Human Neurobiol (1986) 5:121 – 128

**Human Neurobiology**


*Artificial intelligence and brain function (continued)*

# Receptive fields and the representation of visual information

S. W. Zucker and R. A. Hummel*

Computer Vision and Robotics Laboratory, Department of Electrical Engineering, McGill University, 3480 University Street, Montréal, Québec, Canada H3A 2A7

Received: 16 January 1986

**Summary.** Receptive fields in the retina indicate the first measurements taken over the (discrete) visual image. Why are they circular surround with an excitatory/inhibitory structure? We hypothesize that this provides a representation of the visual information in a form suitable for transmission over the optic nerve, a rather limited channel, that can then be extended into a variety of representations at the cortex. These cortical representations span a range of sizes and functionally separate positive and negative contrast data, precisely as is required for further processing. Our scheme is both physiologically and psychophysically plausible. In particular, we derive an explicit formula for constructing large receptive fields from small ones, and introduce the notion of *de-blurring* to derive interpolation filters for hyperacuity. A mathematical requirement of our scheme is a form of separation between positive and negative contrast data, a nonlinearity that we predict will agree with observations. Furthermore, the mathematics that we utilize are more naturally applicable to physiological models based on Gaussians than are (Fourier) spatial frequencies.

**Key words:** Visual representation – Receptive fields – Hyperacuity – Difference of gaussians – Deblurring – x-Pathway – Gabor functions

## 1. Introduction

The structure of receptive fields provides one of the most powerful constraints on visual information processing. They provide a system within which electrophysiologists can classify and compare their research, and they suggest properties of mechanisms. But how do they relate to abstract functional properties of the visual system? For example, why are some structured in a center-surround fashion? Why do they arise in both ON and OFF varieties? How do they support the communication of information from the retina to the cortex, and how can they account for hyperacuity? These are some of the questions we shall address in this paper. In general, all of the answers are related to schemes for representing visual information.

Our plan is to develop a mathematical model for the representation of spatial visual information[1]. Many of the approaches to assessing receptive field structure (and related psychophysics) are based on Fourier techniques: sines and cosines. However, sinusoids are artificial constructs which may not be the most appropriate language for the description of images. On the other hand, they do have certain attractive properties, such as the separation of high and low spatial frequencies and notions of linearity and superposition, to which we shall return. More recent attempts are based on different basis functions: those that come out of D. Gabor's theory of communication (Gabor 1946, Marcelja 1980). These are much closer to what we shall use, although the theoretical motivation is completely different. In particular, Gabor was interested in functions for encoding information that were optimal in the sense that they minimized a certain "uncertainty" relationship in both space and spatial frequency. Our functions are motivated by notions of blurring and de-blurring. While the result is quite similar in appearance, the precise mathematical forms differ.

### 1.1 Receptive fields and image representation

The layout of the visual system implies a need to communicate information between the retina and the cortex, but how does this communication take place? Whatever the process, the result in the cortex is not simply a *re-presentation* of the image sensed in the retina, as would be the case if the optic nerve were an array of perfect optical fibres; but rather is a *representation* of the image

*Present addresses:* *Courant Institute of Mathematical Sciences, New York University. New York, USA

[1] In this study we ignore issues of temporal processing (Fleet et al. 1984) and hence it only approximates those situations in which temporal effects are negligable.

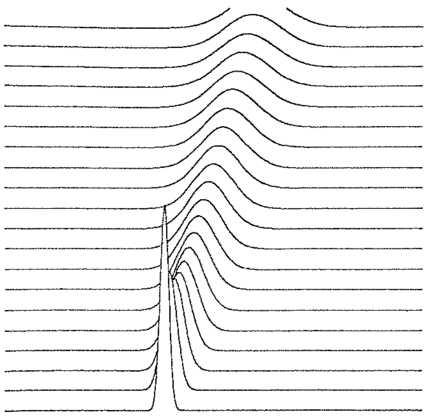

**Fig. 1.** An illustration of successive blurring of a signal. In this example the signal is 1-dimensional (along the x-axis). It is blurred by a Gaussian kernel again and again. The y-axis represents the effective amount of blur; i.e. the spatial parameter for the Gaussian. Note that increasing amounts of blur smooth over details of the initial data, and the peak appears to diffuse into a bump so wide that it will eventually approach a constant. An alternative interpretation of this figure is as a diffusion. Note how the sharp initial pulse (for $t \approx 0$) spreads into a wide, diffuse one for larger values of $t$

information in the form of a sample hierarchy. Such hierarchies arise from the successive application of blurring operators (Witkin 1983, Koenderink 1984); see Figure 1. While such hierarchies are certainly useful within efficient coding schemes (Srinivasan et al. 1982; Burt and Adelson 1983), how should they be constructed? Does the process of constructing the "larger" representations from the "smaller" ones lose information, either in theory or in practice?

