties would tend to make one pessimistic about accomplishing image deblurring, and in particular about discovering deblurring kernels. However, deblurring is a common operation, and is typically accomplished by giving the problem a variational formulation, which can lead to a well-conditioned problem. We describe several variational formulations in the next section, and present deblurring kernels for polynomial domains in the subsequent section.

### 3. VARIATIONAL FORMULATION

We choose a normed space  $\mathcal{N}$  and a closed convex subset  $\mathcal{M} \subseteq \mathcal{N}$  so that  $\Omega_t$  may be regarded as an operator  $\Omega_t: \mathcal{M} \rightarrow \mathcal{N}$ , for all  $t \geq 0$  (note that  $\Omega_0$  is simply the identity operator). We may then pose the deblurring problem in the following form:

$$\text{Given } h \in \mathcal{N}, t > 0, \quad \text{find } f \in \mathcal{M} \text{ minimizing } \|\Omega_t f - h\|.$$

If  $\Omega_t$  is one-to-one and onto over  $\mathcal{M}$ , then the solution  $f$  given  $h$  is precisely the inverse image of  $h$  under  $\Omega_t$ , so that the minimization gives a zero norm. In general, however,  $\mathcal{M}$  is restricted in such a way that  $\Omega_t$  maps into  $\mathcal{N}$ , and so the solution  $f$  is a pseudo-inverse of  $h$ .

Difficulties arise because the operator  $\Omega_t$  on the domain  $\mathcal{M}$  is in general ill-conditioned. There are various approaches that one can take to find a good deblurred signal  $f$  stably from a given  $h$ . A standard approach is to restrict  $\mathcal{M}$  to a sufficiently small set. We mention five possibilities, and our subsequent treatment will use one of these approaches (the second one) as a point of departure.

(1)  $\mathcal{N} = L^2$ , and  $\mathcal{M} = \{f \in \mathcal{M} \mid \hat{f}(\omega) = 0 \text{ for } |\omega| > \rho\}$ , for some fixed constant  $\rho$ . Restricting to the space of band-limited signals (with a specified cutoff) allows stable deblurring of a blurred signal  $h$  by means of the formula

$$\hat{f}(\omega) = \begin{cases} e^{\omega^2/4} \hat{h}(\omega) & |\omega| \leq \rho \\ 0 & |\omega| > \rho \end{cases}$$

The function  $f$  can also be written as a convolution against  $h$ :  $f = K_\rho * h$ , where the Fourier transform of  $K_\rho$  is  $e^{\omega^2/4}$  for  $|\omega| \leq \rho$ , and zero elsewhere. This is a standard method for deblurring, although it is well known that  $K_\rho$  "rings" over a large spatial extent.

In a discrete setting, a similar (discrete) convolution deblurring kernel can be formulated, yielding appropriate band-limited discrete approximations to original signals. It turns out that this method is completely equivalent to computing a pseudo-inverse of the matrix representing the blurring operator by means of a singular value decomposition.

(2)  $\mathcal{N} = L^2(e^{-x^2} dx)$ , and  $\mathcal{M} =$  polynomials of degree  $N$  or less (fixed  $N$ ), which we will denote by  $P_N$ . We will see (in the next section) that  $\Omega_t$  is closed on  $P_N$ . Thus the pseudo-inverse problem can be solved as follows. First,  $h$  is projected onto  $P_N$  by the linear orthogonal projection operator of  $\mathcal{N}$  into  $P_N$ . The resulting polynomial is deblurred by the convolution kernel  $D_N$  to be defined in Section 4 to yield the solution polynomial  $f$ .

The problem with this possibility is that images are typically far more general than polynomials of degree  $N$ . Thus to get even moderately reasonable deblurring,  $N$  has to be very large, and then applying  $D_N$  becomes difficult and numerically unstable. Although we will make use of the deblurring kernels  $D_N$  designed to deblur polynomials, our implementations (Sect. 5) are more general than finding the pseudo-inverse  $f$  among the class  $P_N$ .

