Pattern Recognition Letters 13 (1992) 701-706 October 1992

On a decomposition of 2-D circular convolution

G. Lohar, D.P. Mukherjee and D. Dutta Majumder

National Nodal Centre for Knowledge Based Computing, Electronics and Communication Sciences Unit, Indian Statistical Institute, 203 B.T. Road, Calcutta 700 035, India

Received 8 January 1992
Revised 10 March 1992

Abstract

Lohar, G., D.P. Mukherjee and D. Dutta Majumder, On a decomposition of 2-D circular convolution, Pattern Recognition Letters 13 (1992) 701-706.

A square matrix can be expressed in terms of a set of circulant matrices. This representation is applied to the 2-D circular convolution of two square matrices. The circulant representation of the resulting matrix has the property that each circulant is the product of the corresponding circulants of the two matrices, the product being a 1-D circular convolution. This decomposition offers insight into the process of 2-D convolution. This has applications in image processing.

Keywords. Circulant matrix, 2-D circular convolution.

1. Introduction

The circulant decomposition of a square matrix and its application to image representation has been reported earlier in Wacker and Lohar (1986). However, no efficient algorithm for its calculation or subsequent theoretical development based on this decomposition has been reported. An interesting result which finds application in a decomposition of 2-D circular convolution is the subject of this paper.

The circulant decomposition of a square matrix requires the application of a 1-D Fourier transform only once to the diagonally scanned elements of the matrix. In contrast, the 2-D Fourier transform is calculated by applying a 1-D Fourier transform to the columns, followed by a 1-D Fourier transform on the resulting rows. It is thus easier to physically interpret the circulant decomposition. The significance of the Fourier transform in the context of image processing lies in the fact that convolution in the spatial domain transforms to multiplication in the frequency domain. It is harder to interpret this process in two dimensions than in one dimension both from an analysis and synthesis point of view. This problem is addressed in this paper.

An efficient algorithm for the calculation of the circulant representation is presented in Section 2. This is followed by a discussion on its relationship to the Fourier transform in Section 3. Section 4 presents the main result of this paper. Examples illustrating the advantages of the circulant represen-

Correspondence to: G. Lohar, Indian Statistical Institute, ECSU, 203 B.T. Road, Calcutta 700 035, India.

This work is partially supported by DoE/UNDP fund IND/85/072.

tation are presented in Section 5 followed by conclusions in Section 6.

2. Circulant decomposition and its properties

Consider a $p \times p$ matrix $[H]$. [For ease of notation, matrix rows and columns are numbered from 0 to $p-1$.] A $p \times p$ circulant matrix is defined by the elements of its 0th column. The ith column is generated by i cyclic shifts of the 0th column, where the direction of shift is along increasing row number. Let $[T]$ represent the $p \times p$ Discrete Fourier Transform (DFT) matrix whose klth element is defined by ω^{kl} where $\omega = \exp(-j2\pi/p)$, $j^2 = -1$. Let T_i represent the ith column of $[T]$. The $p \times p$ matrix $[H]$ can be decomposed into circulant matrices as follows

$$[H] = \sum_{i=0}^{p-1} \operatorname{diag}\{T_i^*\}[H_c(i)]. \tag{1}$$

$[H_c(i)]$ is a circulant matrix, $\operatorname{diag}\{T_i^*\}$ is a diagonal matrix whose diagonal elements are the elements of T_i^* ($*$ denotes complex conjugate).

For a given $[H]$, an algorithm to determine the circulant matrices is given below. It is noted that it is sufficient to determine the 0th columns of each circulant matrix. For ease of understanding, a proof is presented for the case $p=3$. The procedure is easily generalised for arbitrary p. Consider the case $p=3$. Let

$$[H] = \begin{bmatrix} h_{00} & h_{01} & h_{02} \\ h_{10} & h_{11} & h_{12} \\ h_{20} & h_{21} & h_{22} \end{bmatrix},$$

$$[H_c(i)] = \begin{bmatrix} h_0(i) & h_2(i) & h_1(i) \\ h_1(i) & h_0(i) & h_2(i) \\ h_2(i) & h_1(i) & h_0(i) \end{bmatrix}.$$