Receptive fields constrain two aspects of early visual information processing: which measurements are taken over the raw retinal image, and how transformations of these measurements provide a representation of visual information rich enough to efficiently support subsequent processing. We shall concentrate on the X-pathway (Orban 1984), along which retinal receptive fields exhibit a circular surround organization with excitatory/inhibitory interactions[2]. The first stages of image analysis, such as orientation selection, take place in the cortex, which raises the problem of how precise visual information can be communicated from the retina onwards (Srinivasan et al. 1982). We propose a formal solution to this problem which leads to two principle observations. First, we derive an explicit formula for describing how larger receptive fields can be constructed from smaller ones. We posit certain non-linearities related to a separation of "positive" and "negative" contrast data. Secondly, it turns out that some degree of additional precision in the information can be obtained by a process of *de-blurring*, which could be relevant to hyperacuity. It also provides an explanation for the additional side-lobes found on smaller cortical receptive fields, which could serve, in functional terms, to aid in the precise localization of contours.

The paper is organized as follows. An overview of our model is presented next, followed by two large Sections. In the first of these (Section 3) we analyze the model mathematically, and in the second (Section 4) we apply it to study several different aspects of receptive field structure. Although much of the treatment in Section 3 is abstract, it is not necessary to follow all of the mathematics in detail. What is necessary are the motivations and intuitions behind it, since it is these that may provide a deeper understanding of certain aspects of receptive fields, as well as of the relationship between abstract models and physiology.

## 2. Overview of the model

We will present a description of the model immediately as a basis for study and subsequent discussion. In order to postpone issues of implementation, i.e., details of how the model maps onto physiology, the presentation is mathematical. However, we do have intuitions about the mapping, and will sketch possibilities throughout the paper as the details of the model are developed.

Let the light intensity distribution imaged on the (retinal) receptor surface be given by $f(x,y)$. The light distribution is not available directly, however. Rather, the initial samples are obtained from $(Df)(x,y)$, where $D$ is a differential operator. For our discussion, we will take $D$ to be a Laplacian operator: $\Delta = (\partial^2/\partial x^2 + \partial^2/\partial y^2)$. Motivated by physiology, we assume that there are two types of differential samples. OFF-center and ON-center, representing, roughly, the positive and negative parts of the Laplacian data. These very local samples are then independently blurred by local weighted averages to multiple levels of resolution; see Figure 2. The initial

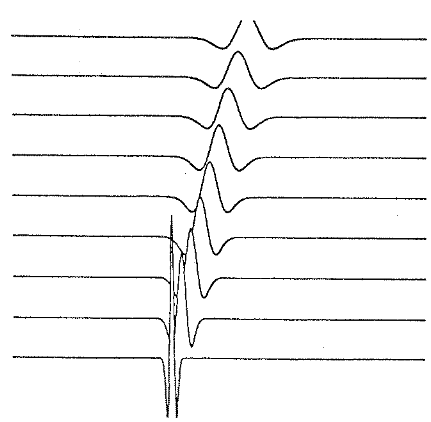

**Fig. 2.** Successive values of $v(x,t) = \Delta K(x,t)$, the Laplacian of a Gaussian, with increasing $t$ along the y-axis. The blur parameter $- t -$ also can be throught of as parameterizing receptive field size from small to large. For simplicity, in this illustration we did not separate the OFF and ON components, which should be thought of as comprising two separate channels

---

[2] In the cortex receptive fields take on an elongated structure, with approximately Gaussian structure in the elongated direction and the excitatory/inhibitory structure in the perpendicular direction. We shall be concerned with constructive mechanisms for both of these types of structure.

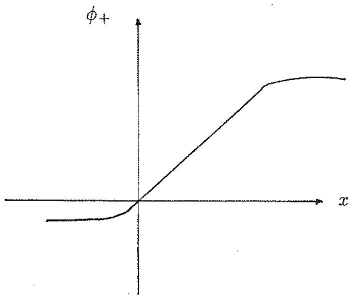


**Fig. 3.** The "positive part" function $\phi(\lambda)$. It is used to separate the positive and negative parts of the Laplacian measurements into separate channels. While the function never goes below 0 mathematically, it is shown here passing through an axis that can be thought of as the background or resting firing rate of a neuron. Values above the axis indicate action potentials that are more frequent than the resting firing rate, while those below are less frequent

measurements and separation into OFF- and ON-center data occur in the retina. Some blurring can take place within the retina as well, although most of the combination into larger receptive fields takes place in the cortex.

To be more precise, we model the separation into OFF- and ON-channel data by a non-negative function $\phi(\lambda)$ which is small but positive for $\lambda = 0$, which increases linearly as $\lambda$ increases until some value where $\phi(\lambda)$ saturates. $\phi(\lambda)$ decreases to 0 as $\lambda$ decreases below $\lambda = 0$ (Figure 3). The function $\phi$ models the firing rate of neurons in, say, the optic nerve. Since firing rate is always a positive number, $\phi(\lambda) \geq 0$. However the curve passes through a significant value when $\phi = \phi_0$, the resting firing rate.

