Generalized Cross-Correlation Functions for Engineering Applications, Part I: Basic Theory

Traditional cross-correlation considers situations where two functions or data sets are linked by a constant shift either in time or space. Correlation provides estimates of such shifts even in the presence of considerable noise corruption. This makes the technique valuable in applications like sonar, displacement or velocity determination and pattern recognition. When regions are decomposed into patches in applications such as Particle Image Velocimetry it also allows estimates to be made of whole displacement/flow fields. The fundamental problem with traditional correlation is that patch size and hence statistical reliability must be compromised with resolution. This article develops a natural generalization of cross-correlation which removes the need for such compromises by replacing the constant shift with a function of time or space. This permits correlation to be applied globally to a whole domain retaining any long-range coherences present and dramatically improves statistical reliability by using all the data present in the domain for each estimate.

1 Introduction

The cross-correlation of two function or data sets $f_1$ and $f_2$ (Weiner, 1949, 1964) is a very common tool in applications as diverse as sonar, flow determination, and pattern recognition in badly corrupted data (Trahey et al., 1969; Coupland and Halliwell, 1992; Richards and Roberts, 1971; Lee, 1960; Matic et al., 1991; Berryman and Blair, 1986; Dejong et al., 1991; Gonzalez and Woods, 1992). Conventional cross-correlation typically applies to situations where the quantities of interest $f_1$ and $f_2$ are related by a simple constant shift $\tau$:

$$f_2(\xi) = f_1(\xi + \tau)$$

and the aim is to obtain a statistically reliable estimate of what will be termed here the transformation parameter $\tau$. In time series work $\xi$ is the time $t$ and $\tau$ a time delay while in spatial applications $\xi$ and $\tau$ define $N$ dimensional displacement vectors, rotations, or some combination of these (Gonzalez and Woods, 1992; Kamachi, 1989).

For applications like particle image velocimetry (Willert and Gharib, 1991; Utami et al., 1991; Adrian, 1986), where a whole flow field is characterized, cellular correlation has been developed (Kamachi, 1989; Leese et al., 1971; Ninnis et al., 1986). The displacement/flow field is made visible in some way with $f_1$ and $f_2$ being consecutive images of the displacement/flow. The images are segmented into patches and cross-correlation is then applied essentially to each patch in turn to determine an average displacement/flow velocity for each such cell.

The main problem with this approach is that increasing spatial resolution means reducing the patch size. As typical applications involve digitized noise corrupted data this reduces the information available in each patch for correlation and thus degrades the reliability of the estimates. If the noise is Gaussian variance of the sample estimate is inversely proportional to patch size. Furthermore, treating each cell independently loses the information theoretic advantages stemming from intercellular coherence in the displacement/flow field. These problems could be avoided if cross-correlation could be generalized to allow spatial or temporal variation of the shift $\tau$ and the present article is concerned with developing such a Generalized Cross-Correlation (denoted as GC-C).

The range of uses of conventional cross-correlation extends far beyond the description of displacement/flow problems cited here in both practical and analytical areas, and consequently the scope of GC-C is expected to be even wider.

The treatment here in Part I is in terms of continuous variables while issues associated with discretisation are addressed in the companion work Part II, Belmont et al. (1997).

2 The Properties Required of a Cross-Correlation Function

The first step in developing the GC-C is to specify those features which a cross-correlation function of any kind should exhibit. These are natural extensions of the characteristics exhibited by conventional cross-correlation (Weiner, 1949, 1964):

1. Cross-correlation should operate upon a pair of functions, or data sets, (in its discrete form), denoted as $f_1$ and $f_2$.
2. If $f_1$ and $f_2$ are connected by some transformation of their independent variables, then the cross-correlation function should exhibit an absolute maximum when a matching transformation is induced by cross-correlation processes. A corollary of this is that the location of the maximum should allow the computation of any parameters associated with the transformation, e.g., $\tau$ in Eq. (1).
3. The cross-correlation function should approach the absolute maximum smoothly.
4. Points 2 and 3 should hold even if $f_1$ and $f_2$ are contaminated by extraneous additive components that are uncorrelated between $f_1$ and $f_2$. 