(3)  $\mathcal{N} = L^\infty(\mathbb{R})$ , and  $\mathcal{M} = \{f \in \mathcal{N} | f \geq 0\}$ . This situation, studied by John [12] leads to partially stable deblurring, as long as the given function  $h$  is sufficiently "blurry." Specifically, suppose  $T > t$ , and  $h = \Omega_T g$ , some  $g \in \mathcal{M}$ . Then the problem is to find  $f \in \mathcal{M}$  such that  $\Omega_t f = h$ . John studies bounds on the deblurring error, where deblurring is accomplished by exactly the same linear process to be discussed in the next section. That is, he constructs an approximation  $\tilde{f}$  to  $f$  by convolving  $h$  with a scaled version of the kernel  $D_N$  given in Section 4. The result is that the error in reconstructing  $f$ ,  $\|\tilde{f} - f\|$ , can be controlled to depend continuously on the error in representing  $h$ . That is, small errors in representing  $h$  can lead to errors in representing  $f$ , but the maximum size of the errors can be bounded. Interestingly, unlike customary notions of stability, the dependence is not linear.<sup>1</sup>

(4)  $\mathcal{N} = L^2(\mathbb{R})$ , and  $\mathcal{M} = \{f | 0 \leq f \leq M\}$ , for some fixed  $M$ . With  $M = \infty$ , this is nearly the same as case (2). Now, however, we consider the possibility of nonlinear deblurring methods. Peleg *et al.* [14] has implemented a deblurring scheme based on a conjugate gradient iterative minimization of  $\|\Omega_t f - h\|$ , constrained by  $f \in \mathcal{M}$ . The constraints are handled, in Peleg's case, by remapping the interval  $[0, M]$  to  $[-\infty, \infty]$ , and then solving an unconstrained minimization problem. By limiting the number of iterations, they obtain only an approximate solution, although the results look very good. They do not study the stability question, but one would expect the same kind of nonlinear stability for partial deblurring as discovered by John.

(5)  $\mathcal{N} = L^2(\mathbb{R})$ ,  $\mathcal{M} = \{f \in L^2 | \|f\| \leq M\}$ , some fixed  $M$ . This case is treated by Carasso *et al.* [15]. They give a relatively simple nonlinear deblurring method, making use of Fourier transforms, to solve the variational problem. The method is not iterative. They also study the stability, and obtain the same kind of stability estimates (for partial deblurring) as John.

An alternative approach to obtaining stable deblurring is to begin by specifying the algorithmic form of the deblurring method, and to optimize with respect to a statistical norm. For example, we can insist on a convolution kernel for deblurring, and seek a kernel  $k$  minimizing

$$E\{\|k * \Omega_t f - f\|\},$$

where  $E\{\cdot\}$  is an expectation operator which presupposes some distribution of unblurred functions  $f$ . Other operators, such as worst-case norm, are also possible. Such methods are studied in the province of information-based complexity [16].

<sup>1</sup>Suppose that  $\tilde{h}$  is the representation of  $h$ , and that  $\tilde{f}$  is the approximate reconstruction of  $f$  using  $\tilde{h}$ . Usual notions of stability would require  $\|\tilde{f} - f\| \leq \mathcal{C} \cdot \|\tilde{h} - h\|$ , using appropriate (and perhaps different) norms. The nonlinear stability that is used in this case, however, asserts that  $\|\tilde{f} - f\| \leq \mathcal{C} \cdot \|\tilde{h} - h\|^{1/k}$ , for some integer  $k \geq 1$ . Thus in order to achieve an accuracy of  $\epsilon$  in the reconstruction of  $f$ , it is necessary to represent  $h$  to an accuracy of  $(1/\mathcal{C})\epsilon^k$ . This might be viewed as *polynomial stability*.

If the distribution of  $f$ 's is concentrated on, or limited to, polynomials in  $P_N$ , then all the functions  $\Omega_t f$  will also be polynomials, and we expect that the optimal deblurring kernel  $k$  will be the one that deblurs polynomials in  $P_N$  (namely,  $D_N$ , given in the theorem of Sect. 4). If the distribution consists of functions that are well approximated locally by polynomials in  $P_N$ , the optimal kernel will not change much.

In the next section, we present the kernel  $D_N$  such that  $D_N * \Omega_t f = f$  for all  $f \in P_N$  (with  $t = \frac{1}{4}$ ). Since  $\Omega_t$  is closed on  $P_N$ , this kernel may be used to deblur all polynomials  $h \in P_N$ . However, motivated by John's studies (case (2) above) and the comments on optimal kernels, we will apply  $D_N$  to functions  $h$  that are not in fact polynomials.

