CONVERGENCE AND SPEED OF THE SINC BASIS MOMENT METHOD

T. J. Cavicchi and W. D. O'Brien, Jr.

Bioacoustics Research Laboratory, Dept. of Elect. & Comp. Engr.
University of Illinois, 1406 W. Green Street, Urbana, IL 61801

Abstract

A method for applying the fast Fourier transform to the convolutions arising in the internal field equations of the sinc basis moment method diffraction tomography algorithm is described. Exactly equivalent results are obtained, while reducing the order of those computations from n^5 to $n^3 \log n$ for an nxn reconstruction.

As originally derived, the sinc basis moment method has an order of computation of n^5 for an nxn reconstruction. To reduce this order, spatial convolutions in the internal field equations were performed by a somewhat unusual application of the FFT. Specifically, if $f_{\phi l}$ and $f_{\phi l}{}^{inc}$ represent, respectively, the total and incident ultrasonic fields evaluated at pixel l due to a transmitter at location ϕ, γ_j is the object function evaluated at pixel j, and C_{lj} are coefficients resulting from the sinc basis expansions and integrations (Johnson and Tracy, 1983),

$$f_{\phi l} = f_{\phi l}{}^{inc} + \sum_{j=1}^{n^2} C_{lj}\gamma_j f_{\phi j} \qquad (1)$$

is a discrete version of the corresponding integral equation, suitable for digital computers. The difficulty in applying the discrete Fourier transform (DFT) to the convolutional sum [the sum in Eq. 1 can be shown to be convolutional (Cavicchi, 1987)] is that the region of support of C_{lj} is $R_{\infty,\infty}$ (numerically, the C_{lj} are significantly nonzero far outside the extent of the object region, $R_{n,n}$). But the region of evaluation of the convolution in Eq. 1 is only the object region. Leaving implicit in the field arguments the spacing scaling of the indices, the convolutional form of Eq. 1 is:

$$f_\phi(i_x,i_y) = f_\phi{}^{inc}(i_x,i_y) + \sum_{j_x=0}^{n-1}\sum_{j_y=0}^{n-1} \gamma f_\phi(j_x,j_y)C(i_x-j_x,\ i_y-j_y) \qquad (2)$$

Note now that because $0 \le i_x,\ i_y \le n-1$ (l in Eq. 1 is a pixel with coordinates $0 \le i_x,\ i_y \le n-1$), the only values of $C(\cdot,\cdot)$ that contribute to the aperiodic convolutional sum are those for which $0 \le i_x - j_x,\ i_y - j_y \le n-1$. Using the maximum value of j_x, n-1, and the minimum value of i_x,0, yields the lower limit -(n-1). Using the minimum value of j_x,0, and the maximum value of i_x, n-1, yields the upper limit, n-1. Thus, pictorially, the region of support that contributes to the convolution within the object region is as shown in Fig. 1. That is, substituting a truncated C as indicated in Fig. 1 will be mathematically equivalent to "using" the entire C. The solution of the present problem becomes evident when an alternative, pictorial interpretation of the previous formalized statements is examined. First note that although one has, mathematically equivalently, the convolution of two finite region of support sequences, the C sequence is centered on the origin; that is, the C are nonzero not only in the first quadrant, but in all four quadrants. (Note that in none of the following is the circular symmetry of the C_{lj} recognized or used.) Yet γf is centered on (n/2-1,n/2-1). This problem cannot be solved by merely shifting the C sequence, for then the values "called" in the sum will be incorrect. Consider the pictorial two-dimensional convolution in Fig. 2a of two finite region of support sequences f_1 and f_2 centered as would be desired in order to use the DFT. Now periodically extending f_2 after zero-padding it out to N, which is chosen to avoid aliasing, results in $\underline{f}_{2p}$, Fig. 2b, (where the underline denotes zero-padding and p denotes periodic extension). Only one repetition of f_1 was drawn because, looking at the convolutional sum,

$$f_1{**}f_2(i_x,i_y) = \sum_{j_x=0}^{N_1-1}\sum_{j_y=0}^{N_1-1} f_1(j_x,j_y)f_2(i_x-j_x,i_y-j_y),$$

it is seen that the argument of f_1 ranges only within R_{N_1,N_1} so that equivalently one need only periodically extend f_2, as far as the evaluation of the convolutional sum is concerned. (There is no harm in periodically extending f_1 as long as the limits of the spatial area of summation included in the aperiodic convolution are kept in mind.) Consider again the periodic convolution in Fig. 2b. Another way of viewing the "aliasing" problem is to remember, as noted before, that if, in particular, one is interested in values of the convolution of f_1 and f_2 on R_{N_1,N_1}, the arguments of f_2 will range from $-(N_1 - 1)$ to $(N_1 - 1)$. Thus, when "reaching back" for values in, say, the third quadrant, one should obtain zeros. But if f_2 is not zero-padded and then periodically extended, then instead of having Fig. 2b, one will have the situation depicted in Fig. 3: now when reaching back for a zero, one will instead obtain a replicated (nonzero) value of f_{2p}. This is another way of showing why and how much the input sequences must be zero-padded in order to simulate aperiodic convolution with periodic convolution.