Contributed by the Applied Mechanics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF APPLIED MECHANICS.
3 A Generalized Cross-Correlation Function

3.1 Parameterizing the Shift Function. To be useful it is preferable that the variable shift should be generated from a parameterized set of known functions, then determining the shift is achieved by estimating the so-called transformation parameters. Consequently the GC-C function must contain correlation parameters that, in effect, are varied to match the transformation parameters.

Such a generalized cross-correlation function of m correlation parameters, \( \mu_1, \ldots, \mu_m \), derived from two multidimensional functions \( f_1 \) and \( f_2 \), will be denoted as \( K_{i,2}(\mu_1, \ldots, \mu_m) \). The development of such functions begins with a consideration of the simplest case denoted as \( K_{i,2}(\mu) \), i.e., where the shift function depends upon only the single parameter \( \mu \) and the data functions \( f_1(\xi) \) and \( f_2(\xi) \) are one-dimensional. This means that \( f_1(\xi) \) and \( f_2(\xi) \) are related by

\[
f_2(\xi) = f_1(\xi + S(\xi, \mu_0))
\]

where the variable shift function \( S(\xi, \mu_0) \) is a member of the family of functions, \( S(\xi, \mu) \), that are continuous over \( \Gamma \) in both \( \xi \) and \( \mu \) up to their second derivatives.

3.2 The Approach Used to Develop GC-C. GC-C will be developed via an extension of the technique used by Wiener to prove that a conventional correlation function exhibits a smooth absolute maximum (Wiener, 1949, 1964). A special case of this treatment has been employed in the development of the Dilatation Correlation Function (Belmont et al., 1991), which is related to Wavelets and Melin transforms and measures stretching transformations. The Dilatation Correlation has been used to investigate expansion effects in combustion gases (Belmont et al., 1991).

The present method establishes an inequality, whose LHS is positive and is independent of the correlation parameter \( \mu \), and also reduces to an equality when the correlation parameter \( \mu \) equals the transformation parameter \( \mu_0 \). The RHS side of this inequality serves to define the GC-C function. Such an inequality is

\[
\int_{\Gamma} \left( f_1(\xi + S(\xi, \mu)) R(\xi, \mu) - W(\xi, \mu) f_2(\xi) \right)^2 d\xi \geq 0 \tag{3}
\]

where now \( R \) and \( W \) are as yet unassigned.

Expanding (3) produces

\[
\int_{\Gamma} f_1(\xi + S(\xi, \mu))^2 R(\xi, \mu) d\xi
+ \int_{\Gamma} f_2(\xi)^2 W(\xi, \mu) d\xi
= 2 \int_{\Gamma} f_1(\xi + S(\xi, \mu)) f_2(\xi) R(\xi, \mu) W(\xi, \mu) d\xi. \tag{4}
\]

The RHS must now be chosen so that the LHS is independent \( \mu \) and exhibits an absolute maximum when \( \mu = \mu_0 \).

3.3 Assignment Options for \( R \) and \( W \). The fact that \( R \) and \( W \) should remove the dependence of the LHS in inequality (4) on \( \mu \) and thus on the form of \( S(\xi, \mu) \), suggests that \( R \) and \( W \) are acting as Jacobians of the transformations. Recalling conventional cross-correlation implies, if possible, they should map the LHS integrals into the form \( \int_{\Gamma} f_1(x, \mu) f_1(w) dw \). These factors suggest

\[
R(\xi, \mu) = \sqrt{1 + \frac{\partial S(\xi, \mu)}{\partial \xi}} \tag{5}
\]

and

\[
W(\xi, \mu) = R(\xi, \mu).
\]

Hence (12) becomes

\[
\int_{\Gamma} f_1(\xi + S(\xi, \mu))^2 \left( 1 + \frac{\partial S(\xi, \mu)}{\partial \xi} \right) d\xi
+ \int_{\Gamma} f_2(\xi)^2 \left( 1 + \frac{\partial S(\xi, \mu)}{\partial \xi} \right) d\xi
= 2 \int_{\Gamma} f_1(\xi + S(\xi, \mu)) f_2(\xi) \left( 1 + \frac{\partial S(\xi, \mu)}{\partial \xi} \right) d\xi. \tag{6}
\]

The first LHS term is independent of \( \mu \), provided

(A) \( f_1(\xi + S(\xi, \mu)) \) is square integrable over \( \Gamma \),

and

(B) \( f_1(\xi + S(\xi, \mu)) \) is zero outside \( \Gamma \).