The results may be analyzed in terms of local approximations by polynomials. Suppose that the initial unblurred data  $f$  is written

$$f = p + \varepsilon,$$

where  $p \in P_N$  and  $\varepsilon$  is an error term. Specifically,  $p$  should be the projection of  $f$  onto  $P_N$  in the space  $L^2(e^{-x^2} dx)$ , so that  $p$  is a good polynomial approximation to  $f$  near the origin, but may differ from  $f$  significantly away from the origin. Applying the deblurring kernel  $D_N$  to the blurred version of  $f$  yields the approximate deblurring

$$\tilde{f} = p + D_N * \Omega_t \varepsilon.$$

Thus

$$f - \tilde{f} = \varepsilon - D_N * \Omega_t \varepsilon,$$

whose norm (in  $L^2(e^{-x^2} dx)$ ) can be bounded by  $\mathcal{C} \cdot \|\varepsilon\|$ . Since the norm measures errors only locally, the result is that we have stable, accurate deblurring for signals that are well approximated locally by polynomial data.

#### 4. POLYNOMIAL DOMAINS

The monomials  $\{1, x, \dots, x^N\}$  form a basis for  $P_N$ . If this basis is orthonormalized with respect to the inner product

$$(f, g) = \int_{-\infty}^{\infty} f(x)g(x)e^{-x^2} dx,$$

then the basis of Hermite polynomials  $\{H_0, H_1, \dots, H_N\}$  result. The Hermites can be represented explicitly,

$$H_n(x) = n! \sum_{m=0}^{\lfloor n/2 \rfloor} (-1)^m \frac{(2x)^{n-2m}}{m!(n-2m)!},$$

or by the Rodrigues formula,

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} (e^{-x^2}).$$

Without loss of generality, we will specialize to the case  $t = \frac{1}{4}$ , and denote  $\Omega_{1/4}$  by  $\mathcal{F}$ . We now prove the aforementioned

**OBSERVATION 1.**  $\mathcal{F}$  is closed on  $P_N$ .

*Proof.* We will show that  $\mathcal{F}H_n \in P_n$  for  $n \leq N$ .

$$\begin{aligned} \sqrt{\pi}(\mathcal{F}H_n)(y) &= \int_{-\infty}^{\infty} e^{-(y-x)^2} H_n(x) dx \\ &= \int_{-\infty}^{\infty} e^{-y^2} e^{2xy} (-1)^n \frac{d^n}{dx^n} (e^{-x^2}) dx \\ &= 2y \int_{-\infty}^{\infty} e^{-y^2} e^{2xy} (-1)^{n-1} \frac{d^{n-1}}{dx^{n-1}} (e^{-x^2}) dx \\ &= \sqrt{\pi} 2y (\mathcal{F}H_{n-1})(y). \end{aligned}$$

Solving this recursion relation, using  $\mathcal{F}H_0 = 1$ , we have

$$(\mathcal{F}H_n)(x) = 2^n x^n. \quad \square$$

As a result of Observation 1,  $\mathcal{F}$  is an isomorphism of  $P_N$ . The inverse of  $\mathcal{F}$  on  $P_N$  is clearly given by

$$\mathcal{F}^{-1} \sum_{i=0}^N (a_i x^i) = \sum_{i=0}^N \frac{a_i}{2^i} H_i(x).$$

Our main result is that  $\mathcal{F}^{-1}$  restricted to  $P_N$  can be represented by a convolution with an explicit kernel  $D_N(x)$ :

**THEOREM.** For  $f \in P_N$  and  $g = \mathcal{F}f$ , then

$$f = D_N * g$$

where

$$D_N(x) = e^{-x^2} \sum_{k=0}^{\lfloor N/2 \rfloor} \frac{(-1)^k}{\sqrt{\pi} k! 2^k} H_{2k}(x).$$

We will give a proof below using direct integration (as opposed to using Fourier transform distributions). Note, however, that  $D_N(x)$  is not the unique function representing  $\mathcal{F}^{-1}$  on  $P_N$ . In general, the kernel can be translated by any function which yields a zero convolution against  $P_N$ . This includes all functions of the form

$$e^{-x^2} H_n(x), \quad n > N.$$

The stated kernel is unique among the class of functions of the form  $e^{-x^2} P(x)$ , where  $P(x)$  is a polynomial of degree  $N$ .

