A VARIANCE-BASED BASIS SELECTION SCHEME FOR THE GABOR IMAGE TRANSFORMATION

Patrick Lau and Nikolaos P. Papanikolopoulos

Department of Computer Science
University of Minnesota, Minneapolis, MN 55455

ABSTRACT

This paper addresses the issues of time and compression efficiency in image transformation. We propose a novel basis selection scheme which improves transformation efficiency by exploiting the energy compacting characteristic in the frequency domain. In a typical complete transformation, a large number of basis functions have low coding efficiency, and thus are not necessary in the encoding process. By removing these functions from the basis set, we improve the time and the compression efficiency of the encoding process while maintaining a high reproduction quality. We have chosen to use the Gabor Transform to demonstrate our proposed method. Experimental results with the Gabor Transform are presented to demonstrate the effectiveness of our method. Finally, issues related to the application of this approach in image sequence encoding and the adaptation of our approach to other transformation schemes are discussed.

1. INTRODUCTION

Transform coding is one of the most widely used techniques in image coding applications. In image transformation, highly correlated image elements are converted into a set of more independent coefficients. These coefficients can be quantized and coded at a much lower bit-rate than the original pixels. Further compression can be achieved by eliminating some of the coefficients that have small values. The fact that these coefficients can be coded at a low bit-rate is due to the energy compacting characteristic of the transformation [2]. Most of the energy in the image elements is transformed into a small number of coefficients in the frequency domain, which gives the coefficients a low entropy. However, the usefulness of this energy compacting characteristic has only been explored in tasks that are performed after the transformation process (e.g., coefficient selection and quantization). We will examine the energy compacting characteristic of the transformation process using the Gabor Transform in the upcoming sections. The objective of this research is to utilize this energy compacting characteristic to improve the overall time and coding efficiency of the transformation process.

It has been found that a carefully parameterized family of Gabor Elementary Functions (or Gabor functions) can capture the salient tuning properties in spatial-frequency, size, orientation, and phase relationship of the cortical simple cells [7]. These selectivities, especially the frequency and orientation preferences, have a direct effect on the performance of the transformation process. However, most of the studies in Gabor Transform have been directed toward the transformation algorithm, and little attention has been paid to the selectivities of the Gabor functions. Similarly, little attention has been paid to the frequency and orientation attributes of the image features in the studies of transformation algorithms.

In this paper, we propose a novel approach to utilize both the frequency and orientation information in the input images. We maximize the performance of the Gabor Transform by selecting the basis functions according to the frequency and orientation attributes of the image features. In particular, we construct an incomplete basis set that can encode an image in less time, produce a higher compression ratio, and yet maintain a high reproduction quality in comparison to the complete basis set. Furthermore, our approach is "domain-oriented," which means the basis set can be used to transform other images that have similar statistical structures without reconstructing a new basis set [8]. We call this basis set the Maximum-Variance (MV) basis set.

In addition to the improvement in time and compression efficiency, the MV approach also has an advantage in the compression of images with peculiar statistical structures (e.g., images of computational fluid flow). Most of the transform-based image compression algorithms are designed to handle natural images whose energy is assumed to be concentrated in the low-frequency regions [12, 13]. As a result, these algorithms may not able to capture the information in images that have peculiar statistical structures and high energy distribution in the high-frequency regions. The proposed MV approach takes into consideration the energy distribution of the image structure and has a great potential for compression applications in the area of scientific visualization.

The rest of this paper is organized as follows: In Section 2, we provide a brief review of some related work. In Section 3, we describe the computational theory behind the MV approach. Experimental results in still image encoding are presented in Section 4 to demonstrate the time and compression rate improvement by the MV approach. In Section 5, we discuss the potential applications of the MV approach in compressing scientific data/images as well as image sequences. Finally, we summarize the contributions of this paper in Section 6.

