IMAGE COMPRESSION BY GABOR EXPANSION

Ľubomír ANDRÁŠ, Ján CHMÚRNY
Department of Theoretical Electrotechnics
Military Academy, PO Box 45
031 01 Liptovský Mikuláš
Slovak Republic

Abstract

Transform-based coding methods are popular in data compression. In the paper, an easily implemented method is proposed for the weighting factors of Gabor decomposition. The method is based on the least-mean- squares error (LMSE) approach. The solution of the LMSE problem shows that the weighting factors can be extracted by simple multiplication between a matrix and the vector of data. If the set Gabor functions are chosen to be independent of the test images, this matrix is constant. Images are reconstructed by multiplying the matrix of Gabor functions and the vector of weighting factors. The choise of Gabor functions in the decomposition allows that the resulting decomposition has a pyramidal structure. In the paper is proposed simple codec system for pyramidal Gabor expansion for image compression.

Keywords

Gabor transform, Gabor elementary function, pyramidal Gabor expansion

1. Introduction

In 1946, D. Gabor [1] suggested the expansion (Gabor expansion) of a one-dimensional signal into a discrete set of Gabor elementary functions (GEF) as follows

$$f(t) = \sum_m \sum_n a_{mn} g_{mn}(t) \tag{1}$$

where m, n are integers, a_{mn} are Gabor coefficients. Thus, a given signal can be expressed by these GEF $g_{mn}(t)$ (Fig. 1) using a set of signal-specific weighting coefficients a_{mn} describing the relative weight of each GEF [2].

The GEF are realisation with window function of translation in time and frequency [4]

$$g_{mn}(t) = g(t - mD)\exp(jnWt) \tag{2}$$

where $g(t) = \exp[-\pi (t/D)^2]$ is Gaussian function, D determines the scale of Gaussian in the time domain, $W=2\pi/D$ is the scale in the frequency domain.

Fig. 1 Real part of five GEF

Window suggested by Gabor minimizes the lowest bound on the joint entropy, defined as the product of effective time and frequency bandwidth. The product Δt and $\Delta\omega$ in the combined time-frequency space achieves the smallest value for Gaussian window

$$\Delta t \Delta\omega \geq 2\pi \tag{3}$$

Each Gabor function in the frequency domain corresponds to a Gaussian window in this domain. The real five typical Gabor elementary functions are shown in Fig. 1.

2. Pyramidal Gabor functions

The power in natural images is spread more or less uniformly within octave frequency bands and octave band division is performed in the human visual system [7]. The above considerations the use of a pyramidal Gabor expansion, with elementary functions expressed in the time domain as:

$$h_{m,s}(t) = \begin{cases} g(t - mD) \quad for \quad s = 0 \\ g\left[2^{|s|-1}\left(t - 2^{-|s|+1} mD\right)\right] \\ \exp\left[jn(|s|-1)Wt\right] \quad for \quad s \neq 0 \end{cases} \tag{4}$$

where

$$n(s) = \frac{3 * 2^{|s|-1} - 1}{2} \operatorname{sgn}(s) \tag{5}$$

for $s = -N, \ldots, -1, 0, 1, \ldots, N$.

In these relations, s represents scaling index, D determines the scale of the Gaussian function in the time domain, $W = 2\pi / D$ is the scale in the frequency domain, *m* and *n* are the spatial and frequency location indices [9].

These functions are concentrated in octave bands in the frequency domain and are optimally located in both domains because each one represents a Gabor function [8]. The 2-D Gabor functions have been widely used in computer vision [7]. These functions have a high degree of location in the spatial and spatial-frequency domain, reaching the minimum of the uncertainty product in the two domains [10]. Therefore, regions of an image can be analysed in the joint spatial and frequency domain with minimal interference from adjacent regions [3].

