Computer Optics Vol. 2, No. 1, pp. 23–28, 1990
Printed in Great Britain

0955–355X/90 $10.00 + 0.00
Copyright © 1990 Pergamon Press plc

NUMERICAL SIMULATION OF TWO-DIMENSIONAL LINEAR SYSTEMS

V. V. SERGEYEV and A. V. USACHEV

Abstract—The problem of numerical simulation is considered for continuous linear systems with constant parameters transforming two-dimensional signals. Ways of reducing the computational complexity of the model are demonstrated. They include rational selection of parameters in the discrete Fourier transform used, partitioning the convolution and use of the Fourier transformation in the Hartly form. The modeling algorithm is given for the case when the impulse response function of the system is known and the input signal is space limited.

Numerical simulation of two-dimensional linear systems transforming spatial signals can have many applications, specifically in computer-aided design of optical imaging and data processing systems.

A numerical model of a system must be capable of approximating with high accuracy to the system's continuous output, using results of sampling involved in the model. Also, a high computational efficiency is required. The fidelity of the model may be ensured by a high sampling rate, but this greatly enlarges the arrays of data and increases the complexity of the algorithm. On the other hand, in reducing the cost of computations we would probably obtain coarser results. There are certain means, however, that allow modeling without losses of accuracy.

The modeling procedure outlined below relates to two-dimensional linear systems with constant parameters.

STATEMENT OF THE PROBLEM

The output $g(x, y)$ of a linear system with constant parameters is the aperiodic convolution of the input signal $f(x, y)$ and the system impulse response function (point-spread function) $h(x, y)$

$$g(x, y) = \int\int_{-\infty}^{\infty} f(\xi, \eta) h(x - \xi, y - \eta)\, d\xi\, d\eta, \tag{1}$$

where x and y are spatial arguments of the signal.

The equivalent expression in the spatial frequency domain is

$$G(\omega_x, \omega_y) = F(\omega_x, \omega_y) \cdot H(\omega_x, \omega_y),$$

where G, F and H are the Fourier images of the functions g, f and h respectively, and ω_x and ω_y are the spatial frequencies.

Suppose that the two-dimensional system is given by its frequency characteristic (transfer function) possessing the property of radial symmetry, *viz.*,

$$H(\omega_x, \omega_y) = H_\rho(\rho), \quad \rho = \sqrt{\omega_x^2 + \omega_y^2},$$

and the two-dimensional input signal $f(x, y)$ is nonzero in a rectangular field. This field is represented by a grid of samples obtained by discretization in the coordinates with a fixed increment (spacing) T, namely,

$$\begin{aligned} f_{mn} &= f(mT, nT), && \text{for} \quad (m, n) \in [0, M-1] \times [0, N-1], \\ &= 0, && \text{otherwise}. \end{aligned}$$

Here, M and N are the dimensions of the meaningful part of the array of input samples.

We assume that the spacing T is sufficiently small for the effects of overlap of the spectra in the discretization to be neglected, i.e. assume that the input-signal spectrum is band limited, *viz.*,

$$F(\omega_x, \omega_y) = 0 \qquad \text{for} \quad |\omega_x| \geqslant \pi/T \quad \text{or} \quad |\omega_y| \geqslant \pi/T.$$

The sample spacing T is assumed to satisfy the Nyquist theorem. Under the circumstances the continuous convolution (1) may be replaced by its discrete analog without loss of accuracy, namely [1],

$$g_{k,l} = T^2 \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f_{mn} h_{k-m,l-n}, \tag{2}$$

where $g_{k,l} = g(kT, lT)$ and $h_{p,q} = h(pT, qT)$.

We seek the samples of the output signal in some rectangular field $(k, l) \in [k_1, k_2] \times [l_1, l_2]$, where k_1, k_2, l_1 and l_2 are the boundaries of the output sample field in question.

A fast technique of computing the convolution (2) consists in taking three discrete Fourier transforms (DFT), two direct transforms and one inverse [2].

In the problem in question the DFT treatment acquires a number of specific features. On the one hand, the description of the system directly in the spatial frequency domain allows one to eliminate one DFT procedure. On the other hand, the spread of the sampled impulse response becomes unknown, that is, one is uncertain about the number of appreciably nonzero samples on the positive half-axis.