From equation (1), matching element by element, it follows that

$$\begin{aligned} h_{00} + h_{11} + h_{22} &= 3h_0(0) + h_0(1)(1 + \omega^* + \omega^{*2}) \\ &\quad + h_0(2)(1 + \omega^{*2} + \omega^{*4}) \\ &= 3h_0(0). \end{aligned}$$

Similarly,

$$h_{10} + h_{21} + h_{02} = 3h_1(0),$$

$$h_{20} + h_{01} + h_{12} = 3h_2(0).$$

Hence,

$$\frac{1}{3}[1 \quad 1 \quad 1] \begin{bmatrix} h_{00} & h_{02} & h_{01} \\ h_{11} & h_{10} & h_{12} \\ h_{22} & h_{21} & h_{20} \end{bmatrix}$$

$$= [h_0(0) \quad h_1(0) \quad h_2(0)]$$

$$= [\text{0th column of } [H_c(0)]]^t$$

(t denotes transpose).

$[H_c(0)]$ is thus determined since it is specified by its 0th column.

To determine $[H_c(1)]$, multiply both sides of equation (1) by $\operatorname{diag}\{T_1\}$. It is now noted that $[H_c(1)]$ is the 0th circulant matrix obtained by decomposing $\operatorname{diag}\{T_1\}[H]$ since $\operatorname{diag}\{T_1\}\operatorname{diag}\{T_1^*\}$ is the identity matrix. Using the above procedure, it is readily determined that

$$\frac{1}{3}[1 \quad \omega \quad \omega^2] \begin{bmatrix} h_{00} & h_{02} & h_{01} \\ h_{11} & h_{10} & h_{12} \\ h_{22} & h_{21} & h_{20} \end{bmatrix}$$

$$= [h_0(1) \quad h_1(1) \quad h_2(1)].$$

$[H_c(1)]$ is thus determined.

To determine $[H_c(2)]$, multiply both sides of equation (1) by $\operatorname{diag}\{T_2\}$ and use the above procedure. It follows that

$$\frac{1}{3}[1 \quad \omega^2 \quad \omega^4] \begin{bmatrix} h_{00} & h_{02} & h_{01} \\ h_{11} & h_{10} & h_{12} \\ h_{22} & h_{21} & h_{20} \end{bmatrix}$$

$$= [h_0(2) \quad h_1(2) \quad h_2(2)].$$

$[H_c(2)]$ is thus determined.

Collecting the above results together,

$$\frac{1}{3}[T] \begin{bmatrix} h_{00} & h_{02} & h_{01} \\ h_{11} & h_{10} & h_{12} \\ h_{22} & h_{21} & h_{20} \end{bmatrix}$$

$$= \begin{bmatrix} h_0(0) & h_1(0) & h_2(0) \\ h_0(1) & h_1(1) & h_2(1) \\ h_0(2) & h_1(2) & h_2(2) \end{bmatrix}.$$

The elements of the ith row on the right are the elements of the 0th column of the ith circulant matrix.

This procedure is easily generalized for arbitrary p. It is noted that the procedure essentially involves diagonally scanning the elements of $[H]$ in a particular order, inserting the diagonally scanned elements into columns, scaling the elements and then applying DFT to the columns of the rearranged matrix.

For a $p \times p$ matrix $[H]$, let $[H_s]$ represent the result of diagonally scanning and rearranging $[H]$. Let $[H_{rc}]$ represent the matrix whose rows contain elements of the 0th columns of the circulant matrices corresponding to $[H]$. Then,

$$\frac{1}{p}[T][H_s] = [H_{rc}] \;\Rightarrow\; [H_s] = p[T]^{-1}[H_{rc}].$$

If $[H]$ is real, then from a property of the DFT applied to real sequences it readily follows that for $i = 1, 2, \ldots, \frac{1}{2}p - 1$ (p even) or $i = 1, 2, \ldots, \frac{1}{2}(p-1)$ (p odd), rows i and $p - i$ of $[H_{rc}]$ are complex conjugates of each other. Rows 0 and $\frac{1}{2}p$ (p even) are real.