2. RELATED WORK

The Gabor Transform has been widely used in image encoding and image compression applications [3, 4, 15]. Among other transform-based image encoding schemes (e.g., the Discrete Cosine Transform and the Fourier Transform),

0-7803-2431-5/95 $4.00 © 1995 IEEE

the Gabor Transform has attracted attention because the Gabor functions are optimally localized in the joint spatial and spatial-frequency domains [3, 6]. In addition, the Gabor Transform presents several computational challenges since most of its implementations require $O(n^3)$ (n is the number of basis functions in the complete case) operations [3, 4]. The Gabor Transform produces low entropy coefficients whose energy in the frequency domain concentrates in the regions that correspond to the most prominent image features.

Recently, Taubman and Zakhor [14] proposed an orientation adaptive coding scheme. The scheme of Taubman and Zakhor resamples small blocks of the image such that the image features in the resampled domain become horizontal and/or vertical. This resampling reduces the activity of the high frequency bands in a subband coding process. In addition, the artifacts introduced when the image is coded with low bit-rates are less disturbing since they lie parallel to the dominant image features. However, the scheme of Taubman and Zakhor is an adaptive coding scheme, which requires the resampling of every input image despite the invariant statistical structure in most natural images [5].

3. OUTPUT VARIANCE AS A SELECTION CRITERION

In the typical, complete Gabor Transform, there are as many basis functions (Gabor functions) as the number of pixels. The basis functions cover all possible frequency ranges from -0.5 to 0.5 horizontally and vertically. However, close inspection of the resulting coefficients reveals that only a small number of the coefficients have large values, while most of them have very small values (see Figure 7 in [3]). In addition, the Gabor functions that produce coefficients with large values are the ones that have frequency parameters that match the dominant image features. In general, a transformation process can be represented by the following equation:

$$\hat{I}[x,y] = \sum_{i=1}^{n} c_i G_i[x,y], \tag{1}$$

where $\hat{I}$ is the reconstructed image, G_i is the basis function, c_i is the corresponding transformation coefficient, and n is the total number of basis functions. In the case of the two-dimensional Gabor Transform, G_i is a Gabor function defined as:

$$\begin{aligned} G_i[x,y] &= \exp(-\pi(\frac{(x-x_i)^2}{\alpha^2} + \frac{(y-y_i)^2}{\beta^2})) \\ &\quad \exp(-2\pi j(u_i x + v_i y)), \end{aligned} \tag{2}$$

where $j = \sqrt{-1}$, (x_i, y_i) and (α, β) are the spatial center and spatial spread of the function, respectively, and (u_i, v_i) is the frequency center of the function. We can see from Equation (1) that the coefficients with small values contribute very little in the reconstruction of the image, which suggests that these coefficients, or correspondingly, the basis functions that produce these coefficients, are unnecessary in the encoding process. Our attention focuses on the derivation of a selection scheme that measures the coding efficiency of the basis functions. By removing the functions with low coding efficiency from the basis set, we reduce the time and the resources needed for encoding the image with little loss in the reproduction quality. We measure the coding efficiency of a

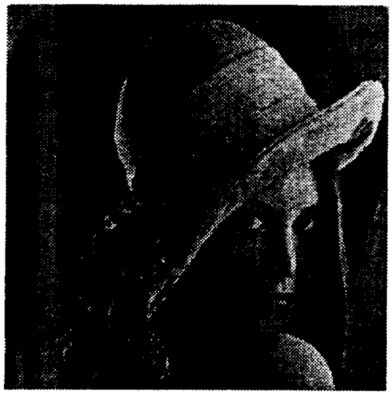

Figure 1: The original "Lenna" image.