ESTIMATING THE SPREAD OF THE IMPULSE RESPONSE FUNCTION

To estimate the spread P of the point spread function $h(x, y)$ we suggest a simple technique using the one-dimensional Fourier transform in place of the two-dimensional procedure. It boils down to the evaluation of the spread of the integrand function in one variable

$$h_\Sigma(x) = \int_{-\infty}^{\infty} h(x, y)\,dy.$$

The Fourier transform of the integrand is the function $H_\rho(\rho)$ continued into the domain $\rho < 0$ as an even function. Clearly, the spread of function $h_\Sigma(x)$ is equal to that of the point spread function $h(x, y)$, as shown in Fig. 1. Accordingly, the parameter P can be determined, say, as the least positive number satisfying the inequality

$$1 - \int_0^{PT} |h_\Sigma(x)|^2\,dx \Big/ \int_0^{\infty} |h_\Sigma(x)|^2\,dx < \varepsilon,$$

where ε is a quantity characterizing the error caused by truncation of the point spread function.

EVALUATING THE DFT SPAN

An enhancement of model efficiency may be gained by taking into account the relations of the input and output windows, i.e. the size of the information signal at the input and the constraints

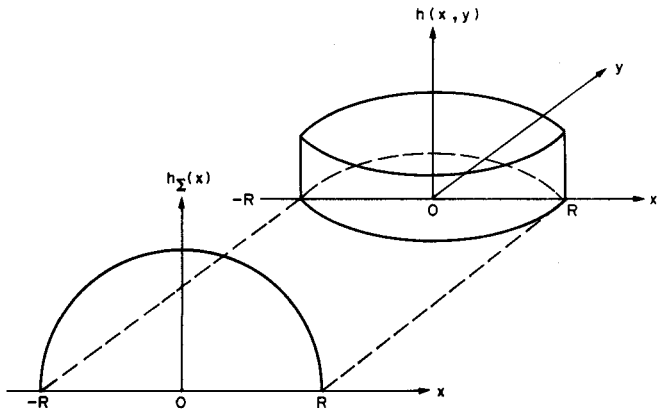

Fig. 1. Construction to evaluate the width of the impulse response function $R = PT$.

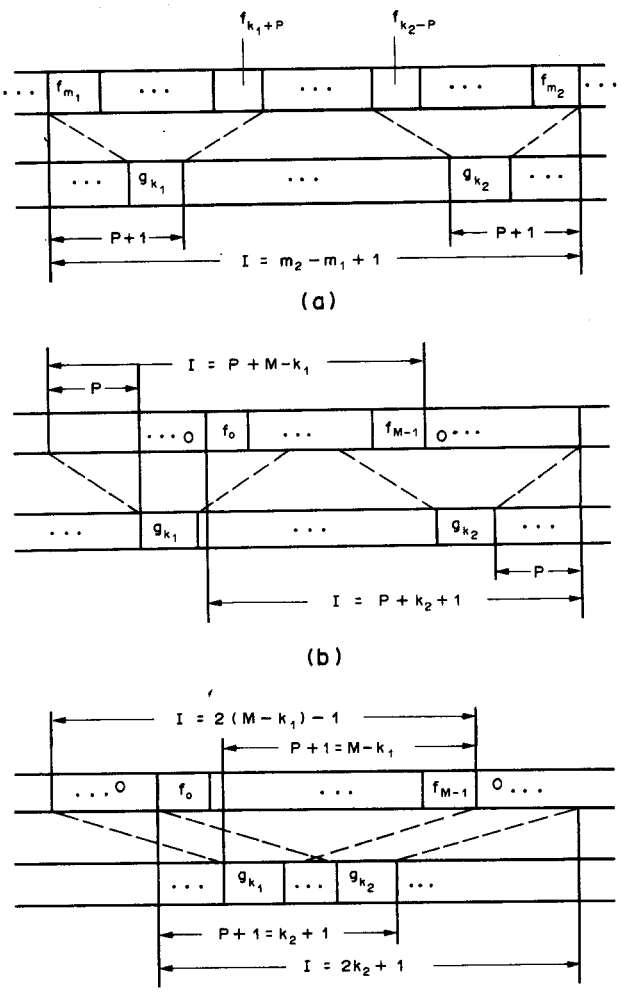

Fig. 2. Evaluation of the least span of DFT: (a) the general case, $I = m_2 - m_1 + 1$; (b) $m_1 < 0$, $m_2 > M - 1$, $P < M$; $I = P + \max\{M - k_1, k_2 + 1\}$; (c) $m_1 < 0$, $m_2 > M - 1$, $P \geqslant M$; $P = \max\{M - k_1 - 1, k_2\}$, $I = 2P + 1$.

on the analysed output frame. This procedure may be adequately described by giving the one-dimensional case:

Compute

$$g_k = \sum_{m=0}^{M-1} f_m h_{k-m}, \quad k \in [k_1, k_2], \tag{3}$$

with $m \notin [0, M-1]$ for f_m, and $h_m = 0$ for $|m| > P$.