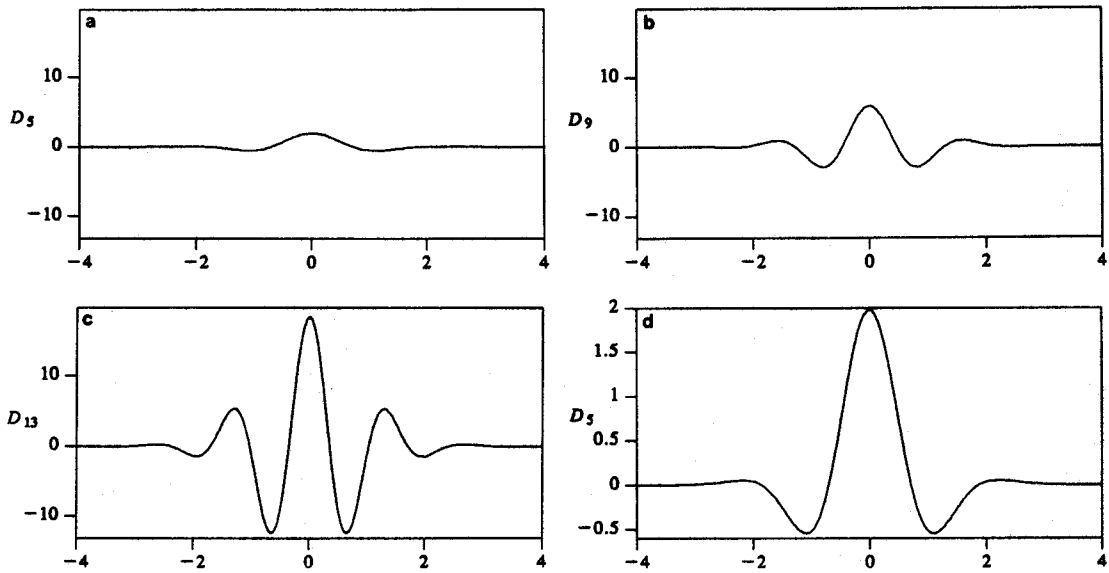


FIG. 1. The deblurring kernels  $D_N$ , for  $N = 5, 9$ , and  $13$  for (a), (b), and (c), respectively. Figures (a), (b), and (c) are drawn on the same vertical scales. Figure (d) shows the deblurring kernel  $D_5$  (the same as Fig. 1(a)) with a different vertical scale to emphasize the structure. The corresponding blurring kernel is  $(1/\sqrt{\pi})e^{-x^2}$ , so that one standard deviation  $\sigma$  is equal to  $x = \pm\sqrt{2}/2$ .

It is interesting to compare the form of  $D_N(x)$  with standard enhancement filters. For example, for  $N = 3$ ,

$$D_3(x) = \frac{2}{\sqrt{\pi}}e^{-x^2}(1 - x^2) = \frac{1}{\sqrt{\pi}}e^{-x^2} - \frac{1}{2} \frac{d^2}{dx^2} \left( \frac{1}{\sqrt{\pi}}e^{-x^2} \right).$$

Thus

$$D_3 * g = \left[ 1 - \frac{1}{2} \frac{d^2}{dx^2} \right] \mathcal{T}g,$$

which is a not uncommon high emphasis filter (see, e.g., the papers by Mach in [17; 18]. In Fig. 1, we display plots of  $D_N$  for several values of  $N$ .

The proof of the theorem depends on several lemmas.

LEMMA 1.

$$A_n = \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}} e^{-x^2} x^n dx = \begin{cases} 0, & n \text{ odd,} \\ \frac{n!}{2^n(n/2)!}, & n \text{ even.} \end{cases}$$

LEMMA 2.

$$c_{2k,2p} = \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}} e^{-x^2} H_{2k}(x) x^{2p} dx = \begin{cases} 0, & p < k, \\ \frac{(2p)!}{2^{2p-2k}(p-k)!}, & p \geq k. \end{cases}$$



*Proofs.* Lemmas 1 and 2 may be found in [19, Sect. 4.16, Example 1].

LEMMA 3. For  $n \geq k$ ,

$$d_k = \int_{-\infty}^{\infty} D_N(x) x^k dx = \begin{cases} 0, & k \text{ odd,} \\ (-1)^{k/2} \frac{k!}{2^k (k/2)!}, & k \text{ even.} \end{cases}$$

*Proof.* For  $k$  odd, and using the definition of  $D_N(x)$ , we observe that  $D_N(x)x^k$  is an odd integrable function, and so integrates to zero. For  $k = 2p$ ,

$$\begin{aligned} \int_{-\infty}^{\infty} D_N(x) x^k dx &= \int_{-\infty}^{\infty} e^{-x^2} \sum_{i=0}^{\lfloor N/2 \rfloor} \frac{(-1)^i}{\sqrt{\pi} i! 2^i} H_{2i}(x) x^{2p} dx \\ &= \sum_{i=0}^{\lfloor N/2 \rfloor} \frac{(-1)^i}{i! 2^i} c_{2i, 2p} = \sum_{i=0}^p \frac{(-1)^i}{i! 2^i} \frac{(2p)!}{2^{2p-2i} (p-i)!} \\ &= \frac{(2p)!}{2^p p!} \sum_{i=0}^p \frac{p!}{i! (p-i)!} (-1)^i (1/2)^{p-i} \\ &= \frac{(2p)!}{2^p p!} (-1/2)^p = (-1)^p \frac{(2p)!}{2^{2p} p!}. \quad \square \end{aligned}$$