Gabor Elementary Function G_i based on the variance of its response over an $M \times N$ image:

$$Var\,\{O_i[x,y], x = 1 \ldots M, y = 1 \ldots N\}. \tag{3}$$

The filter response $O_i[x,y] = I[x,y] \otimes G_i[x,y]$ is the convolution of the Gabor filter with the image. A Gabor function (or basis function) that produces a high output variance transfers more information than one with a low output variance. It has been shown that by maximizing the output variance of a communication channel, we minimize the entropy of the transform coefficients and maximize the performance of the transmission channel [1, 11]. To provide an intuitive explanation, we consider a filter that produces the same (or very similar) output values regardless of the input values. The low output variance of the filter implies that the filter provides no useful information about the input data. Therefore, it is desirable to have basis functions that have high output variances with respect to the input data.

4. EXPERIMENTAL RESULTS

The MV approach was applied in the transformation of a 256×256 "Lenna" image (Figure 1). We have used the Gabor-QR decomposition [9] as the transformation platform. The image was transformed using a latticed approach to avoid dealing with large matrices [10]. We have divided the image into 16×16 lattices, which corresponded to 256 basis functions for a complete basis set. The image was transformed with five different MV basis sets as well as the complete basis set. Figure 2 plots the output variances of the Gabor functions in the frequency domain ("variance map") and the "basis map" for the 131 functions. The variance map and the basis map display the attributes of the Gabor functions in the frequency range of -0.5 to 0.5 in the u and v directions. In the variance map, we use brightness to indicate the value of the output variance. In the basis map, we use the value 1 (bright point) to indicate that the Gabor function with the corresponding frequencies is a member of the basis set, and 0 (dark point) otherwise. The location of the point indicates the frequencies of the function in the u and v directions. Note the similarity between the profiles of the variance map, the basis map, and the power spectrum of the image (right most image in Figure 2), which enforces our hypothesis that the MV basis set captures the dominant features in the image (more examples can be found in [8]).

Table 1 shows the computation time, estimated com-

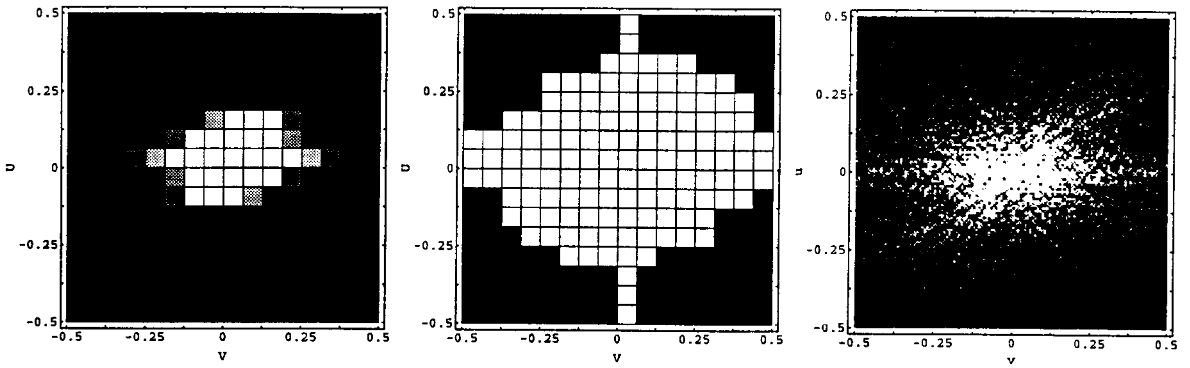

Figure 2: The plots of the output variances (left), the "locations" of the 131 basis functions chosen using the MV approach (center), and the power spectrum of the image (right).