Now consider the contributing block of C for the tomographic convolutions. Again, note that these are the only contributing values to the convolution in the region of interest $R_{n,n}$. If this function is periodically extended, the result is Fig. 4. Now if one chooses the block of C in the first quadrant region bounded by the axes and the dashed lines as the sequence of finite support to be sent to the FFT, and does not zero-pad, then when evaluating the convolutional sum in $R_{n,n}$ and "reaching back" into the third quadrant for replications, one obtains (as always) the replica of the upper right hand side section of the dashed-line finite sequence, which is exactly equal to what would be obtained in the aperiodic convolution; that is, these very contributions are necessary to simulate the aperiodic convolution. Of course, γf must be zero-padded out to the size of C, which is chosen for computational convenience to be the lowest power of two greater than or equal to 2n - 1. Note that, outside of $R_{n,n}$, the result will be garbage, but those values can be thrown away.

For an example of this situation consider two cases. The first case is illustrated in Fig. 5a. In this case, the point of evaluation is inside the object region. Because one replication of C completely covers the object region, the other replications of C make no contribution to the convolutional sum, matching the situation of aperiodic convolution with the original C. In the second case (Fig. 5b) the point of evaluation is outside the object region. Note here that other replications of C lie over part of the object region, making contributions that are erroneous with respect to the aperiodic convolution desired. The final result, therefore, upon inverse inverse DFT of $\{\widetilde{\gamma f_\phi} \cdot \tilde{C}\}$ where $\tilde{C}$ is the DFT of C as described in the above discussion, is shown in Fig. 6a, which is the periodic convolution of γf with the finite extent, periodically extended C_p. The result in the small square is identical to the result in the same region of the aperiodic convolution of γf with the infinite extent C shown in Fig. 6b.

The above theory is verified practically in Fig. 7, which shows the real part of the actual convolutions carried out in the spatial (above) and Fourier (below) domains. They agree exactly in the region of interest and the garbage values from the FFT method outside of these convolutions object region may be discarded. The computational complexity is reduced by this use of the FFT from n^5 to $n^3 \log n$.

Incidentally, if one considers again the possibility of using the FFT by a similar method for the measured scattered field equations

$$f_{\phi m}{}^{sc} = \sum_{j=1}^{n^2} D_{mj}\gamma_j f_{\phi j} \qquad (3)$$

where $f_{\phi m}{}^{sc}$ is the scattered field measured at receiver m due to transmitter ϕ and D_{mj} are the associated coefficients from the sinc basis expansions and integrations, the first consideration will be which values of the D_{mj} contribute to the convolutional sum. Given that the subscript m refers to the receiver and j refers to a pixel in the object region, the difference vectors will fall somewhere on a ring centered a distance R from the origin, given that the receivers are on a ring a distance R from the object region center (see Fig. 8a). Consequently, what was a finite square contributing region of support for the C_{lj} (see figure 8b) becomes roughly a ring contributing region of support, outside of which assuming zero for D_{mj} results in no error. However, such a region is not acceptable for FFT computations demonstrating the infeasibility of trying to use the FFT for the scattered field equations convolutions.

Acknowledgements

Partial support for this research has been provided by the NIH Radiation Oncology Training Fellowship, CA 09067.

References

T.J. Cavicchi, "Diffraction tomography and the sinc basis moment method," Ph.D. thesis, Department of Electrical and Computer Engineering, University of Illinois, 1987.

S.A. Johnson and M.L. Tracy, "Inverse scattering solutions by a sinc basis, multiple source moment method - Part I. Theory," Ultrasonic Imaging, vol. 5, pp. 361-375, 1983.