Condition (A) and a special case of (B) also apply to conventional cross-correlation. Unfortunately, the second LHS term in (6) varies with \( \mu \). One way to remove this dependence is to reassign \( W \) as

\[
W(\xi, \mu) = \sqrt{1 + \frac{\partial S(\xi, \mu)}{\partial \xi}}. \tag{7}
\]

This means that any GC-C definition will explicitly contain the transformation parameter \( \mu_0 \). As a major application of GC-C is the determination of \( \mu_0 \) such an assignment for \( W \) is generally unrealistic. This option is viable for certain analytic approximations taken at the correlation maximum akin to the perturbation method described in Section 7. However, as the basis for a general definition, the explicit presence of \( \mu_0 \) is unacceptable.

An alternative is to simply take the second LHS term over to the RHS then (6) becomes

\[
\int_{\Gamma} f_1(\xi + S(\xi, \mu))^2 \left( 1 + \frac{\partial S(\xi, \mu)}{\partial \xi} \right) d\xi
\]

\[
\leq 2 \int_{\Gamma} f_1(\xi + S(\xi, \mu)) f_2(\xi) \left( 1 + \frac{\partial S(\xi, \mu)}{\partial \xi} \right) d\xi
- \int_{\Gamma} f_1(\xi + S(\xi, \mu))^2 \left( 1 + \frac{\partial S(\xi, \mu)}{\partial \xi} \right) d\xi. \tag{8}
\]

The inequality now exhibits all the desired features and thus the RHS provides the definition for GC-C:

\[
K_{i,2}(\mu) = 2 \int_{\Gamma} f_1(\xi + S(\xi, \mu)) f_2(\xi) \left( 1 + \frac{\partial S(\xi, \mu)}{\partial \xi} \right) d\xi
- \int_{\Gamma} f_1(\xi + S(\xi, \mu))^2 \left( 1 + \frac{\partial S(\xi, \mu)}{\partial \xi} \right) d\xi. \tag{9}
\]

4 Multidimensional Multiparameter Form

The extension to a multiparameter shift \( S(\xi, \mu_1, \ldots, \mu_n) \) is immediate, i.e.,
Using the analysis in Section 3 shows that satisfies the requirements of a cross-correlation function.

Further generalization of Eq. (10) to accommodate multidimensional functions \(f_1(\xi)\) and \(f_2(\xi)\) is almost as immediate. It is achieved by recalling that \(R^2\) and \(W^2\) behave as Jacobians \(J_\eta\) and \(J_\xi\) in the LHS integrals.

4.1 Multidimensional form of \(J\). In the one-dimensional case the Jacobian is simply \((1 + \partial \xi / \partial \mu)\); however, in \(n\) dimensions it becomes the much more complicated determinant

\[
J = \begin{vmatrix}
\frac{\partial \beta_1}{\partial \xi_1} & \ldots & \frac{\partial \beta_1}{\partial \xi_n} \\
\vdots & \ddots & \vdots \\
\frac{\partial \beta_n}{\partial \xi_1} & \ldots & \frac{\partial \beta_n}{\partial \xi_n}
\end{vmatrix}
\]

which defines a transformation between the variables \(\xi_1, \ldots, \xi_n\) and \(\beta_1, \ldots, \beta_n\). \(J\) is conventionally written in the compact form

\[
J = \frac{\partial (\beta_1, \ldots, \beta_n)}{\partial (\xi_1, \ldots, \xi_n)}.
\]