The value $\phi(\Delta f(x,y))$ thus represents an approximation to the positive part of $\Delta f(x,y)$, while $\phi(-\Delta f(x,y))$ is an approximation to the negative part. Our samples will be given by

$$S_{\mathrm{OFF}}(x,y) = \phi(\Delta f(x,y))$$

$$S_{\mathrm{ON}}(x,y) = \phi(-\Delta f(x,y)) \, .$$

These samples are then independently blurred by local weighted averages to obtain the data $v(x,y,t)$ at multiple levels of resolution. In particular,

$$v_{\mathrm{OFF}}(x,y,t) = \iint K(x-x', y-y', t)\, S_{\mathrm{OFF}}(x',y')\, dx'\, dy' = K(x,y,t) * S_{\mathrm{OFF}}$$

and

$$v_{\mathrm{ON}}(x,y,t) = \iint K(x-x', y-y', t)\, S_{\mathrm{ON}}(x',y')\, dx'\, dy' = K(x,y,t) * S_{\mathrm{ON}}$$

where $K(x,y,t)$ is a blurring kernel (typically a Gaussian) in which the amount of blur is parameterized by $t \geq 0$ (Figure 2). As we shall see, these equations are quite important to the theory, because they dictate the mechanism by which large receptive fields are built up from small ones: by a convolution process of Gaussian blurring.

## 3. Mathematical background and analysis

In this section we develop some of the mathematical analysis necessary to understand the structure and power of our model. As you will see, it differs substantially from the Fourier-type analyses prevalent in electrophysiology and psychophysics. Our purpose, in addition to developing the model, is to illustrate the naturalness of Gaussian-related functions for such applications. Witkin (1983) has shown that the Gaussian enjoys interesting uniqueness properties as well.

### *3.1 Blurring and the heat equation*

Our scheme is derived from a diffusion process in which (temporal) spread will become analogous to (spatial) extent of receptive fields[3]. It is formally based on the heat equation, the simplest such diffusion which has all of the necessary mathematical properties. The basic assumption carried by the heat equation is that, for a class of functions, certain spatial differentials (Laplacians) will be formally equivalent to certain (i.e., the first) temporal derivatives. We shall begin by arguing intuitively for this assumption.

Observe that, for a long conducting wire, a unit impulse of heat diffuses into increasingly larger Gaussian distributions as time proceeds. Mathematically, let $f(x)$ denote the initial temperature distribution as a function of the spatial variable $x \in \mathfrak{R}^n$. (Clearly we are interested in the special case when $n = 2$.) Then a solution to the heat equation $u(x,t)$ giving the temperature level as a function of position $x$ and positive time $t$, satisfying

$$u(x,0) = f(x)$$

can be obtained from the convolution

$$u(x,t) = \int_{\mathfrak{R}^n} K(x-x', t) f(x)\, dx' = K(x,t) * f(x)$$

where $K(x,t)$ is the *source kernel* (Widder 1975):

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

Note that this source kernel is just a Gaussian. Since it acts a blurring operator, we can regard the distributions $u(x,t)$ as representing continuously coarser representations of the original data $f(x)$ as $t$ increases. Referring to Figure 1, observe that the successive blurring can now be

[3] That is, what is normally thought of as the time parameter will become a spatial "blur" parameter, as will become clear shortly.

interpreted as a diffusion of the sharp initial pulse into a wide, diffuse one. In fact, assuming $f(x)$ is bounded, $u(x, t)$ as given above is entire analytic. It is the unique bounded solution to the heat equation

$$u_t = \Delta u$$

satisfying $u(x,0) = f(x)$, where $\Delta$ denotes the spatial Laplacian operator $(\partial^2/x^2 + \partial^2/y^2)$ and $u_t$ denotes $\partial u/\partial t$. Other unbounded solutions are technically possible, but the function $u(x, t)$ given by convolution against the Gaussian kernel $K$ is the one that naturally occurs in physical systems.[4]

### *3.2 The smallest non-zero operator*

It is important to note in the equations above that the initial data $f(x) = u(x,0)$ is not blurred at all, and that the parameter $t$ increases continuously from 0. This, of course, could never be realized physically; it can only be approximated finitely. While most of these approximations will not cause problems, one requires special attention: the smallest receptor (size $t = \tau > 0$) that is realizable physically. This operator will become analogous to the smallest receptive fields, and mathematically forces us to think not only of blurring (increasing $t$) but also of beblurring (decreasing $t$).

### *3.3 De-blurring and backwards solutions*

Suppose we take the temperature distribution as representing the image data, but blurred by the Gaussian kernel. Is it possible to reproduce the original data? Specifically, given $g(x) = u(x, \tau)$, for some fixed $\tau > 0$, is it possible to solve the heat equation backwards to recover $u(x, t)$ for $0 < t \leq \tau$? Can $f(x) = u(x,0)$ be recovered? This is the problem of deblurring Gaussian blur.