*Proof of the theorem.* From the formula for  $\mathcal{F}H_n$  computed in Observation 1, it suffices to show that  $D_N * (2^n x^n) = H_n(x)$ ,  $n \geq N$ . We have

$$\begin{aligned} (D_N * 2^n x^n)(y) &= \int_{-\infty}^{\infty} 2^n D_N(x) (y-x)^n dx \\ &= \int_{-\infty}^{\infty} 2^n D_N(x) \sum_{i=0}^n (-1)^k \binom{n}{k} y^{n-k} x^k dx \\ &= \sum_{k=0}^n \frac{2^n n!}{k! (n-k)!} (-1)^k y^{n-k} d_k \\ &= n! \sum_{m=0}^n \frac{(-1)^{2m} 2^n}{(2m)! (n-2m)!} (-1)^m \frac{(2m)!}{2^{2m} m!} y^{n-2m} \\ &= n! \sum_{m=0}^n \frac{(-1)^m}{m! (n-2m)!} 2^{n-2m} y^{n-2m} = H_n(y). \quad \square \end{aligned}$$

The theorem could have been proved using the convolution theorem and by computing the Fourier transform of  $D_N(x)$ . Although we have not taken this approach, we will nonetheless compute  $\hat{D}_N$  in order to show that the multiplier for  $D_N$  approaches, pointwise, the inverse of the multiplier for the operator  $\mathcal{F}$ , i.e.,

OBSERVATION 2.  $\hat{D}_N(\omega) \rightarrow e^{\omega^2/4}$  pointwise as  $N \rightarrow \infty$ .

*Proof.*

$$\hat{D}_N(\omega) = \sum_{k=0}^{\lfloor N/2 \rfloor} \frac{(-1)^k}{\sqrt{\pi} i! 2^k} \mathcal{F} [e^{-x^2} H_{2k}(x)](\omega),$$

where  $\mathcal{F}$  stands for the Fourier transform operator. Now,

$$\mathcal{F} \left[ e^{-x^2} H^{2k}(x) \right] (\omega) = \mathcal{F} \left[ (-1)^{2k} \frac{d^{2k}}{dx^{2k}} (e^{-x^2}) \right] (\omega) = (i\omega)^{2k} \sqrt{\pi} \cdot e^{-\omega^2/4}.$$

Thus

$$\begin{aligned} \hat{D}_N(\omega) &= \sum_{k=0}^{\lfloor N/2 \rfloor} \frac{(-1)^k}{\sqrt{\pi} k! 2^k} \cdot (-1)^k \omega^{2k} \sqrt{\pi} e^{-\omega^2/4} \\ &= e^{-\omega^2/4} \sum_{k=0}^{\lfloor N/2 \rfloor} \frac{1}{k!} \left( \frac{\omega^2}{2} \right)^k. \end{aligned}$$

Hence

$$\lim_{N \rightarrow \infty} \hat{D}_N(\omega) = e^{-\omega^2/4} e^{\omega^2/2} = e^{\omega^2/4}. \quad \square$$

It is interesting to observe that the kernel  $D_N$  is a multiplier where the Fourier transform is the Taylor series approximation to the function  $e^{\omega^2/2}$ , multiplied by the window  $e^{-\omega^2/4}$ .

Also as a consequence of Observation 2, we see that  $D_N(x)$  does *not* converge pointwise to any function as  $N \rightarrow \infty$ , since otherwise the Fourier transform of that function would be  $e^{\omega^2/4}$ , which is impossible.  $D_N(x)$  does converge in  $L^2(e^{-x^2} dx)$ , but that does not imply pointwise convergence to any function. We accordingly have stable deblurring when using the kernels  $D_N(x)$ , where stability is measured in terms of deviation from a polynomial of degree  $N$ , and the  $L^2(e^{-x^2} dx)$  norm is used as the metric.