The variables \(\beta_1, \ldots, \beta_n\) are defined by

\[
\beta_i = \xi_i + S_i(\xi, \mu), \quad i = 1, \ldots, n
\]

where \(S_i\) is the \(i\)th component of the transformation vector \(S\) and the nomenclature \(\mu\) is shorthand for the multiparameter set, i.e., \(\mu\) implies the set \(\mu_1, \ldots, \mu_m\). Consequently in vector notation the expression for an \(n\) dimensional generalized cross-correlation function in \(m\) correlation parameters operating on \(n\) dimensional functions \(f_1(\xi)\) and \(f_2(\xi)\) takes the form

\[
K_{12}(\mu) = 2 \int_{\Gamma_1} \cdots \int_{\Gamma_n} f_1(\xi + S(\xi, \mu)) f_2(\xi) \, d\xi_1 \cdots d\xi_n
\]

\[
- \int_{\Gamma_1} \cdots \int_{\Gamma_n} f_2(\xi + S(\xi, \mu)) f_1(\xi) \, d\xi_1 \cdots d\xi_n
\]

i.e.,

\[
K_{12}(\mu) = 2 \int_{\Gamma_1} \cdots \int_{\Gamma_n} f_1(\xi + S(\xi, \mu)) f_2(\xi) \bigg| \frac{\partial (\xi_1 + S_1, \ldots, \xi_n + S_n)}{\partial (\xi_1, \ldots, \xi_n)} \bigg| \, d\xi_1 \cdots d\xi_n
\]

\[
- \int_{\Gamma_1} \cdots \int_{\Gamma_n} f_2(\xi + S(\xi, \mu)) f_1(\xi) \bigg| \frac{\partial (\xi_1 + S_1, \ldots, \xi_n + S_n)}{\partial (\xi_1, \ldots, \xi_n)} \bigg| \, d\xi_1 \cdots d\xi_n
\]

Using the analysis in Section 3 shows this multidimensional multiparameter form satisfies all the requirements of a cross-correlation function.

4.2 Using \(|J|\) Rather than \(J\). The discussion of the mapping from \(\xi\) to \(\xi + S(\xi, \mu)\) and its generalization in the integrals has been rather informal. In general to effect the required trans-

formation of variables within integrals it is the modulus of \(|J|\) which is needed and thus \(|J|\) will be employed subsequently. However, in most anticipated applications of GC-C it is expected that the transformations of interest are unlikely to be severe and thus \(|J|\) has values close to unity. This means \(|J| = J\) and the modulus requirement can usually be dropped.

5 Two-Dimensional Case

Due to its practical importance the subsequent analysis will concentrate on the two-dimensional case. In which case Eq. (15) becomes

\[
K_{12}(\mu) = 2 \int_{\Gamma_1} \int_{\Gamma_2} f_1(\xi_1 + S_1(\xi_1, \xi_2, \mu), \xi_2) f_2(\xi_1, \xi_2) \bigg| \frac{\partial S_1}{\partial \xi_1} \bigg| \bigg| \frac{\partial S_2}{\partial \xi_2} \bigg| \, d\xi_1 d\xi_2
\]

\[
- \int_{\Gamma_1} \int_{\Gamma_2} f_2(\xi_1 + S_1(\xi_1, \xi_2, \mu), \xi_2) f_1(\xi_1, \xi_2) \bigg| \frac{\partial S_1}{\partial \xi_1} \bigg| \bigg| \frac{\partial S_2}{\partial \xi_2} \bigg| \, d\xi_1 d\xi_2
\]

\[
\frac{\partial S_1}{\partial \xi_1} \frac{\partial S_2}{\partial \xi_2} \bigg| \bigg| \frac{\partial S_1}{\partial \xi_2} \bigg| \bigg| \frac{\partial S_2}{\partial \xi_1} \bigg| \, d\xi_1 d\xi_2
\]

where \(S_1(\xi_1, \xi_2, \mu)\) and \(S_2(\xi_1, \xi_2, \mu)\) are the \(\xi_1\) and \(\xi_2\) components of the displacement vector \(S\).