There are two separate aspects to the answer: whether recovery is possible in principle and whether it is possible in practice. In principle it can be shown that necessary and sufficient conditions for the existence of a solution to the heat equation, $u(x, t)$, $0 < t \leq \tau$, satisfying $u(x, \tau) = g(x)$, $x \in \Re^n$, are that $g(x)$ be analytic, and that the extension of $g(x)$ to an analytic function of several complex variables $g(z)$, $z \in C^n$, satisfy certain growth conditions (John 1955). Both of these conditions, analyticity and bounded growth, fit smoothly into the vision context. Thus deblurring is possible in principle. The question of whether de-blurring can actually be accomplished in practice raises other issues, however, to which we now turn.

[4] In this paper we shall concentrate on the Gaussian kernel as the blurring operator and the heat equation as the partial differential equation (diffusion equation). However the analysis that we do can be extended to other kernels and related partial differential equations. Thus, should the Gaussian turn out to be related only approximately to receptive fields, the structure of our model will still hold.

### *3.4 Stability and positive/negative separation*

Because we have used the nonlinear function $\phi(\lambda)$ in the definition of the primitive sampling elements, it is difficult to analyze the behavior of the model in terms of the spectral characteristics of receptive fields, or in standard analytical terms. This early nonlinearity destroys, for example, superposition.[5] However, interesting behavior within any model generally depends on nonlinearities, and they certainly exist physiologically. Placing the nonlinearity early in the model has certain aesthetic attractions, and certainly does not preclude feasibility.

The potential benefits from using positive- and negative-part nonlinearities in the model are substantial. They accrue from the effects of trying to undo agglomeration: e.g., of deblurring. To illustrate, consider the problem of reconstruction of a sinusoid from blurred samples. When blurred by a Gaussian, a sinusoid transforms into another sinusoid of the same frequency but with smaller amplitude:

$$u(x,1) = K(x,1) * \sin(\omega x) = A(\omega)\sin(\omega x)$$

where

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

and

$$A(\omega) = e^{-\omega^2}.$$

In other words, if we wish to deblur a sinusoid, we must multiply the amplitude by $1/A(\omega) = e^{\omega^2}$. The difficulty is that we may not know the exact value of $\omega$. Worse, $\omega$ can be arbitrarily large. Thus for this fixed amount of blur, arbitrarily large amounts of attenuation may have taken place. In particular, if we wish to deblur a signal which is nearly 0 with some minor perturbations, we can't tell whether the original signal was a fairly smooth one that survived the blur, or was a very large, high frequency sinusoid that has been massively attenuated. The difficulty can be summed up by noting that arbitrarily small errors in the representation of the blurred data can lead to large changes in the deblurred reconstruction. This is what is referred to mathematically as instability. Gaussian deblurring suffers from it.

There are several ways around this instability problem, given that blurring is to be considered inevitable. Our model incorporates a dynamic range limitation, in that the initial receptors are assumed to saturate at some level (where $\phi(\lambda)$ saturates), and separates the range into positive and negative parts. Intuitively this is related to stability as follows. The difficulty with the high frequency sinusoids is that the positive and negative portions quickly blend to cancel out each other. If the

[5] A form of "superposition" still holds for the linear portion of $\phi(\lambda)$, which suggests a number of physiological experiments using Gaussian- and difference-of-Gaussian-probes. These will be developed in a subsequent paper. Some involve combinations of stimuli as in Watson et al. (1983).

initial data is non-negative, then there is less cancellation and significant features are better retained through the blurring process. Of course, within our model the initial intensity data is non-negative, but it is transformed into signed data by the Laplacian operation. The positive and negative parts are separated, by $\phi(\lambda)$, to avoid cancellation during the blurring process.

These intuitions can be given a more precise mathematical formulation; recall the previous discussion of backsolving the heat equation. In a classic paper on the subject, F. John (1955) showed that, in addition to the mathematical conditions required for the existence of a backsolution, if a *nonnegative* backsolution exists, then stable reconstruction of $u(x, t)$ is possible for $a \le t \le \tau$, where $a > 0$. The degree of stability depends on how small $a$ is chosen; i.e., on how much deblurring is attempted, and on the maximum value $\mu$ in the blurred signal; i.e., $0 \le g(x) \le \mu$. For a fixed $a$, John shows that the error in the blacksolution is bounded by a constant (depending on $a$) times $\mu^{(1-\theta)}\varepsilon^{\theta}$. Here $\varepsilon$ is the error in the representation of the blurred data, and $\theta$ is a constant strictly between 0 and 1. Thus for a fixed permissable amount of error $E$ in the (partial) deblurring to a specified $a > 0$, and a fixed upper bound on the blurred signal $g(x)$, the function $g(x)$ will have to be approximated by its representation $\varepsilon$ given by some small constant times $E^{1/\theta}$. Since $0 < \theta < 1$, this says that extremely accurate representation of $g(x)$ will be needed to achieve accurate partial deblurring. This is a kind of "polynomial stability", since, for $1/\theta < N$, we have

$$\varepsilon \le C \cdot E^N.$$

This is not as good as a usual notion of bounded linear stability ($N = 1$), but is better than the completely unstable situation that exists with Gaussian deblurring in the absence of the non-negativity assumption.

Returning to our model, the advantage of separating positive and negative data should now be clear: it permits stable backsolution and deblurring. We now return to vision.

