COMPUTATIONAL ASPECTS OF BISPECTRAL ANALYSIS IN INTERFEROMETRIC IMAGING

Peter W. Gorham
Mailcode 220-47, California Institute of Technology
Pasadena, CA 91125 U.S.A.

I. Introduction

Although many approaches to phase recovery in the techniques of speckle interferometry and discrete-element optical interferometry are now being currently used, the most promising are those which make use of the closure phase principle (Jennison 1957; Rogstad 1968), which provides an observable phase which is immune to atmospheric corruption, and contains the desired object phases. The general mathematical support for the closure phase quantity is provided by the bispectrum function (Hoffmann et al 1983;), which is a third moment of the complex visibility of the observed (and therefore atmospherically corrupted) source distribution. Specifically, triple products of all visibility elements which can be mapped onto a triangle of discrete interferometer elements, are included in the bispectrum. This constraint implies that the bispectrum volume spans four dimensions, since it is effectively a vector product of the aperture plane with itself, corresponding to the two independent legs of the baseline triangles.

A single-dish, large aperture telescope may be treated as a limiting case of a close-packed array of discrete interferometer elements, each having a diameter of order the Fried parameter r_0, ~10 cm for ~1 arcsec seeing. All pairs of sub-apertures will thus combine in the image plane to form an interference pattern (the speckle pattern), and the image is thus the Fourier transform of the source visibility function, convolved with the atmosphere+telescope transfer function. This relation provides the basis for visibility amplitude recovery in speckle interferometry, through deconvolution of the atmospheric transfer function, measured by observing a point source near the object of interest (Labeyrie 1971).

Since the number of complex visibility elements in an $M \times M$ image is $M^2/2$, the bispectrum may be expected to contain $\leq M^4$ elements. For a typical image with $M \geq 10^2$, the bispectrum will contain as many as 10^8 complex elements, requiring close to 1 Gbyte of storage space. In addition, bispectral analysis typically requires the accumulation of the bispectrum function for at least 10^{4-5} "snapshot" frames of the image. Since each frame requires of order 1-2 Gigaflop of computation, the computing needs, as well as the storage requirements, are obviously prohibitive for most presently available machines. One set of solutions to these problems are:

[A] Compute only a reduced portion of the bispectrum, chosen to provide adequate information for the image reconstruction, but perhaps sacrificing some dynamic range in the final map;

[B] Use a CRAY or comparable machine;

[C] Develop algorithms to minimize the computational and storage needs of the full bispectrum and find powerful and cheap computers to do the job.

In the following, we discuss progress we have made with approach [C]. We have assumed that solution [A], which has been successful up to the present, will ultimately be deemed unsatisfactory if these techniques are to be pushed to their limits; and we have avoided solution [B] because we were interested in exploring the alternative of relatively inexpensive concurrent processing computers, which we feel will soon supplant the sequential machines such as the CRAY series for large scale data analysis.

II. The Non-degenerate Bispectral 4-volume

To calculate a bispectrum function which contains only the information which is non-degenerate (e.g, cannot be obtained from some other portion of the function by symmetry or antisymmetry properties) is desireable from the point of view of both memory and computation time. Here we evaluate the constraints on the bispectrum from the properties of the Fourier transform from which it is constructed (these properties completely define the symmetry properties of the bispectrum, as well as its boundaries)

Properties of the Image Fourier Transform. We assume the image (in this case one frame of a set of speckle images) is sampled at a frequency exactly twice that of the highest spatial frequency present, which for a telescope of aperture D is $\sim D/\lambda$, where "spatial frequency" is actually angular frequency on the sky. In this case we have exactly enough information to reconstruct all the information in the image, according to the Shannon sampling theorem. An $M \times M$ square pixel array is oversampled along the diagonals, and the Fourier transform array will thus contain some repetition in the corners, but we can compensate for this by only considering frequencies within a radius $M/2$ of the transform origin.