6 Small Change Approximation

For a variety of well-accepted reasons in flow-field determination the velocities should only induce small changes between \(f_1(\xi)\) and \(f_2(\xi)\) (Kamachi, 1989; Willert and Gharib, 1991, Utami, 1991; Adrian, 1986; Leese, 1971; Ninnis et al., 1986). One of the most obvious reasons is simplicity so that to a good approximation \(S(\xi, \mu_0)\) is proportional to the local velocity \(v(\xi)\) i.e.,

\[
S(\xi, \mu_0) \approx v(\xi) \delta t.
\]

where the parameter \(\delta t\) is the time between images \(f_1\) and \(f_2\).

In any computational procedure aimed at determining the components of \(\mu_0\) the ranges of the \(m\) correlation parameters as represented by the components of \(\mu\), must sensibly be chosen so as to reflect those of the transformation parameters, \(\mu_0\). Therefore, unless extremely pathological circumstances prevail, \(S(\xi, \mu)\) will be of the same order as \(S(\xi, \mu_0)\). Hence not only are \(f_1(\xi)\) and \(f_2(\xi)\) very similar but so are \(f_1(\xi + S(\xi, \mu))\) and \(f_2(\xi)\).

This permits a short Taylor's series approximation of \(f_1(\xi + S(\xi, \mu))\):

\[
f_1(\xi + S(\xi, \mu)) = f_1(\xi) + \frac{\partial f_1}{\partial \xi_1} S_1(\xi, \mu) + \frac{\partial f_1}{\partial \xi_2} S_2(\xi, \mu)
\]

\[
+ \frac{1}{2!} \left[ \frac{\partial^2 f_1}{\partial \xi_1^2} S_1^2(\xi, \mu) + \frac{\partial^2 f_1}{\partial \xi_2^2} S_2^2(\xi, \mu)
\]

\[
+ 2 \frac{\partial f_1}{\partial \xi_1 \partial \xi_2} S_1(\xi, \mu) S_2(\xi, \mu) \right] + O(S^3)
\]

\[
JUNE 1997, Vol. 64 / 323
\]
where \( \mathcal{O}(S^3) \) denotes the order of the error.

Using Eq. (18) in Eq. (16) and retaining terms up to quadratic in the \( S \) functions or their spatial derivatives produces a second-order approximation for \( K_{1,2}(\mu) \). At the maximum of \( K_{1,2}(\mu) \), i.e., at \( K_{1,2}(\mu_0) \), the following system of equations hold:

\[
\frac{\partial K_{1,2}(\mu)}{\partial \mu} \bigg|_{\mu_0} = 0
\]

where \( 1 \leq i \leq m \). This system can be solved for the \( m \) values of the parameters \( \mu_0 \). As will be shown in Section 8 the equation system defined by (19) is linear in the \( \mu_0 \). Determining the \( \mu_0 \) by solving such a system of linear equations is clearly much more efficient than performing a computationally expensive and potentially unreliable multiparameter hill-climb maximum search or optimization processes to estimate these parameters by finding the maximum of \( K_{1,2}(\mu) \). To proceed further it is clearly necessary to define the functional form of \( S(\xi, \mu) \).

6.1 Boundary Conditions on \( f_i(\xi) \). Conditions (A) and (B) in Section 3.5 imply that certain boundary conditions will apply and possibly restrictions exist on the displacement function \( S(\xi, \mu) \). In the small perturbation case on a rectangular domain \([0, L] \times [0, L] \), these issues can be addressed analytically.

After manipulation, including use of Greens Theorem, it can be shown, up to quadratic terms in the Taylor Expansion, that as \( S(\xi, \mu) \to S(\xi, \mu_0) \) the limit \( p_{\xi_0} \to K_{1,2}(\mu) \) is given by

\[
K_{1,2}(\mu) = \int_0^L \int_0^L f_0^2 d\xi_1 d\xi_2
+ \int_0^L \int_0^L [(f_0^2S_{1,0} + f_{1,0}S_{1,0} + 2f_{1,0}S_{1,0}S_{2,0})d\xi_2
- (f_0^2S_{2,0} + f_{2,0}S_{2,0})d\xi_1]
+ \int_0^L \int_0^L [(f_0^2S_{1,0}S_{2,0} + S_{1,0}^2 + S_{2,0}^2)\gamma d\xi_2]
- \int_0^L \int_0^L [(f_0^2S_{1,0}S_{2,0} + S_{1,0}^2 + S_{2,0}^2)\gamma d\xi_1]
\]