## 4. On the structure and function of receptive fields

The strengths of models are in their explanatory and predictive power. We consider some of these in this section, beginning with electrophysiological points and then returning to more mathematical ones.

### *4.1 Why do simple cells have inhibitory flanks?*

Simple cells are those in which the receptive field can be decomposed into separate excitatory and inhibitory areas (Hubel and Wiesel 1977). While there is a great deal of variability within the shape of these receptive fields, for our purposes we can consider an idealized one as shown in Figure 4. It's envelope consists of two main structures:

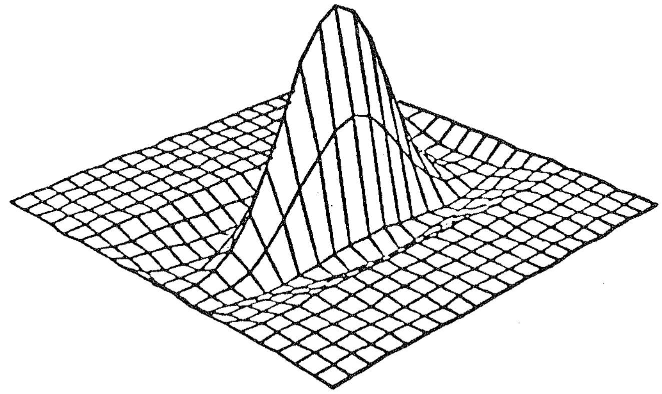

**Fig. 4.** An idealized simple cell receptive field. It consists essentially of a Gaussian envelope in the preferred orientation and a difference of Gaussians across it. Since such cells play a role in orientation selection, the Gaussian envelope can be interpreted as integrating information along the preferred direction, while the difference of Gaussians can be interpreted as "deblurring" information in the perpendicular direction

a Gaussian envelope in the preferred orientation, and a difference of Gaussians across it. Since such cells play a role in orientation selection, it is instructive to consider the function of such a receptive field operator when convolved against a thin line. The Gaussian envelope can be interpreted as integrating information along the preferred direction, providing a maximal response to the oriented stimulus. However, various imaging and neural blurring processes can certainly diffuse the line into a thicker one, and the difference of Gaussians across the receptive field can be interpreted as "deblurring" information. That is, the cross-section profile can function to effectively focus the line into a thinner one. While this is only one aspect of the orientation selection computation (Zucker 1985), it does explain two properties of the shape of these receptive fields that mesh nicely with the other kind of theory.

### *4.2 Hyperacuity and stable backsolutions*

A second application of deblurring ideas is related to the precision with which we can perform various visual tasks. Our visual acuity is given by retinal receptor spacing: if the frequency of a sinusoid is higher than the Nyquist sampling rate derived from this spacing, then the individual fluctuations cannot be resolved. However people can perform tasks (such as vernier alignment) that require spatial resolution higher than this Nyquist rate, an ability referred to as hyperacuity (Westheimer and McKee 1977). One way to account for hyperacuity is to assume the capability of interpolating values of an "image" distribution using the measured sample values. In terms of our model, the available samples are the blurred local differential data separated into positive and negative parts. However, recalling the previous discussion, deblurring is theoretically possible if John's assumptions – including non-negativity – are met. Stable deblurring, moreover, is only possible back to some extent (the constant $a > 0$). Thus our model would suggest that the interpolation filters for deblurring are playing a role in hyperacuity. This is novel both in a functional sense and because it further suggests an answer to the question of why hyperacuity is only as good as it is, and not better! Beyond this point the process becomes unstable.

It is instructive to examine these deblurring filters; i.e., operators for computing backsolutions to the heat equation by convolution, in more detail (Kimia et al. 1984). For technical reasons it is only possible to find a pseudo-inverse to the general Gaussian blur operator. Pseudo-inverses have an order associated with them, so that, intuitively, higher order approximations are capable of deblurring more complex signals; i.e., signals containing more terms in their series expansions. Now, as the order of the pseudo-inverse increases, the deblurring filter acquires more sidelobes: see Figure 5. Such additional sidelobes have been observed physiologically (Movshon et al. 1978, Wilson et al. 1983, 1984), but only on the smaller receptive fields! These are precisely the ones for which our theory predicts they are necessary. Large receptive fields incorporate so much inherent blur that high-order deblurring is completely unnecessary.

Our scheme differs in basic technical ways from others proposed to account for the communication gap and hyperacuity. One class is based on $(\sin x)/x$ reconstruction filters (Barlow 1979, 1981, Crick et al. 1981). The numerical analysis of such filters shows that they require too much spatial support (i.e., several lobes on either side) to function properly with the limited support apparently available (Hummel 1983). In the context of John's stability results, ours is based on Hermite polynomials (Kimia et al. 1984). Not only are these more local than $(\sin x)/x$ filters, but they are derived from Gaussians and their derivatives, the natural set of mathematical basis functions to use in the context of Gaussian receptive fields. They resemble the (even) Gabor functions in shape, although our theory specifies why only certain of them (rather than all) are present. Purely from the point of image communications, leaving aside issues of interpolation, our scheme has a lot in common with, and we have benefitted a great deal from, the discrete "Laplacian pyramids" developed by Burt and Adelson (1983). There is a sense in which our theory provides the foundations for, and a continuous analog of, theirs. However, without our mathematical analysis the connections to contrast separation, de-blurring, and interpolation would not have been clear.