All the sequences involved in the DFT convolution are known to be periodic—with a period I, the span (length) of DFT. We need this convolution to coincide with the computation of the aperiodic convolution (3) for the output window despite the possible alien effects that might arise when going from the series of finite length to the periodic series.

The relation of samples in the input and output series of the aperiodic convolution is shown in Fig. 2. From Fig. 2(a) it follows that in the general case we need information about the values f_m in the input window for $m \in [m_1, m_2]$, with $m_1 = k_1 - P$ and $m_2 = k_2 + P$. Consequently, we may drop the values of the input sequence of f_m outside this interval and assume the cyclic convolution span to be

$$I = m_2 - m_1 + 1 = k_2 - k_1 + 2P + 1. \tag{4}$$

This approach precludes alien effects in the output window elements.

An additional contraction of the DFT span can be achieved when the output window or the size of the impulse response function is such that the input window covers the interval of nonzero values of the f_m sequence, i.e. $m_1 < 0$, $m_2 > M - 1$. If the width of the impulse response function is small ($P < M$), then, observing the cyclic property of the convolution, we can allow harmless alienation in the computation by keeping zeros at one side only, precisely where there are more of them (Fig. 2b). In other words, we assume that

$$m_1 = 0, \qquad I = P + \max\{M - k_1, k_2 + 1\}, \qquad m_2 = I - 1. \tag{5}$$

If the width of the impulse response function is so large that $P \geqslant M$, then to compute every desired sample of the output sequence we need all nonzero samples of the input sequence. In this case from Eq. (3) it follows that only a part of nonzero samples of the point spread function h_m (for $m \in [k_1 - M + 1, k_2]$) will suffice for the computation, the other may be equated to zero.

From the implementation standpoint it is convenient to consider a symmetric truncation of the point spread function while recognizing some redundancy of information, that is, to assume (Fig. 2c)

$$P = \max\{M - k_1 - 1, k_2\}, \qquad I = 2P + 1,$$
$$m_1 = \min\{0, k_1\}, \qquad m_2 = m_1 + I - 1. \tag{6}$$

The choice of an initial value m_1 of the input window is rather arbitrary—it is important that the window embraces all the nonzero values of the input signal. Such an m_1 will affect the distribution only of the samples of the output window in the resultant array of the cyclic convolution.

In evaluating the DFT span of the cyclic convolution and the input window one should not overlook that efficient realization of DFT is not possible for all I. The value of I found by (4), (5) or (6) should be increased to the nearest value allowing the construction of a fast transform algorithm.

PARTITIONING OF CONVOLUTION

For the case shown in Fig. 2(a), that is when the width of the point spread function is small compared with the nonzero part of the input sequence, a further enhancement of efficiency may be achieved by partitioning the convolution.

Partitioning can be effected by two ways of about the same computational complexity [3]. In the method we take first, called overlap with summation, the input sequence of f_m is partitioned into contiguous blocks $f_u^{(v)}$ of length U.

Convolving each of these blocks with the h_n series of length $2P + 1$ results in an output series $g_k^{(v)}$ of length $U + 2P$, which overlap with the adjacent blocks in $2P$ samples and yield at the output the desired samples of the g_k sequence (Fig. 3a).

In the second method, called overlap with accumulation, the input sequence is divided into overlapping blocks. The undesired samples are eliminated at the output, resulting in contiguous blocks that form the required sequence, as shown in Fig. 3(b). This method is somewhat more efficient than the former and therefore preferable.

An optimal size U of an overlap block can be determined by choosing the minimum computational complexity of the convolution, i.e. the least time consumption for complex addition and multiplication.

For an output window size large compared with the width of the point spread function, we made an analysis of DFT algorithms with multiple "chopping in time" [2] with respect to various bases. This indicated that (for both one- and two-dimensional cases) an optimal block length should be such that the span of the corresponding cyclic convolution (and DFT) should be an integer power of two. This fact facilitates the search for an optimum and confines the class of algorithms to straightforward DFT procedures with base 2 chopping.

USE OF HARTLY TRANSFORM

In the case when both the signal and the point spread function of the system are real-valued, the complexity of computations and the amount of data to be processed may be further reduced to about one half by using a real valued DFT in the form of Hartly [4].