\[\mathcal{O}(S^3)\]

where \( C \) is the boundary contour of the domain and \( f \) is equivalent to \( f_i(\xi) \). The symbols \( S_{1,0} \) and \( S_{2,0} \) denote the \( S \) values corresponding to \( \mu = \mu_0 \) and the subscripts, \( \xi_1 \) and \( \xi_2 \), have the conventional meaning of partial differentiation. Equation (3) reveals the boundary conditions upon the various quantities and their spatial derivatives. As would be expected from the lack of any special distinction between the \( \xi_1 \) and \( \xi_2 \)-axes, it is also possible to derive an alternate form of Eq. (20) in which the roles of \( S_1 \) and \( S_2 \) are reversed.

Examination of Eq. (20) and its alternate form shows that conditions (A) and (B) in Section 3.5 are satisfied if any of the following six restrictions hold:

\[\lim_{p_{\xi_0} \to K_{1,2}(\mu)} \int_0^L \int_0^L f_i^2(\xi_1, \xi_2) d\xi_1 d\xi_2\]

**Condition 1.** The most important case for applications is that data function, \( f_i(\xi) \), is zero on the boundary. The ability to impose an appropriate window function on \( f_i \) means that this condition can always be forced if it is not present naturally. It is very important to note that in this case no restrictions arise wrt the displacement function \( S(\xi, \mu) \) and consequently the velocity \( v \).

The need for \( K_{1,2}(\mu) \) to exhibit a maximum for some \( S(\xi, \mu) \) means that a first-order approximation is inadmissible.

\[\text{Condition 2.} \quad \text{If either} \ S_{1,0} \text{or both} \ S_{2,0} \text{are zero on the boundary.}\]

\[\text{Condition 3.} \quad \text{If} \ f_i \text{and both} \ S_{1,0} \text{and} \ S_{2,0} \text{are zero on the boundary.}\]

\[\text{Condition 4.} \quad \text{The alternate form of Eq. (20) shows that conditions 2 and 3 also hold with} \ S_{1,0} \text{replaced by} \ S_{2,0}. \text{This produces the two further conditions.}\]

\[\text{Condition 5.} \quad \text{Various interrelationships between the} \ f_i \text{and the values and derivatives of} \ S_{1,0} \text{and} \ S_{2,0} \text{on the boundaries can also be found; however, these are not expected to be of much practical value.}\]

While these results are formally limited to the case of a rectangular domain, intuition does suggest that this condition will also apply to other shaped regions.

7 A Specific Model for \( S(\xi, \mu) \)

In order to implement \( K_{1,2}(\mu) \) in any particular case, whether in general or as the small change form, a sensible choice must be made for the mathematical form of \( S(\xi, \mu) \). In the present context "a sensible choice" is one which will adequately describe the situation at hand to the required precision using the least number of parameters, with some attention being paid to analytic simplicity.

Many anticipated applications will be for "localized flow process," e.g., flow within the cylinder of an internal combustion engine, or more open problems like cyclonic weather systems. The "rotational" nature of such flows suggests a set of two-dimensional sinusoids of various, i.e., implying a trigonometric polynomial model would appear to be "a sensible choice." Thus

\[S(\xi, \mu) = \sum_{j=\text{Max}}^{\text{Max}} \sum_{r=\text{Max}}^{\text{Max}} \mu_{j,r} e^{i k_j \xi_1 + k_r \xi_2}. \tag{21}\]

The coefficient vectors \( \mu_{j,r} \) have components, \( \mu_{j,0} \) and \( \mu_{0,r} \), which are complex numbers

\[\mu_{0,r} = \mu_{1,1} \hat{k}_1 + \mu_{2,2} \hat{k}_2. \tag{22}\]

and the wave vector \( \mathbf{k}_r \), is

\[\mathbf{k}_r = k_{1} \hat{k}_1 + k_{2} \hat{k}_2. \tag{23}\]

where \( \hat{k}_1 \) and \( \hat{k}_2 \) are the unit basis vectors for \( \mathbf{k}_r \). The unit vectors \( \hat{\mathbf{k}}_1 \) and \( \hat{\mathbf{k}}_2 \) are the basis for \( \mathbf{v} \), these typically but not exclusively correspond to those for \( \xi \).