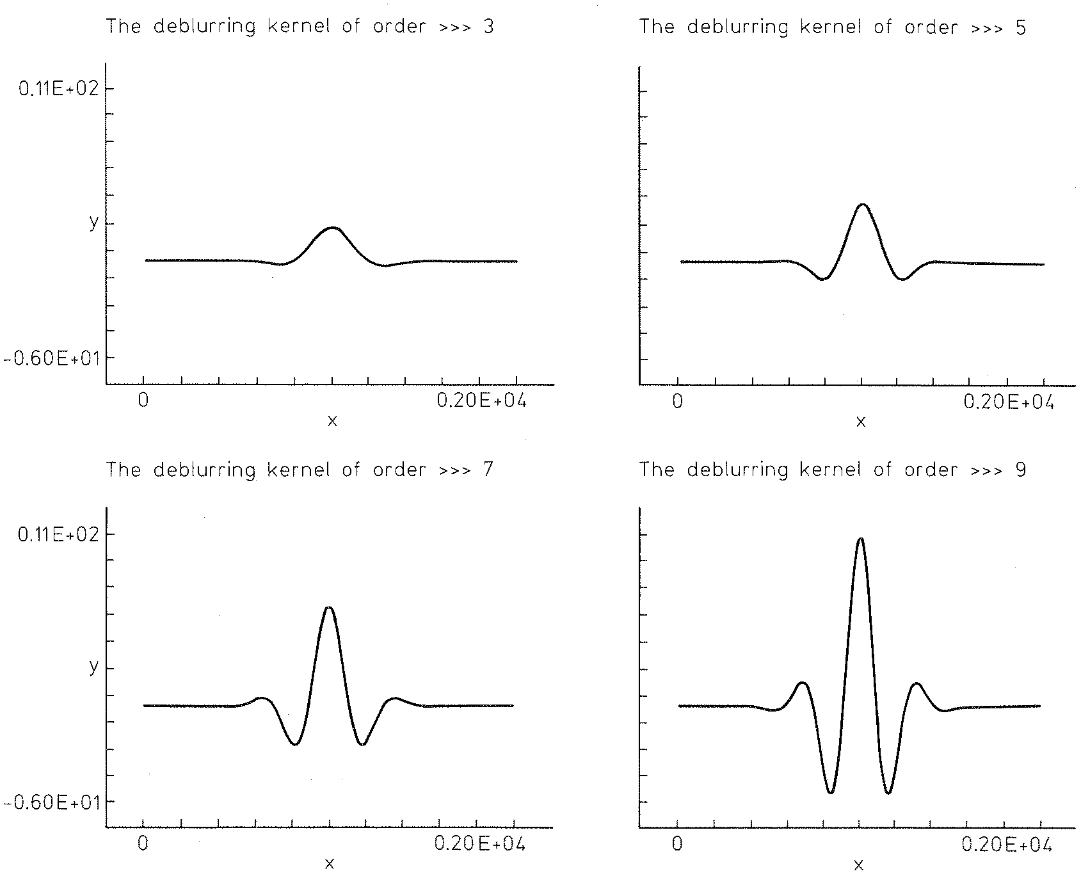


**Fig. 5.** One-dimensional deblurring kernels of order 3, 5, 7, and 9. Note how, as the order increases, the number of side-lobes increases as well. Such kernels are visually indistinguishable from certain Gabor functions, and hence from the cross-section through certain simple cells. The theory predicts that it is only the smaller kernels that require side lobes, exactly as has been observed physiologically

### 4.3 *Difference of Gaussian receptive fields*

Circular surround receptive fields have been modeled by kernels given either as the difference of two Gaussians (Rodieck 1965, Enroth-Cugell and Robson 1966), or as the Laplacian of a Gaussian (Marr and Hildreth 1980). Although these kernels are distinct, we can interpret the former as a discrete analog of the latter. This follows since the heat kernel $K(x,t)$, a solution of the heat equation, satisfies

$$\Delta K(x,t) = \frac{\partial K(x,t)}{\partial t} \approx \frac{K(x,t_1) - K(x,t_2)}{(t_1 - t_2)},$$

which is the difference of two Gaussians. We therefore take

$$v(x,t) = \int_{\Re^n} \Delta K(x-x',t) f(x')\,dx'$$

as a continuous parameterization of variable receptive field sizes. As $t$ increases, the spread of the Gaussian increases, which implies that the size of the receptive field increases. It is precisely these operators that were plotted in Figure 2.