The image has M^2 real values; the transform has $M^2/2$ complex values within the region $-M/2 \leq k_1 \leq M/2$; $0 \leq k_2 \leq M/2$, where $(k_1,k_2)=\mathbf{k}$ is the transform position vector. Each $\mathbf{k}$ corresponds to a unique set of "fringes" across the image plane, which are normal to the direction of $\mathbf{k}$ and proportional in frequency to $|\mathbf{k}|$. These fringes in turn correspond to a pair of sub-apertures with separation $2\mathbf{k}D/M$ in the aperture plane, and thus there exists a map from the aperture to the transform plane. This map is not one-to-one, since many different pairs of sub-apertures (with congruent baselines) can correspond to the same $\mathbf{k}$; however, if translations are ignored, the map preserves orientations and relative baseline lengths.

If we define the values of the discrete Fourier transform of the image to be $B(\mathbf{k})$, then the following properties apply:

Conjugate symmetry to point reflections around the origin:

$$B(\mathbf{k}) = B^*(-\mathbf{k}) \tag{1}$$

Periodicity with period M, for $M \times M$ image:

$$B(k_1,k_2) = B(k_1 \pm M,k_2) = B(k_1,k_2 \pm M) = B(k_1 \pm M,k_2 \pm M) \tag{2}$$

Also, the combination of (1) and (2) above implies that the real part of $B(\mathbf{k})$ is an even function, and the imaginary part of $B(\mathbf{k})$ is an odd function in $\mathbf{k}$. This property implies that there is a square lattice of "reflection points" spaced by $M/2$ rather than M; thus we get the limits on $\mathbf{k}$ mentioned previously.

A Modified Bispectrum. We have noted that there is a map between $\mathbf{k}$ and pairs of sub-apertures; thus a pair of $\mathbf{k}$-vectors determines a triangle in the aperture plane, and thus a closure phase around the triangle, which includes the origin as one of the corresponding aperture positions. The bispectrum is conventionally defined as:

$$B^{(3)}(\mathbf{u},\mathbf{v}) = B(\mathbf{u})B(\mathbf{v})B^*(\mathbf{u}+\mathbf{v}). \tag{3}$$

(e.g., Bartelt, et al 1984). If we now choose $\mathbf{u},\mathbf{v}$ from the $\mathbf{k}$-plane, we will not get the triangle we expect since $-(\mathbf{u}+\mathbf{v})$ does not give the final leg to the triangle $\mathbf{O},\mathbf{u},\mathbf{v}$. We are better off choosing $\mathbf{u},\mathbf{u}+\mathbf{v}$ as the initial pair, since these will form a triangle closed by $\mathbf{v}$. We can also choose $\mathbf{v}$ from the "head" of $\mathbf{u}$ but this computationally cumbersome, since we must translate the entire plane to the head of $\mathbf{u}$ in order to retain the proper domain limits for $\mathbf{v}$. It is more sensible to keep the domain of $\mathbf{u},\mathbf{v}$ the same, and redefine the bispectrum slightly so that it corresponds more closely to our way of thinking about triangles and closure phases.

A suitable redefinition, which preserves all the symmetry/anti-symmetry properties of the old bispectrum, is:

$$B^{(3)}(\mathbf{u},\mathbf{v}) \equiv B(\mathbf{u})B^*(\mathbf{v})B(\mathbf{v}-\mathbf{u}). \tag{4}$$

It is easy to verify that the following properties apply:

$$B^{(3)}(\mathbf{u},\mathbf{v}) = B^{(3)*}(-\mathbf{u},-\mathbf{v}) \tag{5}$$

$$B^{(3)}(\mathbf{u},\mathbf{v}) = B^{(3)*}(\mathbf{v},\mathbf{u}) \tag{6}$$

$$B^{(3)}(\mathbf{u},\mathbf{v}) = B^{(3)}(\mathbf{u},\mathbf{u}-\mathbf{v}) \tag{7}$$

where we need only use equation (1) on each term in the domain of the bispectrum. Note that equation (5) implies antisymmetry on point reflection around the origin (which gives the "anti-triangle") , equation (6) implies antisymmetry for interchange of pairs of triangle legs, and equation (7) implies symmetry for a cyclic permutation of vertices. We will use the bispectrum of equation (4) in what follows.