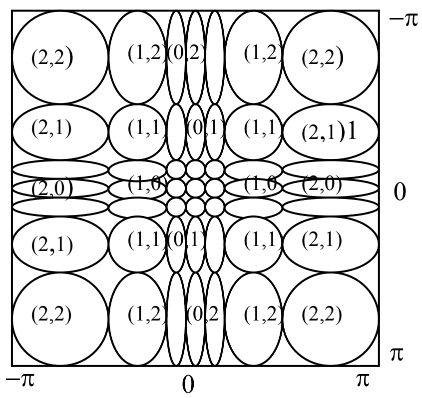

Fig. 2 Spatial frequency partitioning for 2-D pyramidal GEF

A 2-D separable Gabor function can be expressed in the spatial domain as:

$$g_{m1,m2,n1,n2}(k_1,k_2) = g_{m1,n1}(k_1) g_{m2,n2}(k_2) \quad (6)$$

where

$$g_{m,n}(k) = \exp\left[-\pi\left(\frac{k-mD}{D}\right)^2\right]\exp(jnWk) \quad (7)$$

For the same reason as in the 1-D case, a 2-D separable pyramidal Gabor functions can be defined as:

$$h_{m1,m2,s1,s2}(k_1,k_2) = h_{m1,s1}(k_1) h_{m2,s2}(k_2) \quad (8)$$

where (9)

$$h_{m,s}(k) = \begin{cases} \exp\left[-\pi\left(\frac{k-mD}{D}\right)^2\right] & for\ s=0 \\ \exp\left[-\pi\left(\frac{k-2^{-|s|+1}mD}{2^{-|s|+1}D}\right)^2\right].\exp\left[jn(|s|-1)Wk\right] & for\ s\neq 0 \end{cases}$$

Fig. 2 shows how these functions are positioned in the spatial-frequency domain. Each ellipse represents the extent of the Gaussian envelope of functions. Note that in the pyramidal Gabor expansion presented here, the spatial-frequency resolution is higher in the horizontal and vertical directions than in oblique directions and this property is in accordance with the human visual system [7].

3. Discrete Gabor expansion

Consider the sampled 1-D signal *f*(*k*) and its discrete expansion into a weighted sum of elementary functions $g_l(k)$:

$$f(k) = \sum_{l=0}^{N-1} g_l(k) a(l) \quad \text{for} \quad k = 0,\dots M-1, N \le M \quad (10)$$

In this relation, *l* represents the ordering subscript of elementary functions in the expansion *l*=(*m*, *n*). The above expansion can be expressed in matrix notation as:

$$f = G.a, \quad (11)$$

where

$$f = \begin{bmatrix} f(0) \\ . \\ f(M-1) \end{bmatrix}; \quad a = \begin{bmatrix} a(0) \\ . \\ a(N-1) \end{bmatrix}$$

$$G = \begin{bmatrix} g_0(0) & . & g_{N-1}(0) \\ . & . & . \\ g_0(M-1) & . & g_{N-1}(M-1) \end{bmatrix} \quad (12)$$

If $N<M$, (11) cannot be satisfied exactly [9]. Hence, a criterion should be defined to find the best solution of (11). The LMSE criterion is used. Using this criterion, a solution to (11) is the vector $a = \overline{a}$ minimizing $(G.a - f)^T (G.a - f)$, which can be found as:

$$\overline{a} = B.f \quad (13)$$

where $B = (G^T G)^{-1} G^T$ is the pseudo-inverse of *G*. A $N\times N$ matrix $(G^T G)$ cannot always be inverted. To avoid the problem of matrix inversion in such situation, singular value decomposition can be used.

Singular value decomposition states that any $M\times N$ matrix *G* can be decomposed as the product of a $M\times N$ column-orthogonal matrix *U*, a $N\times N$ diagonal matrix *W* with positive or zero elements, and the transpose of $N\times N$ orthogonal matrix *V*:

$$G = U.\begin{bmatrix} w_0 & . & . & 0 \\ . & w_i & . & . \\ . & . & . & . \\ 0 & . & . & w_{N-1} \end{bmatrix}.V^T \quad (14)$$

with $w_i \ge 0$ for all *i*. The solution to (13) is then:

$$\overline{a} = V\left[diag\left(\frac{1}{w_i}\right)\right]U^T f \quad (15)$$

If the matrix *G* is singular or close to singular, the diagonal matrix *W* will contain zero or very small values on its diagonal. One needs simply to replace the associated diagonal elements $1/w_i$, which are very large or undefined, by zero. This automatically eliminates the redundant lines in the G matrix [6].