3. Relation of circulants to Fourier coefficients

The Fourier transform of $[H]$ is defined as $[T][H][T] = [H_f]$ (Pratt (1978)). Each circulant matrix $[H_c(i)]$ is composed of a set of p Fourier coefficients of $[H]$ in a select order, as will be shown below. The p Fourier coefficients are from both low and high frequency regions. The circulant matrices are indexed by a single number i. In other words, a single index i is used to label p Fourier coefficients. Hence, two-dimensional regions in the Fourier domain are represented by p-dimensional column vectors, the 0th columns of the circulant matrices.

Let $[I] = [I_0 \; I_1 \; \cdots \; I_{p-1}]$ denote the $p \times p$ identity matrix where I_i is the ith column of $[I]$. Consider the 2-D circular convolution of $[I]$ and $\mathrm{diag}\{T_i^*\}[H_c(i)]$. It is now shown that for $i = 0$, the result is $p[H_c(0)]$ and for $i \neq 0$, the result is a zero matrix. The Fourier transform of $[I]$ is $[T][I][T] = [T]^2$. As can be checked,

$$[T]^2 = p[I_0 \; I_{p-1} \; I_{p-2} \; \cdots \; I_1].$$

The Fourier transform of $\mathrm{diag}\{T_i^*\}[H_c(i)]$ is

$$[T]\mathrm{diag}\{T_i^*\}[H_c(i)][T].$$

This can be expressed as

$$[T]\mathrm{diag}\{T_i^*\}[T]^{-1}[T][H_c(i)][T]^{-1}[T][T].$$

An important property of DFT and circulant matrices is that a DFT matrix diagonalizes a circulant matrix (Hunt (1971)). Specifically, $[T][H_c(i)][T]^{-1}$ is a diagonal matrix whose jth diagonal element, say $H_j(i)$, is the jth Fourier coefficient of the DFT of the 0th column of $[H_c(i)]$. The term $[T][H_c(i)][T]^{-1}[T]^2$ can be expressed as

$$p[H_0(i)I_0 \;\; H_{p-1}(i)I_{p-1} \;\; H_{p-2}(i)I_{p-2} \;\cdots\; \cdots \; H_1(i)I_1].$$

As can be checked,

$$[T]\mathrm{diag}\{T_i^*\}[T]^{-1} = \begin{bmatrix} I_{p-i}^t \\ I_{p-i+1}^t \\ \vdots \\ I_0^t \\ I_1^t \\ \vdots \\ I_{p-i-1}^t \end{bmatrix}$$

which is obtained by cyclic shifts of the rows (considered as units) of $[I]$ i times, where the direction of shift is along increasing row number (rows cyclicly shifted downwards) (the indices of I are to be interpreted as being periodic with period p). This implies that the Fourier transform of $\mathrm{diag}\{T_i^*\}[H_c(i)]$ has a structure similar to $[T][H_c(i)][T]^{-1}[T]^2$ except for a change only in the indices of the columns of $[I]$. The change is expressed by i cyclic shifts towards the right of only the I_j terms. The convolution of $[I]$ and $\mathrm{diag}\{T_i^*\}[H_c(i)]$ is calculated by multiplying the corresponding Fourier coefficients term by term and using the inverse Fourier transform. It is now clear that for $i \neq 0$, the positions of the non-zero Fourier coefficients of $[I]$ and $\mathrm{diag}\{T_i^*\}[H_c(i)]$ do not match. Hence, multiplication of the Fourier coefficients term by term results in a zero matrix. For $i = 0$ on the other hand, the result of convolving $[I]$ and $\mathrm{diag}\{T_0^*\}[H_c(0)] = [H_c(0)]$ is $p[H_c(0)]$.

In equation (1), consider the product $\mathrm{diag}\{T_j\}$

$[H]$. Using the fact that $\text{diag}\{T_j\}\,\text{diag}\{T_j^*\} = [I]$, it is clear that $[H_c(j)]$ is the 0th circulant matrix obtained by the circulant decomposition of $\text{diag}\{T_j\}[H]$. Therefore, in general, $[H_c(i)]$ is $1/p$ times the result of convolving $[I]$ and $\text{diag}\{T_i\}[H]$.