The trigonometric choice is further endorsed by the work on Proper Orthogonal Decomposition for describing turbulent shear flows (Moser, 1990; Lumley, 1967, 1970). However, such functions may not be parametrically parsimonious in other applications. An obvious example is inelastic deformation problems where orthogonal polynomials normally be a far more efficient choice for \( S(\xi, \mu) \).

Equations (21) and (23) place no restrictions upon the frequencies present and thus \( S(\xi, \mu) \) is an Almost Periodic Function, Bohr (1947, 1968). This is the obvious form to use if there is additional information available concerning the values of the important spatial frequencies present in the problem. However, if these are not known in advance, then it is probably sensible to restrict Eq. (21) and (23) to having integer multiples of some fundamental wavelength in which case

\[\mathbf{k}_{\xi_1, \xi_2} = \frac{2\pi}{L_1} \left( \frac{\xi_1}{L_1} + \frac{\xi_2}{L_2} \right). \tag{24}\]
where \( L_1 \) and \( L_2 \) are the longest wavelength components in the \( \zeta \) and \( \xi \) directions, respectively. In the absence of additional information \( L_1 \) and \( L_2 \) would normally be set to the bounding dimensions of the system; thus making Eq. (21) a conventional two-dimensional Fourier Series.

The word, normally, is italicized to emphasize the importance of incorporating any prior knowledge about the problem of interest into Eq. (21) as this could considerably improve the efficiency of the transformation model in terms of the number of components required. This is the case for the partially forced vortex examined in Part II (Belmont et al., 1997), which has one spatial rotation in the domain of interest and falls to zero towards the walls and the center. Modeling even the main features of such a flow with a conventional Fourier Series based upon a fundamental spatial wavelength \( L \) requires many harmonics. However, making the fundamental \( 2L \), i.e., a half-period over the domain, means that Max = 2 in Eq. (21) is sufficient to describe all the major aspects of this flow.

8 Determination of the Parameters

Determining a set of parameter values which maximize a function is a common computational problem. It can be treated as a hill-climb process or, more indirectly, as an optimization process with Eq. (19) defining the cost function. Typically such procedures are computationally expensive if a large number of parameters must be determined. However, for typical applications when the small change approximation described in Section 8.1, are efficient in conventional cross-correlation work. Thus, benting the Consequences. For a constant shift \( S \) the analysis in Section 8 of Eq. (19) produces a set of integral equations in the real and imaginary parts of \( \mu_{ij} \) and \( \mu_{ji} \). The quantities \( \Theta(p, q) \) are produced from the functions \( f_1(\zeta) \) and \( f_2(\xi) \) via the Finite Interval Fourier Transform shown in Eq. (27). Such integrals are typified by that for \( \Theta_1(p, q) \), which is given by

\[
\Theta_1(-p, -q) = \int_{\Gamma_1} \int_{\Gamma_2} \left( 2 f_1(\xi) f_2(\zeta) - f_1(\zeta) e^{-i2\pi L_2 \xi} \right) d\xi d\zeta \tag{27}
\]

where \( \Gamma_1 \) and \( \Gamma_2 \) denote the domains of \( \zeta \) and \( \xi \), respectively. Using the symbol \( \leftrightarrow \) to denote an integral transform pair the six \( \Theta \) functions are given by

\[
\Theta_1(-p, -q) = \frac{2\partial f_1}{\partial \xi} f_2(\zeta) \tag{29}
\]
\[
\Theta_1(-p, -q) = \frac{2\partial f_2}{\partial \xi} f_1(\zeta) \tag{30}
\]
\[
\Theta_1(-p, -q) = \frac{2\partial f_1}{\partial \zeta} f_2(\xi) \tag{31}
\]
\[
\Theta_1(-p, -q) = \frac{2\partial f_2}{\partial \zeta} f_1(\xi) \tag{32}
\]
\[
\Theta_1(-p, -q) = \frac{2\partial^2 f_1}{\partial \xi^2} f_2(\zeta) \tag{33}
\]

The equation system (25)–(33) uses integer-related sinusoids as this is probably the form of most practical interest. However, with no extra difficulty above the results can be derived in terms of arbitrary noninteger spatial frequencies.

