Spatial Vision, Vol. 2, No. 4, pp. 273–293 (1987)

The Gaussian derivative model for spatial vision: I. Retinal mechanisms

RICHARD A. YOUNG
Computer Science Department, General Motors Research Laboratories, Warren, Michigan 48090–9055, USA

Received 9 December 1985; revised 22 May 1987; accepted 30 September 1987

Abstract—Physiological evidence is presented that visual receptive fields in the primate eye are shaped like the sum of a Gaussian function and its Laplacian. A new 'difference-of-offset-Gaussians' or DOOG neural mechanism was identified, which provided a plausible neural mechanism for generating such Gaussian derivative-like fields. The DOOG mechanism and the associated Gaussian derivative model provided a better approximation to the data than did the Gabor or other competing models. A model-free Wiener filter analysis provided independent confirmation of these results. A machine vision system was constructed to simulate human foveal retinal vision, based on Gaussian derivative filters. It provided edge and line enhancement (deblurring) and noise suppression, while retaining all the information in the original image.

INTRODUCTION

A major trend in image processing is the convergence of biological and machine vision approaches to understanding the basic principles of image-analyzing systems. Insights into how the eye and brain organize visual data may provide novel and powerful computational paradigms for image processing.

Mach (1868) first recognized that not just light intensities, but intensity *changes* (i.e. derivatives) are important for influencing what we see. Assume a visual scene with a light intensity distribution $I = f(x, y)$. For a point in that scene, Mach (1868) stressed that the output of the retina at a corresponding point is given not just by I, but also by the sum of the second derivatives of I with respect to x and y, which Mach (1906) recognized was the Laplacian of I or $\nabla^2 I$. That is, what we see is influenced not just by I, but by its second derivatives. Indeed, enhancing the second derivatives of a picture (by adding the negative of the Laplacian of the picture to the picture itself) markedly improves overall sharpness (Kovasznay and Joseph, 1953). In photography this technique is known as 'unsharp masking' (Pratt, 1978; Cannon and Hunt, 1981), seen by adding a defocussed negative to a normal positive.

Marr and Hildreth (1980) proposed Gaussian filtering to reduce noise, in addition to the Laplacian operation. They termed this combination a Laplacian of a Gaussian or a $\nabla^2 G$ operator. Their development of this operator was inspired in part by a widely-known description of cat ganglion cell receptive fields (Rodieck and Stone, 1965; Enroth-Cugell and Robson, 1966), also proposed for humans (Schade, 1959), termed the 'difference-of-Gaussians' or DOG mechanism. (The ganglion cells form the optic tract, and so contain the final message sent from the eye to the brain. A 'receptive field' is defined as the local region in the overall field of view in which a cell responds to light.) The DOG mechanism describes ganglion cell receptive field shapes in terms of two concentric positive and negative Gaussians with different standard deviations. Marr and Hildreth (1980) showed that with certain ratios of standard deviations, and with equal volumes of center and surround Gaussians, such a DOG filter is a close approximation to a $\nabla^2 G$ filter.

However, a quantitative test of the $\nabla^2 G$ model with actual biological data has not previously been made. I now report such tests, with data from three species—monkey, cat, and human. Several new findings emerged. First, a Gaussian term had to be added to the $\nabla^2 G$ term to describe ganglion cell data. Second, a new paradigm called a DOOG (difference of *offset* Gaussians) was discovered as a description of the underlying neural mechanism for generating such fields. The DOOG mechanism provided a closer match to the receptive field structure of the predominant class of monkey ganglion cells (de Monasterio, 1978) than the DOG mechanism previously proposed for older neurophysiological data. Since the formal definition of a Gaussian derivative function is given mathematically in terms of the offset differences of a Gaussian function, it was hypothesized that retinal receptive field mechanisms describable in terms of DOOG mechanisms should provide a better approximation to a Gaussian derivative model of receptive field structure than to other competing models.

A rapid new image-processing algorithm was inspired by these biological findings, and experimentally tested to learn the possible functional usefulness of applying such operators to images.

METHODS

For a spatial system, a point-weighting function (PWF) is its response or sensitivity to a point source. It is useful to describe a biological visual system in terms of its PWF (or its line-weighting function LWF), because: (1) the output to any arbitrary intensity pattern can be derived, given the assumption of superposition; and (2) the behavior of the system can be duplicated by a machine using digital filters with identical weighting functions. Accordingly, a monkey PWF (de Monasterio, 1978) and cat LWF (So, 1979) for ganglion cells were selected for quantitative analysis. These are the only known examples of ganglion cell receptive field shape quantitative data measured directly in the spatial domain. A human LWF for the overall visual system (Fiorentini, 1972) was also selected for analysis. A monkey retinal cell LWF measured directly in the spatial domain has not yet been published to my knowledge, so an estimate was made by orienting a line parallel to (and therefore summing along) the y-axis in the 2-D PWF field mechanism, for different positions along the x-axis. The best-fit parameters for the Gaussian derivative, DOOG, DOG, Gabor (Kulikowski *et al.*, 1982) and Cauchy (Klein and Levi, 1985) models were sought for in all cases by the SAS NLIN non-linear regression curve-fitting program (SAS Institute, 1982), using equations given in Young (1985a).

To simulate the biological behavior in a machine, a Gaussian digital filter (Burt and Adelson, 1983) was implemented. It used the weights of 0.05, 0.25, 0.40, 0.25 and 0.05, representing five points along a Gaussian curve with $\sigma = 1$ and area $= 1$, for reasons given by Burt and Adelson (1983). (These weights were not critical—the binomial weights of 1, 4, 6, 4, 1 each divided by 16 in fact form an even better discrete approximation to a Gaussian.) A 2-D Gaussian convolution of an image was accomplished by the following algorithm. Let $I(i,j)$ be an m by m image intensity input matrix (where i is the row index and j is the column index); let $T(i,j)$ be a temporary matrix; and let $G(i,j)$ be the final Gaussian-filtered output matrix. Compute $T(i,j) = 0.40I(i,j) + 0.25\,[I(i-1,j) + I(i+1,j)] + 0.05[I(i-2,j) + I(i+2,j)]$ for i from