A Gaussian Derivative-Based Transform

Jeffrey A. Bloom and Todd R. Reed

Abstract—**This correspondence describes a new image transform that decomposes an image using a set of Gaussian derivatives. The basis functions themselves have been shown to effectively model the measured receptive fields of simple cells in the mammalian visual cortex. Based on these functions, it can be expected that this transform can provide a mechanism for exploiting the properties of the human visual system in image processing algorithms.**

I. Introduction

The quality of the results of image processing is often judged by human viewers. In these cases, errors are significant only if perceived. With an understanding and model of the human visual system (HVS), we can exploit its properties to achieve resulting images that *look* better.

One such model of the HVS is the Gabor model that represents the receptive fields of the visual cortex with Gaussian modulated complex exponentials. Like the receptive fields, the Gabor functions are spatially local and consist of alternating bands of excitation and inhibition in a decaying envelope. There has been activity in the last 15 years using Gabor functions as filters [1] and as basis functions for a transform to a spatial/spatial-frequency domain [2]. The Gabor transform has been applied to image coding and image sequence coding to exploit the HVS model by introducing quantization error in a way that is not visually offensive [3].

In 1987, Young proposed a receptive field model based on the Gaussian and its various derivatives [4]. These functions, like the Gabor, are spatially local and consist of alternating regions of excitation and inhibition in a decaying envelope. Young showed that Gaussian derivative functions more accurately model the measured receptive field data than do the Gabor functions [5]. As had been done with Gabor functions, the Gaussian derivatives have been used as filters [5]. In related work, Gaussian derivatives have been interpreted as the product of Hermite polynomials and a Gaussian window [6], where windowed images are decomposed into a set of Hermite polynomials. The image transform introduced in this correspondence decomposes an image using Gaussian derivatives and, like the Gabor transform, results in a spatial/spatial-frequency domain and is completely invertible.

From the image processing perspective, the locality of the Gaussian derivatives can be a significant benefit. This allows the processing of an image based on local image properties rather than on the global makeup. In image compression, for example, it is common to apply square, nonoverlapping blocks in order to parse the image into many smaller "local images." Each of these local images is then transformed with a global process such as the DCT. Gaussian derivatives are inherently local themselves, and thus, this blocking of the image is not necessary.

It is reasonable to expect that the local nature of the Gaussian derivatives and their ability to model the early response of the HVS will result in a transform that is well suited for image applications.

Manuscript received November 1, 1994; revised June 21, 1995. The associate editor coordinating the review of this paper and approving it for publication was Prof. Rama Chellappa.

The authors are with the Department of Electrical and Computer Engineering, University of California, Davis, CA 95616 USA.

Publisher Item Identifier S 1057-7149(96)01798-7.

Artifacts introduced through operations in the transform domain (such as quantization) will be localized to a particular region and will be distributed in harmony with the HVS, thus reducing artifact perception and offensiveness.

II. Gaussian Derivative Functions

The set of functions that are the various derivatives of the Gaussian are known to be the product of Hermite polynomials and the original Gaussian [4]. Since Hermite polynomials of various orders are orthogonal to each other, the set of Gaussian derivatives *at a particular location* are also orthogonal. The set of basis functions used in this Gaussian derivative transform includes Gaussians and their derivatives at multiple spatial locations. These functions will necessarily overlap in the spatial domain, and the overall set will be nonorthogonal.

Two issues of concern will be addressed here. First, we consider which functions of the Gaussian derivative family should be used as basis functions. We need to consider the spatial placement as well as choose the derivatives to be included. Second, we must address the nonorthogonality of the basis set and the accompanying challenge of computing the decomposition.

By allowing this transform to be separable, we can investigate these issues in 1-D and then make the necessary extension to 2-D.

III. Basic Set

The family of 1-D Gaussian derivatives centered at the origin can be defined as

$$g_0(x) = e^{-x^2/2\sigma^2} \tag{1}$$

$$g_n(x) = \frac{d^{(n)}}{dx^{(n)}} g_0(x). \tag{2}$$

This set can then be shifted to any desired location. The variance σ^2 defines the extent of the functions in the spatial domain. There is an inverse relationship between the spatial and spectral extents, and the value of this variable may be constant or may vary with context.

We can choose the basis set to cover both the spatial and spatial-frequency domains. In order to arrange the functions equally in an N-point domain (spatial or spatial-frequency), N/D centers are identified in which D is the spacing between centers. The functions in the shifted Gaussian derivative family are linearly independent, which implies that the basis set will be complete if the total number of functions in the decomposition is equal to the original number of data points. Thus, there should be D functions at each of the N/D locations.