## 5. HIGHER DIMENSIONS

The Gaussian blur operator is given by

$$\mathcal{T}f(x) = \int_{\mathbb{R}^n} \pi^{-n/2} e^{-(x-y)^2} f(y) dy.$$

Due to the separability of the kernel and Fubini's theorem,  $\mathcal{T}$  can be decomposed into  $n$  iterated blurrings:

$$\begin{aligned} \mathcal{T} &= \mathcal{T}_1 \circ \mathcal{T}_2 \circ \dots \circ \mathcal{T}_n \\ (\mathcal{T}_i f)(x) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}} e^{-(x_i - y_i)^2} f(x_1, \dots, y_i, \dots, x_n) dy_i. \end{aligned}$$

Consider a polynomial in  $\mathbb{R}^n$ ,

$$\begin{aligned} f(x) &= \sum_{|\alpha| \leq N} a_\alpha x^\alpha \\ \alpha &= (\alpha_1, \alpha_2, \dots, \alpha_n), \quad \alpha_i \in \mathbb{Z}, \alpha_i \geq 0, \quad |\alpha| = \sum \alpha_i, \quad x^\alpha = x_1^{\alpha_1} \dots x_n^{\alpha_n}. \end{aligned}$$



FIG. 2. An original image (a) digitized to 480 by 512 pixels. In (b), the image has been blurred by a Gaussian  $(1/\pi)e^{-x^2}$ , with the interpixel distance  $h = 0.15$ . Figures (c), (d), and (e) show Figure (b) restored, using the deblurring kernels  $D_5$ ,  $D_9$ , and  $D_{13}$ , respectively.

For fixed  $n$ , the function of one real variable

$$g(y_i) = f(x_1, \dots, y_i, \dots, x_n)$$

is a polynomial of degree no greater than  $N$ , so

$$D_N^*(\mathcal{T}g) = g$$

where  $\mathcal{T}$  is the standard 1-dimensional blurring operator (2.1). Combining we find that

$$f(x) = \int_{\mathbb{R}^n} D_N(y_1) D_N(y_2) \cdots D_N(y_n) (\mathcal{T}f)(x - y) dy$$

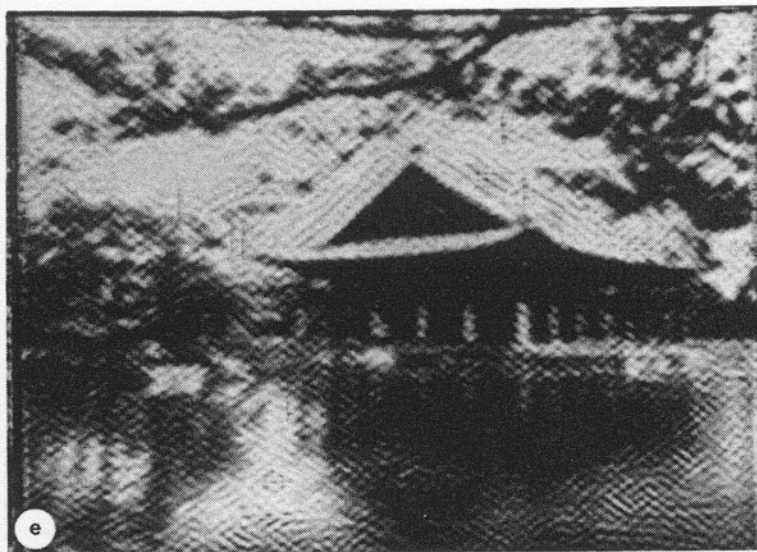
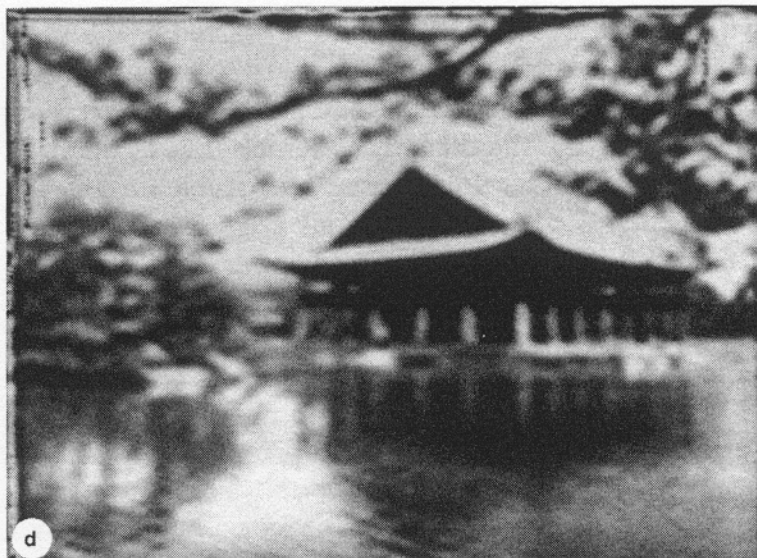