It should be noted that this formulation, using the Laplacian of a Gaussian kernel, corresponds to a difference of Gaussians scheme in which the two Gaussians have nearly the same extent, i.e., $t_1 \approx t_2$. Interestingly, the (Laplacian of a Gaussian) or difference of two similar Gaussians can be obtained by Gaussian blurring the difference of two dissimilar but very local Gaussians. Observe:

$$\begin{aligned}&[K(x,t+\varepsilon_1) - K(x,t+\varepsilon_2)] * f \\ &\quad = K(x,t) * [K(x,\varepsilon_1) - K(x,\varepsilon_2)] * f.\end{aligned}$$

If $t \gg \varepsilon_1$ and $t \gg \varepsilon_2$, then $(t+\varepsilon_1) \approx (t+\varepsilon_2)$ as required.

Our scheme differs from other difference of Gaussian schemes (e.g., Burt and Adelson 1983, Marr and Hildreth 1980) in which the separation between the two Gaussians $(t_1 - t_2)$ increases at coarser resolutions.

### 4.4 *Reconstruction in principle: how much information is available in receptive fields?*

The continuous family of measurements suggests a way to obtain the initial image data $f(x) = u(x,0)$ from them. While this scheme is probably only of theoretical interest, it does provide an approach to determining the completeness of the representation. It will also lead to a formula suggestive of how large receptive fields can be constructed from smaller ones. Since

$$\begin{aligned}v(x,y,t) &= \Delta K * f \\ &= \Delta(K * f) \\ &= K * \Delta f,\end{aligned}$$

$v(x,t)$ can be interpreted either as the Laplacian of the blurred intensity data, i.e., as $\Delta(K \cdot f) = \Delta u(x,t)$, or as the bounded solution to the heat equation using the initial data $\Delta f(x)$. From the former interpretation and the fact that $u(x,t)$ satisfies the heat equation, we have that $v(x,t) = \Delta u(x,t) = \partial u(x,t)/\partial t$, so

$$-\int_0^T v(x,y,t)\,dt = u(x,y,0) - u(x,y,T).$$

Now, $u(x,T)$ is nearly constant for sufficiently large $T$, so the above integral can be used to recover $f(x)$ modulo an additive constant. That is, if the entire family of measurements $v(x,t)$, $t \in [0,T]$ were available, then the original could be trivially reconstructed (up to its mean) by simply adding them up.

The above reconstruction scheme requires all measurements, from $v(x,y,0)$ to $v(x,y,T)$. How can the smallest, physiologically unrealizable ones (those smaller than $t = \tau$) be obtained? The alternate interpretation of $v(x,t)$ above also shows how we can obtain these values. Since $v(x,t)$ is itself a solution to the heat equation, the values of $v(x,t)$, for $0 < a \le t \le \tau$, can be obtained by backsolving the heat equation using $v(x,\tau)$ as initial data. These backsolved data can then be used to evaluate the above integral. This is the same point made previously in the discussion of hyperacuity.

Of course, our model separated the initial data into two channels, one positive and the other negative, and then blurred each channel separately. But approximately,

$$S_{\mathrm{ON}}(x,y) - S_{\mathrm{OFF}}(x,y) \approx (\Delta f)(x,y).$$

Thus

$$\begin{aligned}v_{\mathrm{ON}}(x,y,t) - v_{\mathrm{OFF}}(x,y,t) &= K(x,y,t) \\ &\quad * [s_{\mathrm{ON}}(x,y) - s_{\mathrm{OFF}}(x,y)] \\ &\approx K(x,y,t) * \Delta f \\ &= v(x,y).\end{aligned}$$

So the preceeding discussion of reconstructing $f(x,y)$ from $v(x,y)$ applies to theoretical reconstruction from data supplied by our model by setting

$$v(x,y,t) = v_{\mathrm{ON}}(x,y,t) - v_{\mathrm{OFF}}(x,y,t).$$

### 4.5 *Construction of large receptive fields*

The previous equation, $v(x,t) = K * \Delta f$, to emphasize connections with Section 2, also shows how to construct larger receptive fields, out of smaller ones: simply convolve them with a Gaussian $K(x,t)$ for a suitably large $t$. (Actually the positive and negative parts must first be separated; see Section 2.) This is how our model constructs effective receptive fields from the initial local measurements. It should be stressed that the pure

Laplacian operating on the pure image data $\Delta f(x)$ is a mathematical idealization. In practice, the construction of $\phi(\Delta f)$ and $\phi(-\Delta f)$ can take place using local discrete approximations.

## 5. Summary and conclusions

In summary, in this paper we attempted to lay out a new approach to mathematically studying receptive field structure. As in other approaches, blurring motivated the approach, with larger receptive fields successively constructed from smaller ones. We began with initial image intensity measurements, but these were immediately combined with a local differential operator. Thus both positive and negative data arise. At this point our approach takes on a significant difference from others, in that we then separated the data into positive and negative parts. These positive and negative parts were treated separately, with larger receptive fields built by blurring smaller ones within each contrast channel.

The separation of positive and negative data has a number of advantages, given the mathematical connection between blurring operators and differential equations. While it is not critical, we focused on Gaussian blurring and the heat equation to facilitate analysis. It then became possible to consider issues of deblurring and backwards solutions of the heat equation. This lead to conjectures about the shape of receptive fields and hyperacuity. The mathematical issue that we introduced was stability, which shed new light on the reasons why positive and negative data should be sêparated.