Domain Constraints on the Bispectrum 4-Volume. We see now that the process of constructing the bispectrum using the FFT approach consists of choosing pairs of vectors out of the **k**-plane. However, equations (5)-(7) imply that choosing one pair of vectors actually determines a number of other bispectrum points, which are degenerate in their value of $B^{(3)}$ with the original point. In fact there are 12 possibilities for each **u**,**v**:

$$\begin{array}{cccc} B^{(3)}(\mathbf{u},\mathbf{v}) & B^{(3)*}(-\mathbf{u},-\mathbf{v}) & B^{(3)*}(\mathbf{v},\mathbf{u}) & B^{(3)}(-\mathbf{v},-\mathbf{u}) \\ B^{(3)}(\mathbf{u},\mathbf{u}-\mathbf{v}) & B^{(3)*}(\mathbf{u}-\mathbf{v},\mathbf{u}) & B^{(3)}(\mathbf{v}-\mathbf{u},\mathbf{v}) & B^{(3)*}(\mathbf{v},\mathbf{v}-\mathbf{u}) \\ B^{(3)*}(-\mathbf{u},\mathbf{v}-\mathbf{u}) & B^{(3)}(\mathbf{v}-\mathbf{u},-\mathbf{u}) & B^{(3)*}(\mathbf{u}-\mathbf{v},-\mathbf{v}) & B^{(3)*}(-\mathbf{v},\mathbf{u}-\mathbf{v}) \end{array} \tag{8}$$

all of which are equal. Physically these 12 points correspond to the six permutations of three vertices and their reflections.

Equation (8) tells us that, for a given set of (**u**,**v**) we have a 12-fold degeneracy in the bispectrum. The standard 8-fold degeneracy given by Bartelt et al 1984, and Lohmann et al 1983, appears to apply only for a 1-dimensional image, since in this case all of the **k**-plane triangles are collapsed, and four of the corresponding symmetries are thus frozen out.

We can now make an upper limit estimate to the non-degenerate bispectrum volume. We assume that our (**u**,**v**) are taken in pairs from the domain given by the **k**-plane, extended to $k_2 = -M/2$ to include the values of $-\mathbf{u},-\mathbf{v}$, so that all points in equ. (8) may be found in the domain. We assume that all of the vector sums of **u** and **v** also lie in this domain (in fact they don't, but this will only provide tighter constraints later). The 4-volume $V^{(4)}$ of the bispectrum is then:

$$V^{(4)}(M) \leq \frac{M^2(M^2-1)}{2!} \times \frac{1}{12} \approx \frac{M^4}{24}\,. \tag{9}$$

In practice, it is unphysical to include a (**u**,**v**) for which $\mathbf{v}-\mathbf{u}$ is not in the domain, since it then corresponds to a higher frequency than any that is present. In fact, the limit is even more stringent than this: the three vertices **O**,**u**,**v** must all lie within a circle of diameter $M/2$ in the **k**-plane, corresponding to the similar requirement in the aperture plane. However, such a constraint is not easy to parameterize. Equation (9) also does not account for the fact that some triangles are equal to one another under the interchange of (**u**,**v**), thus reducing the 12-fold degeneracy for some fraction of the points, and making equ. (9) an overestimate nominally. However, such triangles are a small fraction of the total.

We may improve the estimate in equ. (9) by applying the constraint that our pairs (**u**,**v**) must come from within a radius $M/2$ in the **k**-plane, which provides a constraint somewhat weaker than, but similar to, the aperture containment constraint discussed above. This yields:

$$V^{(4)}(M) \leq \frac{\pi^2 M^4}{32} \times \frac{1}{12} \approx \frac{M^4}{39} \qquad (11)$$

which should now be close to the actual value that we will need to work with. Equ. (11) still contains unphysical triangles, for example, the right triangle with two legs of length $M/2$ and an unphysical hypotenuse; it has still "crossed out" some valid triangles that reflected into themselves, but these two groups tend to cancel one another in the counting.

To further refine these domain constraints analytically is unnecessary; they are already sufficiently complicated as to be best applied numerically to a given problem. However, we have already achieved a factor of ~40 reduction in the storage and computation requirements, and we find a factor of at least five discrepancy between this and earlier estimates of the bispectrum non-degenerate 4-volume (c.f Wirnitzer 1985).