The specific Fourier coefficients of $[H]$ which compose $[H_c(i)]$ are now considered. Since $[H_c(i)]$ is $1/p$ times the convolution of $[I]$ and $\text{diag}\{T_i\}$ $[H]$, the Fourier coefficients of $\text{diag}\{T_i\}[H]$ which correspond to the positions of the non-zero elements of $[T]^2$, the Fourier transform of $[I]$, compose $[H_c(i)]$. The Fourier transform of $\text{diag}\{T_i\}[H]$ is

$$[T]\,\text{diag}\{T_i\}[H][T] = [T]\,\text{diag}\{T_i\}[T]^{-1}[H_f].$$

As can be checked, $[T]\,\text{diag}\{T_i\}[T]^{-1}$ is obtained by cyclic shifts of the rows (considered as units) of $[I]$ i times, where the direction of shift is along decreasing row number (rows cyclicly shifted upwards). The Fourier coefficients corresponding to the positions of the non-zero elements of $[T]^2$ are thus

$$H_f(i,0), H_f(i-1,1), H_f(i-2,2), \ldots, H_f(i+1,p-1)$$

(the Fourier indices are to be interpreted as being periodic with period p). It is therefore evident that all Fourier coefficients $H_f(j,k)$ such that $j+k=i$ $(\text{mod}\,p)$ go into the composition of $[H_c(i)]$. It is noted that these Fourier coefficients are from the low band and high pass range.

4. Decomposition of 2-D circular convolution

Consider the 2-D circular convolution of the $p\times p$ matrices $[H]$ and $[G]$ to yield the $p\times p$ matrix $[F]$. As in equation (1),

$$[H] = \sum_{i=0}^{p-1} \text{diag}\{T_i^*\}[H_c(i)], \tag{2}$$

$$[G] = \sum_{i=0}^{p-1} \text{diag}\{T_i^*\}[G_c(i)], \tag{3}$$

$$[F] = \sum_{i=0}^{p-1} \text{diag}\{T_i^*\}[F_c(i)]. \tag{4}$$

It is asserted that for $i=0,1,\ldots,p-1$,

$$[F_c(i)] = [G_c(i)][H_c(i)] = [H_c(i)][G_c(i)]. \tag{5}$$

Equation (5) implies that the 0th column of $[F_c(i)]$ is the 1-D circular convolution of the 0th columns of $[G_c(i)]$ and $[H_c(i)]$.

A short heuristic proof based on the results of Section 3 is now presented. In the Fourier domain, every Fourier coefficient of $[F]$ is the product of the corresponding Fourier coefficients of $[G]$ and $[H]$. As shown in the last section, $[F_c(i)]$ is composed of p Fourier coefficients of $[F]$, the Fourier coefficients being the DFT of the 0th column of $[F_c(i)]$. But the DFT coefficients of the 0th column of $[F_c(i)]$ are the product of the corresponding DFT coefficients of the 0th columns of $[G_c(i)]$ and $[H_c(i)]$. Hence, the 0th column of $[F_c(i)]$ is the 1-D circular convolution of the 0th columns of $[G_c(i)]$ and $[H_c(i)]$. This can be expressed as

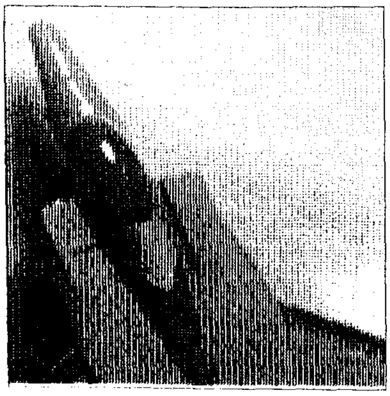

Figure 1.

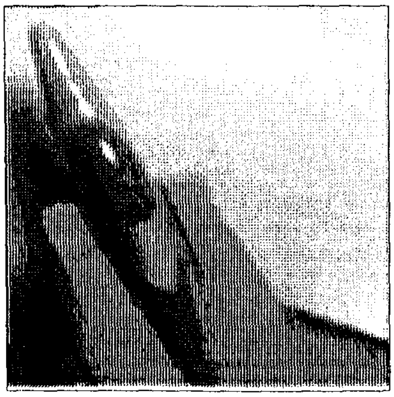

Figure 2.