It should also be noted that if $M=A$ and the elementary functions are independent, the expansion becomes an invertible transform in the sense that the reconstructed data $f = G.a$ will be the exact reconstruction and $A = G^{-1}$.

For 2-D case, the expansion of a 2-D signal (image) into a set of 2-D Gabor functions is defined as [2]:

$$f(k_1, k_2) = \sum_{l_1=0}^{N_1-1} \sum_{l_2=0}^{N_2-1} g_{l1,} g_{l2}(k_1, k_2).a(l_1 l_2)$$

$$for \quad k_1 = 0,...M_1 - 1, \quad N_1 \le M_1$$

$$for \quad k_2 = 0,...M_2 - 1, \quad N_2 \le M_2 \qquad (16)$$

Here, l_1 and l_2 represent the ordering subscripts. The Gabor functions are separable, then can be written as :

$$g_{l1,l2}(k_1, k_2) = g_{l1}(k_1) g_{l2}(k_2) \qquad (17)$$

In this case, eqn. (11) can be written in matrix form as:

$$F = G_1.A.G_2^T \qquad (18)$$

where F is the 2-D data matrix and A is the 2-D expansion coefficient matrix:

$$F = \begin{bmatrix} f(0,0) & ... & f(0, M_2 - 1) \\ ... & ... & ... \\ f(M_1 - 1, 0) & ... & f(M_1 - 1, M_2 - 1) \end{bmatrix}$$

$$A = \begin{bmatrix} a(0,0) & ... & a(0, N_2 - 1) \\ ... & ... & ... \\ a(N_1 - 1, 0) & ... & a(N_1 - 1, N_2 - 1) \end{bmatrix} \qquad (19)$$

and G is the same expansion matrix as in eqn. (12).

Again, for $N_i < M_i$ the exact relation cannot be always satisfied. Here also the two-step LMSE criterion can be used to define the best solution (according to this criterion), given by:

$$\bar{A} = \left(G_1^T G_1\right)^{-1} G_1^T.F.G_2 \left(G_2^T G_2\right)^{-1} \qquad (20)$$

The use of singular value decomposition is also valid here and the solution is given by:

$$\bar{A} = V_1 [diag\ (1/w_{i1})] U_1^T.F.U_2 [diag\ (1/w_{i2})] V_2^T \qquad (21)$$

The solution of the LMSE problem shows that weighting factors can be extracted by simple multiplication between a matrix and the vector of data.

4. The proposal simple codec system

Recursive pyramidal expansion is realised as: at first the signal is decomposed into a set of Gabor functions and then the same operation can be applied recursively to the result of the previous operation. The recursive expansion is comparable to recursive sub-band decomposition.

The information about different frequency bands and orientations can be extracted using these expansions by one simple matrix multiplication applied to the entire image.

Fig. 3 illustrates a simple codec system, which is used to demonstrate the performance of the recursive pyramidal Gabor expansion for image compression.

The input image $F(x,y)$ is first decomposed by matrix multiplication $B^T.F.B$ to obtain the coefficients of expansion. The resulting coefficients are then uniformly quantized and thresholded to retain only the most significant coefficients. The uniform quantization followed by an entropy coder is the best solution in the LMSE distortion sense for a given bit rate. The threshold can be considered the dead zone of the above uniform quantizer. The dead zone diminishes the bit rate without increasing the subjective quantization noise.

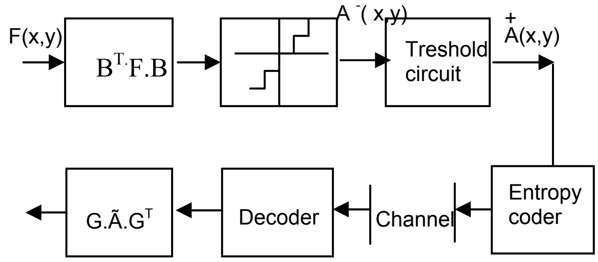

Fig. 3 The simple codec system used for 2-D compression

5. Conclusion

Gabor transform is a useful tool in signal processing. Gabor elementary functions are chosen so that the resulting decomposition has a pyramidal structure. The use of the pyramidal structure is motivated both by the observations of the early human visual system and by the fact that it lends itself easily to progressive transmission. A technique for coding of still image based on a pyramidal Gabor decomposition is discussed. The GEF of the decomposition form a class of basic functions that covers the frequency domain in octave bands.