III. Computing The Bispectrum.

Since the bispectrum depends on the four coordinates u_1, u_2, v_1, v_2, we could naively approach the computation of the bispectrum with a set of nested indexed loops, in which we perform simple indexing operations to determine which combinations of indices we will accept as valid bispectrum points. By "simple indexing operations" for computing the bispectrum, we mean that we can define a contiguous sequence $i_0, j_0, k_0, m_0 \cdots i_p, j_q, k_r, m_s$ where the indices represent the components (**u**,**v**) on a one-to-one basis, which covers all of the (**u**,**v**) required to get a complete bispectrum. This implies that the 4-volume is in fact a right-handed Euclidean space (a 4-polygon). The indices may also be linear functions of one another successively (except for the first one which must be independent). However, operations like "compare this triangle x to the fourth permutation of previous triangle y" are very difficult to code in this fashion, if it can be done at all. Thus, even if such coding as indexing functions can lead to the bispectrum, one pays a large computing price in the index computation, which then begins to dominate the whole calculation.

The solution is to pre-compute all the (**u**,**v**) necessary for the bispectrum, once for all time, since it will be the same for every bispectrum for a given M. For $M \leq 256$ we can store each index in a single byte (or ASCII character); thus only 4 bytes are required per (**u**,**v**). Using the estimate of equation (11) above and M=128, with 12 bytes per bispectrum element (4 for the (**u**,**v**), and 8 for each complex element) we will require $\sim$80 Mbyte for the full bispectrum.

The symmetries in equations (5)-(8), as well as the aperture constraint, may be applied as domain constraints in three steps:

[1] **u**,**v** ,**v**—**u** must all lie in the positive half of the **k**- plane, thus eliminating point-reflection symmetries;

[2] The angle of **v** with respect to the $k_1 = 0$ axis must be $\leq$ the corresponding angle of **u**, thus eliminating skew symmetries;

[3] As stated above, the triangle **O**, **u**,**v** must fit within the spatial frequency limit of the telescope, e.g, the circle that circumscribes the baseline triangle must have a diameter $\leq M/2$.

We have coded these constraints and checked the resulting bispectral components by hand for completeness and non-degeneracy on small scale images (M = 8,16). The boundaries on each constraint may be open or closed as required, and elements where one of the baselines have zero length may be included, since they represent the power spectrum information, and may be necessary for bias corrections. We have also tested the completeness of the information content of the resulting bispectrum through a simulation (Smith 1987) which included Monte Carlo generated speckle frames (using Kolmogorov theory), and image reconstruction using an SNR-weighted phase-recursion algorithm based on that of Bartelt (1984).

The Bispectrum 4-Volume Distribution. In figure 1, we plot of the number density of non-degenerate bispectrum elements determined by the constraints above, vs. two parameters that characterize the type of aperture-plane closure triangle to which a given bispectrum element corresponds: (a) the length of the minimum baseline; and (b) the sum of all three triangle baselines. Parameter (a) indicates the distance of a given element from the axes of the bispectrum (the on-axis information is the image power spectrum, and thus has no phase information); and Parameter (b) indicates the degree of higher- spatial frequency information in a given closure triangle. The lengths in fig. 1 are