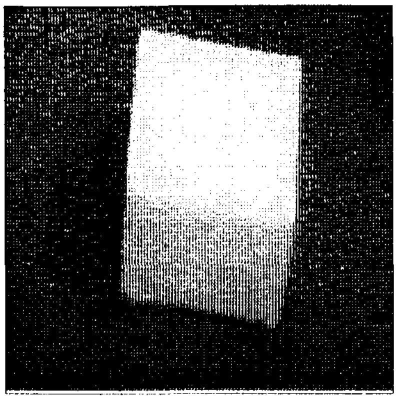

Figure 3.

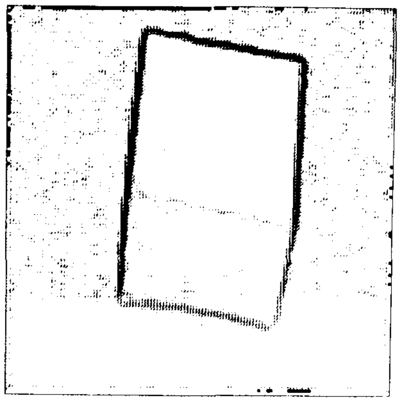

Figure 4.

$$0\text{th column of } [F_c(i)]$$
$$= [G_c(i)] \times 0\text{th column of } [H_c(i)]$$
$$= [H_c(i)] \times 0\text{th column of } [G_c(i)].$$

From this fact, equation (5) follows.

In terms of the matrices $[F_{rc}]$, $[G_{rc}]$ and $[H_{rc}]$, it follows that each row of $[F_{rc}]$ is the 1-D circular convolution of the corresponding rows of $[G_{rc}]$ and $[H_{rc}]$. If the matrices $[G]$ and $[H]$ are real, then due to the complex conjugate property among the circulants, approximately half of the total number of 1-D circular convolutions need not be calculated.

5. Example

An example illustrating the data compression capability of the circulant representation is presented in Figures 1 and 2. Figure 1 represents a 128×128 image and Figure 2 represents the reconstructed image using the first 25 circulants (and their appropriate complex conjugates). To reconstruct the original image exactly requires 65 circulants (and their appropriate complex conjugates). This representation is sensitive to gray level constancy along the diagonals which show up early during reconstruction.

Figure 3 represents a 128×128 image and Figure 4 is the result of applying an edge detection filter to the image of Figure 3. Figure 5 is the result of filtering with the first 31 one-dimensional convolutions (and their appropriate complex conjuagtes). In this context, it should be noted that in the reconstruction of the filtered image, the appropriate complex conjugates of the convolutions need not be calculated.

The circulant representation has a compact notation and is easy to use in a programming environment. This is to be contrasted with the two-dimensional Fourier transform. From the 2-D circular convolution decomposition, it follows that filtering of the original image is equivalent to the corresponding filtering of each circulant. This fact can be used in the analysis of existing filtering techniques or in the design of filters using one-dimensional techniques. Due to the complex conjugate property among the circulants and the data compression capability, the circulant representation has computational advantages.

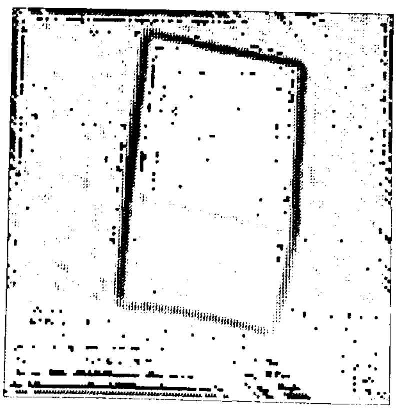

Figure 5.

6. Conclusions

An algorithm for the circulant representation of a square matrix has been presented. This representation when applied to a 2-D circular convolution leads to a decomposition in terms of 1-D circular convolutions. This decomposition offers a new insight in and interpretation of the process of 2-D convolution. It has been noted that this decomposition in the context of image processing applications can be used to analyse or synthesise filtering techniques more simply than conventional two-dimensional approaches. There is also a computational advantage in practical applications. It is hoped that this technique being of general applicability will find use in other contexts.

References

Hunt, B.R. (1971). A matrix theory proof of the discrete convolution theorem. *IEEE Trans. Automat. Control* 19, 285–288.

Pratt, W.K. (1978). *Digital Image Processing*. Wiley, New York.

Wacker, A.G. and G. Lohar (1986). Image representation via its symmetrical components. *Proc. IGARSS '86 Symposium*, Zurich, 8–11 Sept., 761–766.