The choice of GEF in the Gabor decomposition is very important. If the set of GEF is chosen to be independent of the image, this matrix is constant [5]. GEF are not suitable for numerical computations and very expensive to compute. GEF are not orthogonal and the computation of Gabor transform is complicated.

Image sequence compression is performed by decomposition of each individual image and by transmission the significant differences between the coefficients of successive frames. The coefficients of expansion are computed frame by frame and the coefficients of each frame are uniform quantized, similar to the still image coding method. Because the Gabor transform is applied to the entire image and GEF have overlapping support, the blocking defects in the highly compressed images are reduced.

Pyramidal Gabor decomposition is very efficient in image analysis and coding.

References

[1] GABOR, D.: Theory of Communication. Journal IEE. London, vol. 93, November 1946, pp. 429-457.

[2] DAUGMAN, J.G.: Complete Discrete 2-D Gabor Transform by Neural Networks for Image Analysis and Compression. IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 36, no. 7, July 1988, pp. 1169-1179.

[3] WANG, H., YAN, H.: Efficient Implementation of Gabor Transform for Image Compression. Electronics Letters, vol. 28, no. 9, April 1992, pp. 870-871.

[4] ANDRÁŠ, L., CHMÚRNY, J.: Gabor Transform and its Applications. Journal of Electrical Engineering, vol. 48, No. 3-4, 1997, pp. 101-104.

[5] BASTIAANS, M. J.: A Sampling Theorem for the Complex Spectrogram, and Gabor's Expansion of a Signal in Gaussian Elementary Signals. Optical Engineering, vol. 20, no. 4, 1981, pp. 594-598.

[6] STEWART, D.F.: Minimal Time-Frequency Localization Techniques and their Application to Image Compression. Dissertation Thesis, The Ohio State University, 1994.

[7] PORAT, M., ZEEVI, Y.Y.: The Generalized Gabor Scheme of Image Representation in Biological and Machine Vision. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 10, no. 4, 1988, pp. 453-468.

[8] DUNN, D., HIGINNS, W.E., WAKELEY, J.: Texture Segmentation Using 2-D Gabor Elementary Functions. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 16, no. 2, 1994, pp. 130-149.

[9] WEXLER, J., RAZ, S.: Discrete Gabor Expansions. Signal Processing, vol. 21, no. 3, November 1990, pp. 207-221.

[10] ANDRÁŠ, Ľ., ŠURIANSKY, J.: The Use of Gabor Transform for Image Processing. In Proceedings of the National Scientific Conference ELEKTRO 95, Žilina, 1995 pp. 159-162.

About authors...

Ľubomír ANDRÁŠ (Ing, PhD) was born in Vojčice (Slovakia). He received his Ing (MSc) degree from the Military Academy in Liptovský Mikuláš in 1986 and PhD degree in 1998. Since 1988 he is with the department of Theoretical Electrotechnics, Faculty of Support of Command, Military Academy in Liptovský Mikuláš. His current research interest are connected to Gabor transform, short-time Fourier transform and its application in signal processing.

Ján CHMÚRNY (Dr.h.c., Prof., Ing., DrSc) was born in Martin (Slovakia). He received his undergraduate Diploma degree in electrical engineering from the Slovak Technical University of Bratislava in 1947, the CSc (PhD) degree in Radioelectronics from the Moscow Technical University in 1954, and the DrSc degree in Radioelectronics from the Technical University of Brno (Czech republic) in 1984. He is curently Professor of Electrical Engineering in the Military University of Liptovský Mikuláš. His research interest are in the areas of image processing, coding, neural networks and mobile communication system.

RADIOENGINEERING REVIEWERS

July 2001, Volume 10, Number 2