Number of basis functions	Computation time (sec)	Compression rate (bpp)	NMSE
256 (complete)	387.00	1.09	0.019
224	308.61	1.03	0.020
209	277.17	1.00	0.020
183	229.54	0.93	0.022
163	192.72	0.91	0.024
131	138.95	0.82	0.029

Table 1: Experimental results (obtained on a Sun SPARC IPX workstation).

pression rates, and the Normalized Mean Squared Errors (NMSE) of the reconstructed images. In estimating the compression rates, the AC coefficients were uniformly quantized to 65 levels while the DC coefficients were not quantized. Since fewer functions were used in the MV basis sets, both the compression ratio and the transformation time were improved as compared to the complete transformation. Nevertheless, the encoded images retained high reproduction quality. With 131 basis functions, the MV approach required a total of 139 seconds, in contrast to 387 seconds in the complete case, which corresponded to an improvement of 64%. The image encoded with 131 out of 256 basis functions is shown in Figure 3.

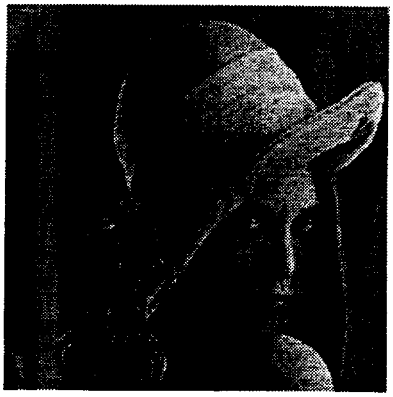

Figure 3: Image encoded with an MV basis set at 0.82 bits per pixel. Only 131 out of 256 basis functions were used in this transformation. Computation time was reduced from 387 seconds to 139 seconds.

5. FUTURE WORK

In the following sections, we look into two different applications where the proposed MV approach may have an impact on the performance of the image encoding process.

5.1. Scientific Image Encoding

In addition to the improvement in time and compression efficiency, the MV approach also has the advantage of being able to capture the statistical structure of the input images. In most transform-based image encoding schemes, a large part of the compression comes from the pruning or coarse quantization of coefficients in the high-frequency regions. These methods assume the input images have few high-frequency details. However, some classes of images (e.g., images of computational fluid flow used in scientific visualization) can have dominant high-frequency image features and a complex statistical structure that does not fall into the category of natural images. An example of such an image and its variance map are shown in Figure 4. Using the MV approach, we will be able to select the basis functions that capture the dominant image features in both the high-frequency and low-frequency areas.

5.2. Image Sequence Encoding

The MV approach can be applied in conjunction with the differential coding technique for image sequence encoding [8]. In differential coding, instead of encoding each frame of the image sequence separately, we encode the difference of two successive frames. A "difference image" usually contains little detail and has features with limited frequency ranges; therefore, very few basis functions are needed to encode this information, especially if the basis functions are well matched to the dominant image features (e.g., the MV basis functions). Preliminary studies [8] have shown that the difference images can be encoded with as few basis functions as 10% of the complete set and the image sequence is still reconstructed with high accuracy. Figure 5 shows some experimental results in reconstructing the second frame of a two-frame sequence using the first frame and the encoded difference image.

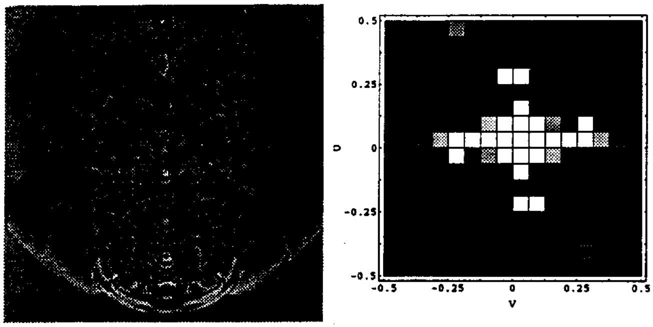

Figure 4: A computational fluid flow image (left) and its variance map (right).

6. CONCLUSION

In this paper, we proposed a variance-based basis selection scheme for image transformation. The proposed MV approach improves the compression and computation efficiency of the transform process by exploiting the frequency and orientation selectivities of the basis functions. The proposed scheme was applied to the selection of basis functions for the Gabor Transform. Experimental results showed that the MV approach improved the computation time and compression rate of the Gabor Transform by 64% and 25%, respectively, and yet maintained a high reproduction quality. Finally, we discussed the potential application of the MV approach to encoding images for scientific visualization and image sequences.