The Gaussian derivative function spectra are bimodal (except for that of the original Gaussian, which is itself a Gaussian) with modes centered at $\pm\Omega_n$ rad/pixel as in (3), where n is the derivative order. This equation can be rearranged to denote the derivative necessary to center a spectral mode at Ω_n, as shown in (4)

$$\Omega_n = \frac{\sqrt{n}}{\sigma} \tag{3}$$

$$n = \Omega_n^2 \sigma^2. \tag{4}$$

Using (4) to identify the necessary derivatives and realizing that the maximum frequency representable is π rad/pixel, we can space the D functions equally in the spatial frequency domain. It is important to note that the D functions need not be equally spaced. There is evidence to suggest that octave spacing may be more appropriate

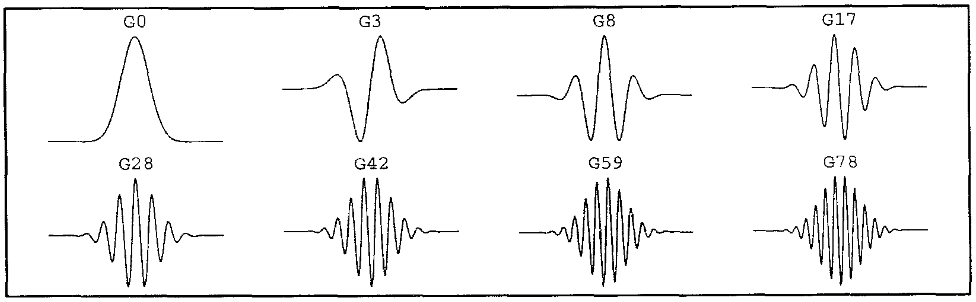

Fig. 1. Basis set.

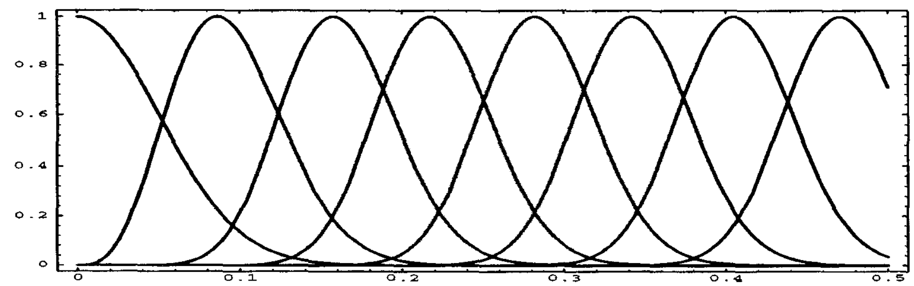

Fig. 2. Basis set spectra.

for models of the HVS. The current implementation uses a standard deviation of 3 pixels and equal spacing of $D = 8$ pixels, which results in the following eight derivatives: 0, 3, 8, 17, 28, 42, 59, and 78. These eight functions are shown in Fig. 1, and the magnitude of their spectra are shown, for positive frequency, in Fig. 2. The functions used are analytic derivatives of a noiseless Gaussian function.

IV. Decomposition and Reconstruction

Due to the nonorthogonality of the basis set, standard inner product techniques for calculating the decomposition are not appropriate. This transform is similar to the Gabor transform with regard to the nonorthogonality of the basis set. Approaches to Gabor decompositions have included neural networks [7], which are iterative, filter banks via biorthogonal functions [8], and a LMS linear systems approach [2]. This final approach will be applied to the Gaussian derivative basis set.

A. LMS Error Solution

The following derivation was presented by Ebrahimi *et al.* for decomposing images using the nonorthogonal Gabor functions [2]. As the authors point out, this technique is not dependent on the actual basis used and can be used to find the weighing coefficients for any type of basis functions.

The 1-D discrete signal $f(k)$ can be decomposed into a weighted sum of elementary functions $g_n(k)$, regardless of the orthogonality of the g_n

$$f(k) = \sum_{n=0}^{N-1} c(n) g_n(k) \qquad 0 \le k \le N-1. \tag{5}$$

The expansion can be written in matrix notation as

$$\underline{f} = G\underline{c} \tag{6}$$

where

$$\underline{f} = \begin{vmatrix} f(0) \\ \vdots \\ f(N-1) \end{vmatrix}, \quad \underline{c} = \begin{vmatrix} c(0) \\ \vdots \\ c(N-1) \end{vmatrix} \tag{7}$$

and

$$G = \begin{vmatrix} g_0(0) & \cdots & g_{N-1}(0) \\ \vdots & & \vdots \\ g_0(N-1) & \cdots & g_{N-1}(N-1) \end{vmatrix}. \tag{8}$$

From (6), we can define the error as

$$E = G\underline{c} - \underline{f} \tag{9}$$

and then, find the $\underline{c}$ that minimizes the squared error

$$E^2 = (G\underline{c} - \underline{f})^T (G\underline{c} - \underline{f}) \tag{10}$$

which can be shown to be

$$\hat{\underline{c}} = (G^T G)^{-1} G^T \underline{f} \tag{11}$$

$$\hat{\underline{c}} = A^T \underline{f}. \tag{12}$$

Note that this method is independent of the actual basis functions used as long as they are linearly independent. In addition, note that $A^T = (G^T G)^{-1} G^T$ is independent of the signal and can be precomputed. The decomposition becomes the matrix multiplication in (12), and the reconstruction becomes the matrix multiplication in (6).