FIG. 2.—Continued.

for any multivariate polynomial  $f(x)$ . Thus deblurring of blurred polynomials of degree  $N$  can be accomplished by convolution against the kernel

$$D_N^*(x) = D_N(x_1)D_N(x_2) \cdots D_N(x_n).$$

Thus the situation in higher dimensions is similar to the 1-dimensional case. The deblurring convolution kernel is separable and will be of the form  $e^{-x^2}P(x)$ , where  $P(x)$  is a polynomial of degree  $nN$  in  $x \in \mathbb{R}^n$ . Figure 2 shows a plot of  $D_N^*$  for  $n = 2$ ,  $N = 3$ . Using a separable kernel for deblurring has computational advantages, but can lead to artifacts in diagonal directions when applied to certain images. We will see these effects in the next section.

## 6. EXPERIMENTS IN DEBLURRING

We experimented with the deblurring kernels  $D_N$  using both modest and large amounts of blur. An original image (Fig. 2a) has eight bits of grayscale information digitized on a 480 by 512 grid. The image is regarded as a piecewise constant function of two continuous real variables and is blurred by convolution against the Gaussian  $(1/\pi)e^{-x^2-y^2}$ , where for Fig. 2b the interpixel spacing is taken to be  $h = 0.15$  (corresponding to a standard deviation  $\sigma = 4.7$  pixels). Figures 2c, d, and e show the results of applying the deblurring kernels  $D_5$ ,  $D_9$ , and  $D_{13}$ , respectively. The diagonal artifacts arise due to the use of separable kernels and become more pronounced for higher  $N$ . The computations were done with 12-bit fixed-point arithmetic on a VICOM image processing computer. The results by using floating point arithmetic on a general purpose computer (a VAX) were essentially identical, although the effect of the diagonal artifacts is very slightly reduced.

For comparison, we display in Fig. 3 the results of convolving the blurred image in Fig. 2b by a kernel  $K_\rho$  whose Fourier transform is a truncated version of  $e^{\omega^2}$ . The kernel size is arbitrarily limited to 100 by 100 pixels, and  $\rho$  is chosen small enough



FIG. 3. The blurred image in Figure (b) deblurred using a kernel whose Fourier transform is a truncated version of the deblurring multiplier. The result is a pseudo-inverse under the blurring operator. Difficulties arise because the deblurring kernel "rings."