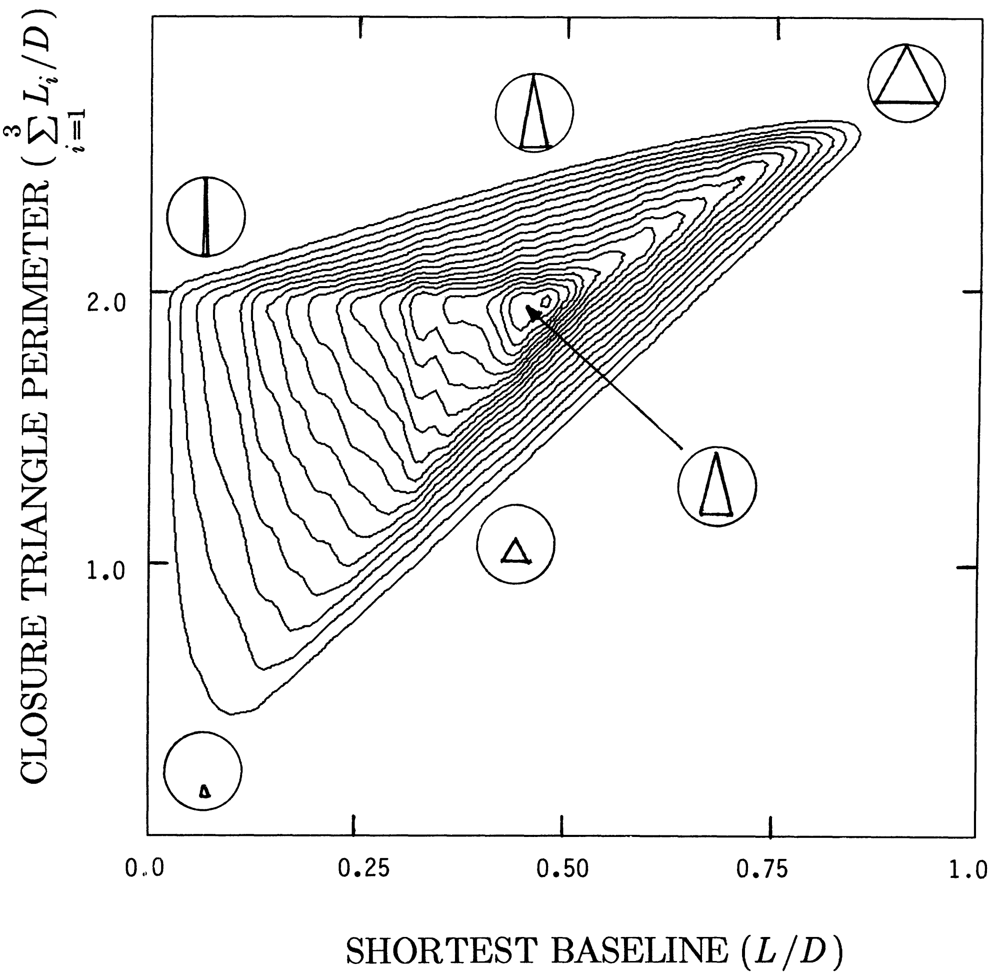

FIGURE 1: Contours of equal number density of bispectrum elements for a fully-filled telescope aperture, plotted according to the length of the shortest baseline, and the sum of all three baselines in the aperture plane closure triangle corresponding to each element. The lengths are normalized to telescope diameter D. Contours go from 5% to 95% of the maximum, and the form of triangles associated with different portions of the plot are shown scaled within a circular aperture.

normalized to the diameter of the telescope, and the types of triangles and their relative size in the aperture are indicated near the corresponding portions of the plot. Contours run from 5% to 95% of the maximum value.

Although this plot contains no information about the relative signal-to-noise ratio per bispectrum element that can be expected in practice, it does provide a summary of the distribution of bispectrum volume in terms of domain parameters, and provides an intuitive understanding of the type of closure phase information that can be derived from different portions of the bispectrum. At present, most results on bispectral analysis have limited themselves to the region very near the leftmost portion of the plot, where the closure triangles all contain one very short leg. We note that such triangles have baseline covariance properties which may be undesireable for unbiased image reconstruction (Kulkarni, this proceedings); certainly an interferometer array is not generally designed to conform to such a pattern.

Exploiting Concurrency in the Bispectrum Computation. Once a lookup table for the bispectrum elements if generated under these domain constraints, the computation of the triple products is conceptually simple. Yet although algorithms for sequential computation of the bispectrum for a sequence of speckle frames are straightforward to develop in principle, in practice they require unrealistically large memories or floating point speed, for values of $M \geq 100$, thus limiting their use for larger image sizes to only the highest performance supercomputers. Alternatively, much less expensive, and therefore much more accessible, multi-processor computers are achieving performance levels comparable to the CRAY X/MP-class of supercomputers, and will certainly continue to improve in the near future, probably achieving Teraflops performance with near Terabyte fast memories within the next decade or so, at a small fraction of the cost per Mflop or Mbyte of the present high end of the supercomputer class (c.f. Fox et al 1987).