This 1-D approach can be easily extended to two dimensions by considering separable 2-D functions. With the same G matrix described above, the signal $f(k_1, k_2)$ can be decomposed to yield the coefficients $c(n_1, n_2)$, and the matrix notation is

$$F = GCG^T. \tag{13}$$

The 2-D decomposition equation then becomes

$$\hat{C} = A^T F A \tag{14}$$

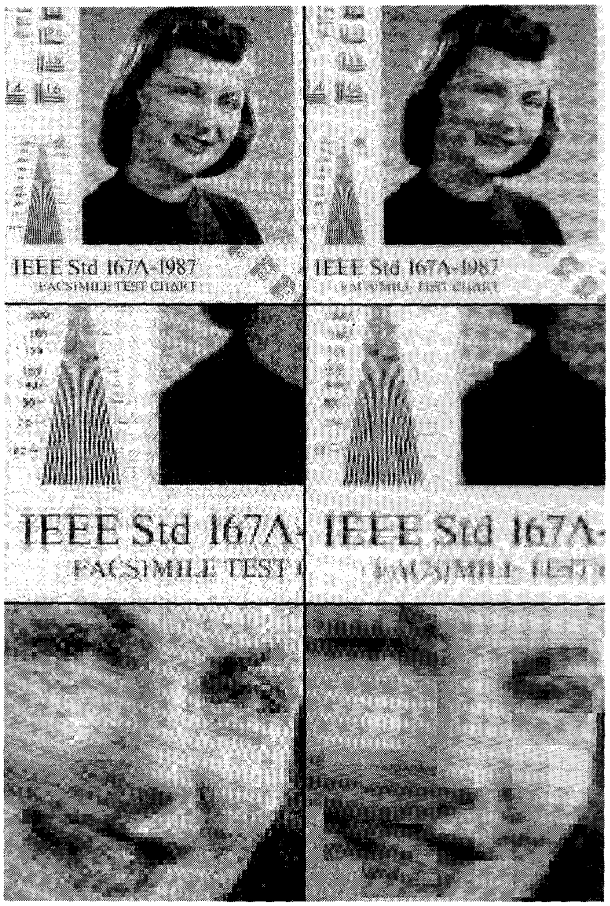

Fig. 3. Compressed images: 5 b, SNR = 19 dB, DGT (left) and DCT (right).

where $\boldsymbol{A}$ is as defined earlier. Once again, $\boldsymbol{A}$ can be computed offline, and the decomposition and reconstruction are matrix multiplications.

V. Image Coding

In an image coding application, the transform coefficients of an image are typically quantized for data reduction. In order to take advantage of local pixel correllation with global transforms, such as the DCT, the image is broken into blocks that are processed separately. Errors are distributed throughout the entire block, making the block boundaries visible. Since the transform presented in this work has local support in the spatial domain, no such blocking is necessary. Fig. 3 shows the reconstruction of a facimile test image after the transform coefficients had been quantized uniformly to 5 b. The images on the left result from the use of the derivative of Gaussian transform, and those on the right from an 8×8 block-based DCT. In both cases, the SNR of the reconstruction is approximately 19 dB.

The blocking artifacts in the image coded with the DCT are offensive in the smooth regions and can make the text unreadable. The DGT coder distributes the same error more uniformly resulting in a more readable image. The increased sharpness at high contrast edges comes at the expense of smoothness in the flat regions. Current work includes implementation of this transform with larger Gaussian variances at low-order derivatives to allow representation of smooth regions and smaller variances at high-order derivatives to maintain sharpness at high contrast edges.

VI. Conclusion

We have introduced a nonorthogonal, 2-D local transform and shown that the decomposition and reconstruction can be calculated efficiently. The preliminary set of basis functions have been shown to yield promising results for image compression. These results are expected to improve further with a more sophisticated basis selection and with a quantization scheme that uses the basis function's general form to exploit properties of the HVS.

References

[1] J. M. H. du Buf, "Gabor phase in texture discrimination," *Signal Processing,* vol. 21, pp. 221–240, 1990.

[2] T. Ebrahimi, T. R. Reed, and M. Kunt, "Video coding using a pyramidal gabor expansion," in *SPIE Proc. VCIP-90,* vol. 1360, 1990, pp. 489–502.

[3] A. E. Soohoo and T. R. Reed, "Low-bit-rate coding of sequences using the Gabor transform," *Soc. Inform. Display Dig.,* vol. XXV, pp. 641–644, 1994.

[4] R. A. Young, "The Gaussian derivative model for spatial vision: I. Retinal mechanisms," *Spatial Vision,* vol. 2, pp. 273–293, 1987.

[5] ——, "Oh say, can you see? The physiology of vision," *Proc. SPIE,* vol. 1453, pp. 92–123, 1991.

[6] J. Martens, "The Hermite transform–theory," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. 38, pp. 1595–1606, 1990.

[7] J. G. Daugman, "Complete discrete 2-D Gabor transform by neural networks for image analysis and compression,"*IEEE Trans. Acoust., Speech, Signal Processing,* vol. 36, pp. 1169–1179, 1988.

[8] M. J. Bastiaans, "Gabor's expansion of a signal into Gaussian elementary signals," *Proc. IEEE,* vol. 68, pp. 538–539, 1980.