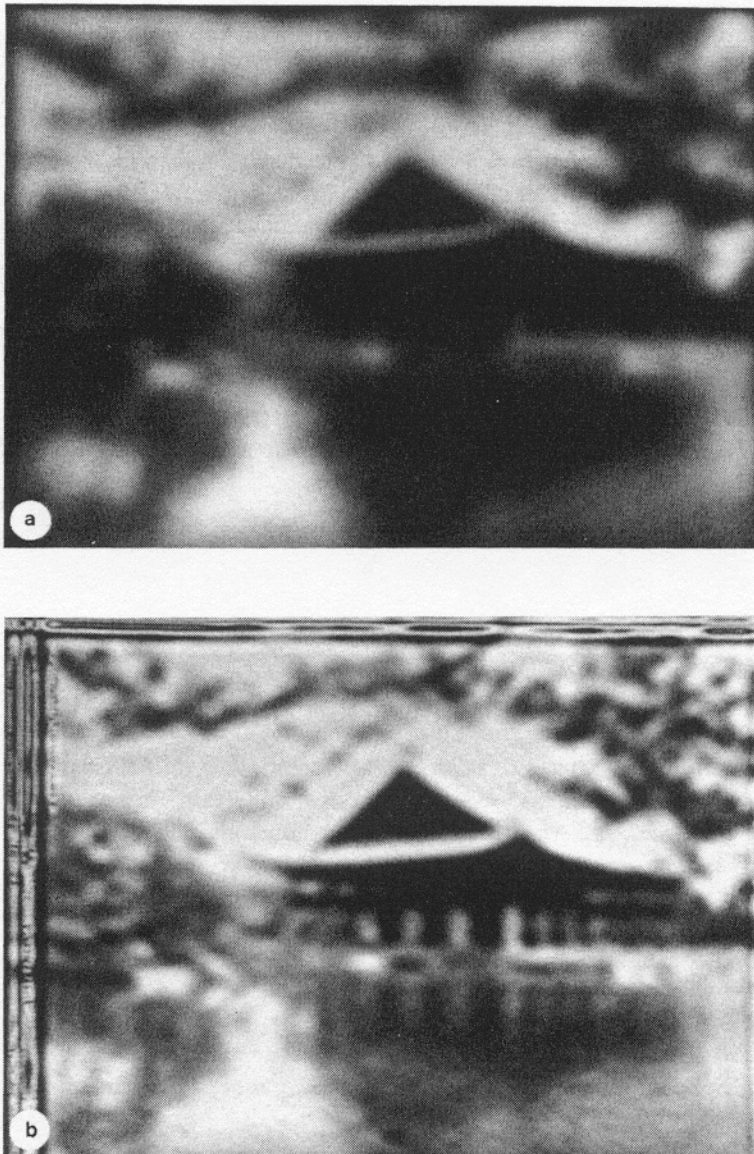


FIG. 4. The original image blurred by a Gaussian with interpixel distance  $h = 0.10$ , and the result of deblurring using the kernel  $D_{11}$ .

so that the kernel elements are considerably smaller on the periphery of the kernel. If  $\rho$  were chosen too large, the magnitude of the kernel elements would decay extremely slowly. Clearly, deblurring by the kernel  $D_N$  is far superior, due to the elimination of the sharp cutoff in its spectral properties. The difficulty with a sharp cutoff is that it leads to "ringing" of the spatial kernel, and evidence of ringing can be seen in the example of deblurring.

In Fig. 4, we show the blurred image using an interpixel distance of  $h = 0.1$  (corresponding to a standard deviation  $\sigma = 7.1$  pixels). The blurring kernel, to beyond three standard deviations, is roughly a 50 by 50 mask. Figure 4b shows this image deblurred using the kernel  $D_{11}$ .

The experimental results show that deblurring certainly improves blurred images, even with the given inherent lack of stability. Better results may be obtainable with

nonlinear or stochastic techniques, although our experiments certainly demonstrate the human visual system's sensitivity to visual quality of deblurrings.

## REFERENCES

1. G. Herman, *Image Reconstruction from Projections*, Academic Press, New York, 1980.
2. D. Marr and E. Hildreth, Theory of edge detection, *Proc. R. Soc. London B*, 1980, 187.
3. P. Burt and T. Adelson, The laplacian pyramid as a compact image code, *IEEE Trans. Commun.* 1983, 532.
4. J. Crowley, *A Representation for Visual Information*, Ph.D. thesis, CMU Robotics Institute, 1982.
5. J. J. Koenderink, The structure of images, *Bio. Cybernet.* **50**, 1984, 363–370.
6. S. W. Zucker and R. A. Hummel, Receptive fields and the reconstruction of visual information, *Human Neurobiol.* **5**, 1986, 121–128.
7. W. Pratt, *Digital Image Processing*, Wiley-Interscience, New York, 1978.
8. R. Feynman, R. Leighton, and M. Sands, *The Feynman Lectures on Physics*, Addison-Wesley, Reading, Mass., 1963.
9. L. Bers, F. John, and M. Schechter, *Partial Differential Equations*, Amer. Math. Soc., Providence, R.I., 1964.
10. D. Widder, *The Heat Equation*, Academic Press, New York, 1975.
11. F. John, *Partial Differential Equations*, Springer-Verlag, New York, 1975.
12. F. John, Numerical solution of the equation of heat conduction at preceding times, *Ann. Math.* **17**, 1955, 129–142.
13. L. Hormander, *The Analysis of Linear Partial Differential Operators I*, Springer-Verlag, New York, 1983.
14. H. Shvaytser and S. Peleg, Pictures as elements in a vector space, *IEEE Conf. Computer Vision and Pattern Recognition*, June 1983; see also Shvaytser and Peleg, Inversion of picture operators, *IEEE Conf. on Systems, Man, and Cybernetics*, Bombay and New Delhi, December 1983, p. 202–206; *Pattern Recognit. Lett.*, in press.
15. A. Carasso, J. Sanderson, and J. Hyman, Digital removal of random media image degradations by solving the diffusion equation backwards in time, *SIAM J. Numer. Anal.* **15**, 1978, 344–367.
16. H. Wozniakowski, A survey of information-based complexity, *J. Complexity* **1**, 1985.
17. F. Ratliff, *Mach Bands: Quantitative Studies on Neural Networks in the Retina*, Holden-Day, San Francisco, 1965.
18. A. Rosenfeld and A. Kak, *Digital Picture Processing*, Academic Press, New York, 1976.
19. N. N. Lebedev, *Special Functions and their Applications*, Prentice-Hall, Englewood Cliffs, N.J., 1965.