One such class of concurrent processing machines for which there is much local expertise at Caltech is that which is based on a *hypercube* topology for the interprocessor communication network. In the hypercube topology, there are $N = 2^K$ processors, each in direct communication with K other processors, which thus forms a topological cube in K dimensions with each processor at one "corner" of the cube. Such a network avoids the $\sim N^2$ communication links required if all processors are directly linked to one another, yet a message from any processing node to any other node will require forwarding at most $K-1$ times. The hypercube topology also contains as subsets many other useful topologies, and in fact represents the optimal communication topology for distributed FFT (Fast Fourier Transform) algorithms in a concurrent processing environment. However, present concurrent processing machines generally cannot communicate as quickly as they can calculate; thus algorithms which minimize interprocessor communication tend to come closest to achieving the maximum performance on such machines.

One relatively straightforward adaptation of the bispectrum computation to a hypercube environment is outlined as follows:

(a) Distribute list of **u**,**v** bispectrum elements: each processor will compute $\sim(M^4/39)/N_{proc}$ elements for its frame

(b) Divide data frames among processors; $FFT(I_j(\mathbf{x}))\rightarrow B_j(\mathbf{u})$ for j_{th} processor

(c) calculate $B_j^{(3)}(\mathbf{u},\mathbf{v})$ and accumulate

(d) shift $B_j(\mathbf{u})$ to proc $(j+1)mod(N_{proc})$

(e) loop over (c) - (d) $N_{proc}-1$ times

(f) loop over (b) - (e) N_{frame}/N_{proc} times

The hypercube parallelism is very efficient here: the communication is done in large packets, using a *ring topology* , in which each processor communicates directly with two adjacent processors, and no message forwarding is necessary. The FFT and the bispectrum calculation are done locally on each node, avoiding any loss of efficiency to communication in a distributed FFT.

Status of the Concurrent Implementation. The algorithm described here constrains some of the characteristics of the processing nodes on which it is implemented; they must be relatively *large grained* in order to perform significant computational steps such as the local FFT on an image with $M \geq 100$, for example. Our present implementation in being done on a 512-processor hypercube made by the NCUBE corporation. Each processor has 1/2 Mbyte of memory, and is capable of $\sim$200 Kflops. Inter-processor communication is done with a library of communication routines which are accessible to the user program as function calls from the processor programs, in either C or FORTRAN. The user supplies a single parent program which runs individually on each processor in parallel, and operations which are processor dependent are flagged by processor number within this program. Thus code for sequential machines is often quite easily portable to the concurrent environment with minimal changes.

We have coded the bispectrum computation for the Caltech NCUBE, and we are in the process of analyzing data from the Hale 5 meter telescope taken in September 1987, and April 1988. We find at present that we achieve sustained floating point speeds of about 30 Mflops with 512 nodes; this value is about 30% of the maximum 100 Mflops that the machine is capable of with no communication. Inter-processor communication inefficiency accounts for this discrepancy, but we find that we are able to process $\sim$3500 frames/hour of 128^2 images, and we can thus reduce a 10^5 frame data set in $\sim$1 day. In addition, we are developing concurrent algorithms for image reconstruction using a minimization approach on the entire bispectrum, as discussed by Cornwell (1987).

The author wishes to thank G. Smith, T. Nakajima, S. Kulkarni, A. Readhead, and T. Prince for useful discussion and criticism, and G. Fox and the Caltech Concurrent Supercomputing Initiative for support. This work was supported in part by the U.S. National Science Foundation, the U. S. Dept of Energy, and the W. M. Keck Foundation.

References

--Bartelt, H., Lohmann, A. W., and Wirnitzer, B. Appl. Optics, **23** , pp 3121, (1984).
Cornwell, T. J., Astr. Astroph., in press (1987).
--Fox, G. C., et al, *Solving Problems on Concurrent Processors* , (Inglewood Cliffs, N.J.: Prentice Hall), (1988).
--Jennison, R. C., MNRAS, **118,** pp 276, (1958).
--Labeyrie, A., Astron. Astrophys., 6, pp 85 (1970).
--Lohmann, A. W., Weigelt, G., and Wirnitzer, B., Appl. Optics, **22** , pp 4028, (1983).
--Rogstad, D. H., Appl. Optics, **7** , pp 585, (1968).
--Smith, G., Caltech Internal Report, (1987).
--Wirnitzer, B., J. Opt. Soc. Am. A, **2** , pp 14, (1985).