This transformation, being essentially the sum of the real and imaginary parts of the Fourier

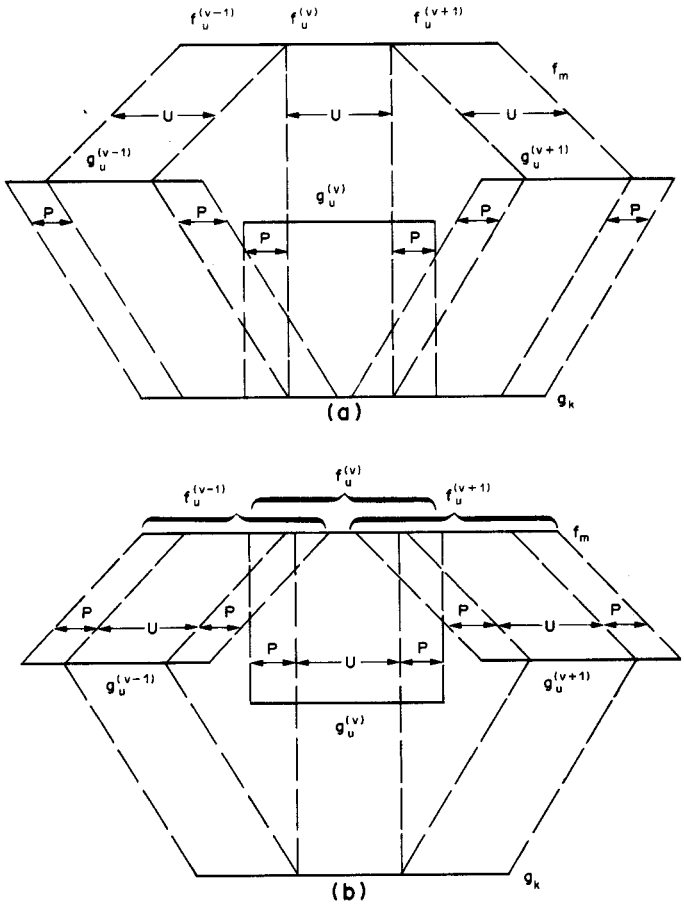

Fig. 3. Schematic representation of the arrangement of blocks of a partitioned convolution (a) for the method of overlapping with summation, and (b) for the method of overlapping with accumulation.

transform, may be used in computing the convolution provided the spectrum of at least one of the convolute sequences is real valued. This procedure is efficient because the one-dimensional Hartly transform can be computed through a one-dimensional DFT of half size. The Hartly transform is more advantageous for computing a convolution than the combined DFT algorithm [5] because in the former case the intermediate data, as well as input and output data, are real valued, whereas the Fourier transform of a real-valued function is in general complex valued.

GENERAL SIMULATION ALGORITHM

Given the point spread function of the system and the space limited signal, the algorithm for modeling a two-dimensional linear system becomes:

(1) Determine the dimensions of the sampled point spread function by one-dimensional inverse DFT, or by inverse DFT in the Hartly form if the impulse response function is real valued.
(2) Compute the parameters of the cyclic convolution based on the dimensions of the (nonzero) input data samples, size of the impulse response function and the constraints imposed on the input sampling domain by means of (4), (5) and (6). Evaluate the most efficient span of the transformation (Fourier or Hartly) or the optimal size of a block of partitioned convolution for the case shown in Fig. 2(a).
(3) Perform the direct (Fourier or Hartly) transformation for the two-dimensional array of input data, or for each block in the case of a partitioned convolution.
(4) Multiply the resultant discrete spectrum by the samples of the point spread function.
(5) Perform the inverse (Fourier or Hartly) transformation.

(6) Obtain the desired samples of the output signal, in the case of a partitioned convolution—after preliminary matching of the overlap blocks.

The entire method of numerical modeling for continuous two-dimensional linear systems ensures an adequate transformation of spatial signals in a computationally efficient manner.

REFERENCES

1. W. K. Pratt. *Digital Image Processing*. Wiley, New York (1977).
2. L. A. Rabiner and L. R. Gold. *Theory and Application of Digital Signal Processing*. Prentice Hall, Englewood Cliffs, N.J. (1975).
3. G. Nussbaumer. *Fast Fourier Transform and Convolution Algorithms*. Radio i Svyaz, Moscow (1984).
4. L. M. Soroko and T. A. Strizh. *Frequency Transformations on a Computer*. OIYaI Press, Dubna (1972).
5. L. P. Yaroslavsky. *An Introduction to Digital Image Processing*. Nauka, Moscow (1979).