The model provided tools for constructing and understanding receptive fields. We were able to provide functional explanations for the antagonistic structure of simple cell receptive fields orthogonal to their preferred orientation (accurate positioning of lines); and also for the extra side lobes found in smaller receptive fields (deblurring and interpolation for hyperacuity). This contributed not only to understanding putative mechanisms for hyperacuity, but also suggested why it is only as good as it is, and not better. Finally, in order to study how much information is stored receptive fields, we also developed a reconstruction technique that works in principal.

But there is certainly more to receptive fields than just the representation of visual information. The most pressing questions relate to how our visual system infers the structure of the world. The present theory compliments these latter investigations by providing functional constraints: how much deblurring is possible; how much hyperacuity is possible, and so on. And it does it in a language that seems more natural for this purpose than, say, Fourier basis functions. Whether the analysis techniques that we propose will stand up to more detailed mappings onto physiology still remains. The more precise the mapping, the tighter the constraints.

*Acknowledgements.* This research was sponsored by the Natural Science and Engineering Research Council (Canada), grant A4470, and by the National Science Foundation (USA), grant DCR-8403300. We thank Yvan Leclerc for comments on the manuscript. This report is a revised version of TR 83-17. Computer Vision and Robotics Lab., McGill University, 28 September, 1983. Some of the earlier material has been presented at the *Canadian Society for Computational Studies of Intelligence*, London, Ont., May, 1984, and at the *Seventh International Conference on Pattern Recognition,* Montreal, July, 1984. S. W. Zucker is a Senior Fellow of the Canadian Institute for Advanced Research.

## References

Balow H (1981) Critical limiting factors in the design of the eye and the visual cortex. Proc Roy Soc Lond [Biol] 212:1 – 34

Barlow H (1979) Reconstructing the visual image in space and time. Nature 279:189 – 190

Burt P, Adelson E (1983) The laplacian pyramid as a compact image code. IEEE Trans Comm COM 31:532 – 540

Crick F, Marr D, Poggio T (1981) An information processing approach to understanding the visual cortex. In: Schmitt, Warden, Aldeman, Dennis (eds) The organization of the cerebral cortex. MIT Press, Cambridge, Mass Oxford

Enroth-Cugell C, Robson J (1966) The contrast sensitivity of retinal ganglion cells of the cat. J Physiol (Lond) 187:517 – 552

Fleet D, Jepson A, Hallett P (1984) A spatio-temporal model for early visual processing. RCBV-TR-84-1. Computer Science Dept University of Toronto

Gabor D (1946) Theory of communication. Proc IEE (Lond) 93:429 – 457

Hubel D, Wiesel T (1977) Functional architecture of macaque monkey visual cortex. Proc Roy Soc Lond [Biol] 198:1 – 59

Hummel R (1983) Sampling for spline reconstruction. SIAM J Appl Math 43:278 – 288

John F (1955) Numerical solution of the equation of heat conduction for preceeding times. Ann Matematica 17:129 – 142

Kimia B, Hummel R, Zucker SW (1986) Deblurring gaussian blur. IEEE Conf. Computer Vision and Image Proc, Washington, DC, June, 1983; Technical Report 83-15. Computer Vision and Robotics Lab, McGill University, Montreal. Computer Vis Graph Image Process (in press)

Koenderink JJ (1984) The structure of images. Biol Cybernetics 50:363 – 370

Marcelja S (1980) Mathematical description of the responses of simple cortical cells. J Opt Soc Am 70:1297 – 1300

Marr D, Hildreth E (1980) Theory of edge detection. Proc Roy Soc Lond [Biol] 207:187 – 217

Movshon J, Thompson I, Tolhurst D (1978) Spatial summation in the receptive fields of simple cells in the cat's striate cortex. J Physiol (Lond) 283:53 – 77

Orban G (1984) Neuronal operations in the visual cortex. Springer, New York Berlin Heidelberg

Rodieck R (1965) Quantitative analysis of cat retinal ganglion cell response to visual stimuli. Vis Res 5:583 – 601

Srinivasan M, Laughlin S, Dubs A (1982) Predictive coding: a fresh view of inhibition in the retina. Proc Roy Soc Lond [Biol]: 427 – 459

Watson A, Barlow H, Robson J (1983) What does the eye see best? Nature 302:419 – 422

Westheimer G, McKee S (1977) Spatial configurations for visual hyperacuity. Vis Res 17:941 – 947

Widder D (1975) The heat equation. Academic Press, New York

Wilson H, Gelb D (1984) Modified line-element theory for spatial-frequency and width discrimination. J Opt Soc Am [A] 1:124 – 131

Witkin A (1983) Scale space filtering. Proceedings of the 8th International Joint Conference on Artificial Intelligence. Karlsruhe, FRG, pp 1019 – 1022

Zucker SW (1985) Early orientation selection: tangent fields and the dimensionality of their support. Computer Vis Graph Image Process 32:74 – 103