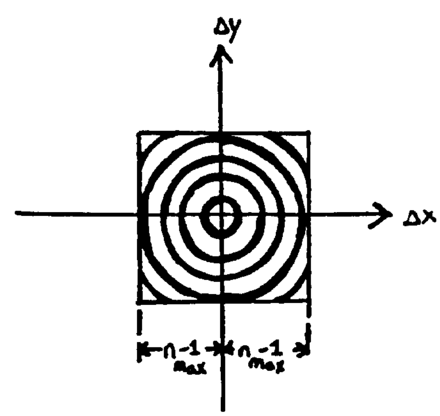
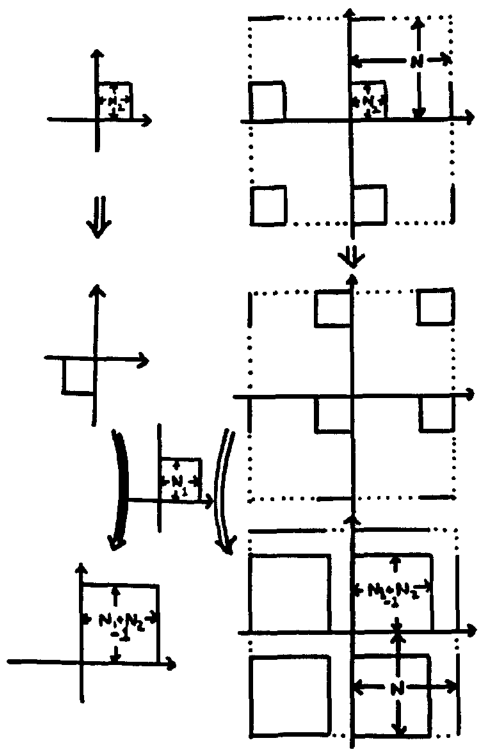


Figure 1 Spatial region for which C_{lj} contribute to the convolutional sum in the internal field Eq. 1, which is evaluated within the object region only. Thus, even though the C_{lj} have an infinite region of support, in this context they may be treated as having finite support.



(a) (b)

Figure 2 The aperiodic (Fig. 2a) and the periodic (Fig. 2b) convolution of a function of spatial extent of support $N_1 x N_1$ with another function of spatial extent of support $N_2 x N_2$. The result is shown at the bottom, where that for Fig. 2b is a periodically repeated version of that for Fig. 2a, with period N. In Fig. 2b, N was chosen large enough (via zero-padding) so that the repetitions of f_2 do not interfere with each other (no aliasing).

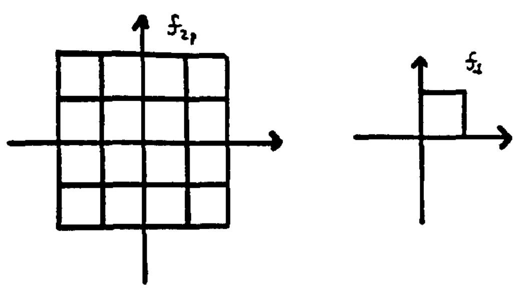


Figure 3 In Fig. 2b, when in the convolutional sum a value of f_2 was called in the third quadrant, e.g., a zero was correctly obtained as in the aperiodic convolution because of zero-padding. If, however, the function f_2 is merely repeated without first zero-padding, a nonzero value from another repetition will be obtained from the third quadrant, causing the periodic convolution to differ from the corresponding aperiodic convolution.

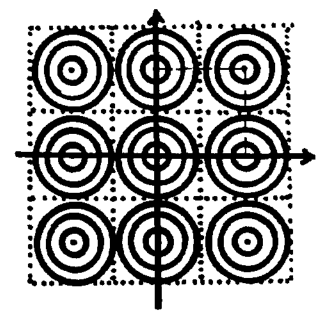
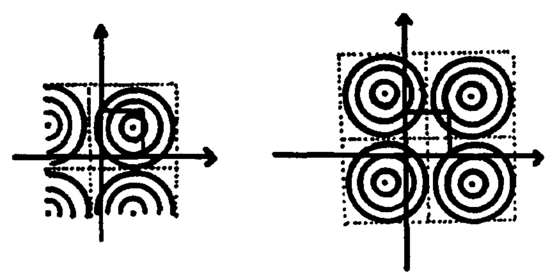
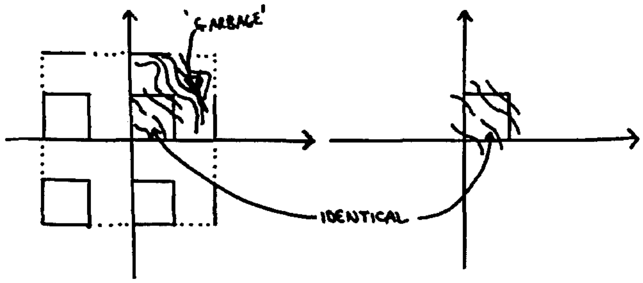


Figure 4 Periodic extension of the region of C_{1j} contributing to the convolutional field sum evaluated in the object region (see Fig. 1). The first quadrant region bounded by the axes and the dashed lines is the sequence chosen as input to the FFT.