Since the MV basis selection scheme explores the spatial-frequency selectivity of the basis function, other transformation schemes that transform image data from the spatial domain into those of the frequency domain (e.g., the Discrete Cosine Transform) are potential applications that can benefit from the MV approach.

7. REFERENCES

[1] H. Barlow and P. Földiák. *Adaptation and decorrelation in the cortex*, chapter 4, pages 54–72. Addison-Wesley, 1989.

[2] R. Clarke. *Transform coding of images*. Academic Press Inc., 1985.

[3] J. G. Daugman. Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 36(7):1169–1179, 1988.

[4] T. Ebrahimi and M. Kunt. Image compression by Gabor expansion. *Optical Engineering*, 30(7):873–880, 1991.

[5] D. Field. Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America*, 4(12):2379–2394, 1987.

[6] D. Gabor. Theory of communication. *J. Inst. Elec. Eng.*, 93:429–457, 1946.

[7] J. P. Jones and L. A. Palmer. An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *Neurophysiology*, 58:1233–1258, 1987.

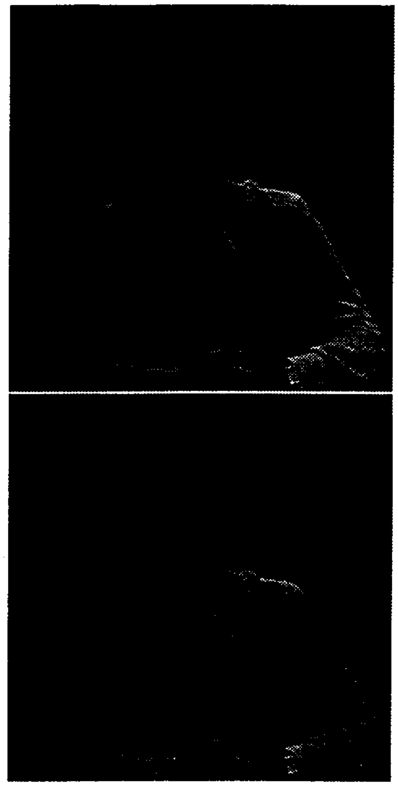

Figure 5: Reconstructing the second frame with the differential coding method. Percentages of basis set used are (approximately) 30% (top) and 10% (bottom). The NMSE's are 0.006669 and 0.011479, respectively.

[8] P. Lau. A maximum-variance approach to image transformation. Technical Report TR 94-04, Department of Computer Science, University of Minnesota, Minneapolis MN 55455, 1994.

[9] P. Lau, N. P. Papanikolopoulos, and D. L. Boley. Gabor-QR decomposition for image encoding. *Electronics Letters*, 29(25):2182–2183, Dec. 1993.

[10] P. Lau, N. P. Papanikolopoulos, and D. L. Boley. A note on the Gabor-QR decomposition. *Proceedings of the IEEE International Conference on Image Processing*, 1:815–819, Nov. 1994.

[11] R. Linsker. Self-organization in a perceptual network. *IEEE Computer*, pages 105–117, Mar. 1988.

[12] H. G. Musmann, P. Pirsch, and H.-J. Grallert. Advances in picture coding. *Proceedings of the IEEE*, 73(4):523–548, 1985.

[13] A. N. Netravali and J. O. Limb. Picture coding: A review. *Proceedings of the IEEE*, 68(3):366–406, Mar. 1980.

[14] D. Taubman and A. Zakhor. Orientation adaptive subband coding of images. *IEEE Transactions on Image Processing*, 3(4):421–437, July 1994.

[15] J. Yao. Complete Gabor transformation for signal representation. *IEEE Transactions on Image Processing*, 2(2):152–159, Apr. 1993.