(a) (b)

Figure 5 Periodic convolution process carried out in the spatial domain. The convolvees are the periodic extension of truncated C (see Fig. 4) and γf zero-padded out to the truncation size of C. If the convolutional sum is evaluated within the object region (Fig. 5a), one repetition of the contributing block of C completely covers the object region (shown boxed), and the periodic convolution equals the aperiodic convolution. If the convolutional sum is evaluated outside the object region (Fig. 5b), parts of more than one repetition cover the object region, making erroneous contributions to the convolutional sum.



(a) (b)

Figure 6 The result of periodic (C truncated) (Fig. 6a) and aperiodic (Fig. 6b) convolutions of C and γf. Within the object region (shown boxed), the two results are identical, but outside the object region they differ. As only values within the object region are needed, the other values returned from the periodic convolution are simply discarded.

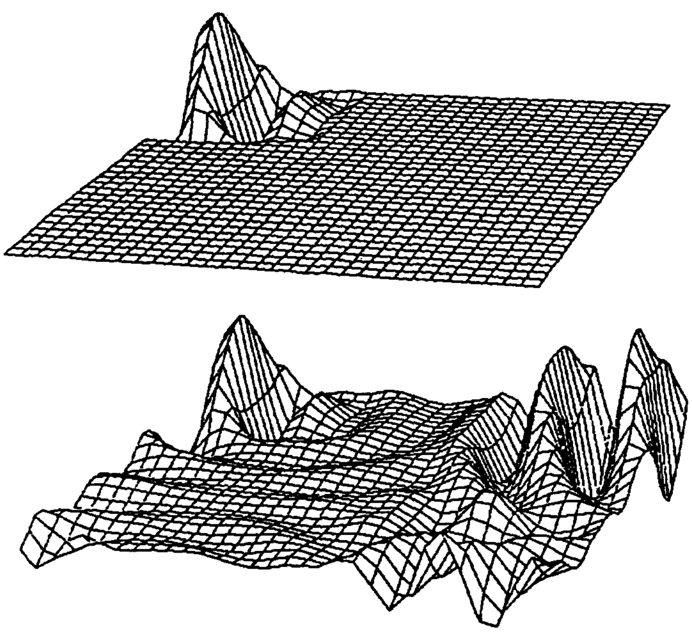


Figure 7 Real part of the aperiodic (above) and periodic (below) convolutions of C and γf as defined in Eq. 1. Within the object region (square region where the aperiodic convolution is nonzero) the aperiodic and periodic results are identical; outside the object region, the periodic results are discarded.

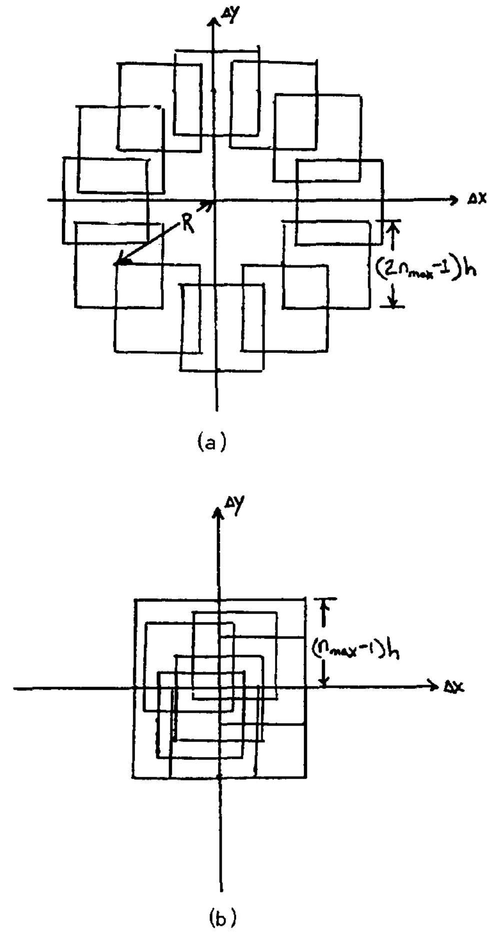


Figure 8 Contributing regions of D_{mj} (Fig. 8a) and C_{lj} (Fig. 8b) for individual points of evaluation of the convolutions in Eqs. 3 (Fig. 8a) and 1 (Fig. 8b), where h is the pixel spacing. The points of evaluation are receivers (Fig. 8a) and object region pixels (Fig. 8b). The situation in Fig. 8b is amenable to FFT application, whereas that in Fig. 8a is not.