8.2 Relationship to Conventional Cross-Correlation and the Consequences. For a constant shift \( S \) the analysis in Section 3 allows \( K_2(\mu) \) to have the same form as the conventional cross-correlation function. This raises the very tempting speculation as to whether a generalisation of the Wiener Khintchine theorem can be found implying the existence of some very general and potentially powerful new class of integral transforms.

If this were the case then it may open up the route to very efficient evaluation of the \( \mu_0 \) in the same way as spectral routes are efficient in conventional cross-correlation work. Thus, be-

---

This means that Eq. (21) has much of the character of a cosine transform whose origin lies at the center of the domain.
Beyond the purely applications driven viewpoint are much wider mathematical horizons.

Resume and Concluding Comments

The definition of a cross-correlation function has been extended so as to allow the underlying coordinate shift processes operating to vary from point to point over the region of interest. As the parameters in the shift function are computed from global data the usual compromises between resolution and reliability encountered when a region is correlated in a set of patches (Kamachi, 1989; Leese, 1971; Ninnis, 1986) is completely avoided. This issue is of particular importance in view of the central role that patch-based correlation has come to play in techniques like Particle Image Velocimetry (Willert and Gharib, 1991; Utami et al., 1991; Adrian, 1986).

An approximate scheme has been developed which allows efficient calculation of the correlation parameters when the changes in the objects of interest are small. As the vast majority of applications are likely to be of this kind, such an approach is probably the preferred one in practice. The form used for the displacement function in the small change scheme is that of a trigonometric polynomial. This choice was made partly for mathematical utility and partly because such functions tend to be employed any desired set of functions to describe the displacement $\delta(t, \vec{\mu})$. The only real change that occurs is in the form of the integral in Eq. (27) and hence of the functions $\Theta\{k_n, k_{s_n}\}$.

The emphasis here has been mainly upon establishing the basis for Generalized Cross-Correlation Functions, all of the results have been developed in terms of continuous variables. The consequences of using sampled data sets for $f_1$ and $f_2$ are dealt with in Part II, Belmont et al. (1997).

References


Generalized Cross-Correlation, Part II: Discretization of Generalized Cross-Correlation and Progress to Date in Its Implementation

The companion article, "Generalised Cross-Correlation Functions, Part I," introduced a generalization of cross-correlation in which the constant shift used in traditional cross-correlation is replaced by a function of time or space. This allows correlation to be applied globally to the whole domain of interest avoiding the need to compromise spatial resolution with statistical reliability. The development in Part I was entirely in terms of continuous variables. This article examines the issues that arise when Generalized Cross-Correlation is applied to discrete variable situations. Topics discussed include sampling rate requirements, noise rejection, and efficient approximate algorithms, with special attention being paid to the condition number for the system.

1 Introduction

Part I (Belmont and Hotchkiss, 1997), develops the basic theory for generalizing the well-known cross-correlation function. The properties of the Generalized Cross-Correlation function, (abbreviated to GC-C), were discussed together with its advantages in applications. While GC-C has a wide range of uses the specific problems discussed concern the determination of displacement or flow fields, usually from pairs of images of interest will involve sampled data sets. Consequently, it is necessary to examine the effects that discretization has on GC-C and this is the main role of the present article.

In order to avoid duplication the required definitions and results of the continuous variable GC-C theory are all quoted directly from Part I. A convention is adopted that a primed equation number must be read in conjunction with Part I.

It is important to note that examining the effects of discretization GC-C is not the same as considering the details of applications. The focus of attention here is on aspects such as spatial sampling, noise, and numerical issues associated with the reliable determination of parameters. While consideration is given to the effects these factors have on applications this article is not concerned explicitly with actual experimental data. All the numerical studies employ simulation.

The problem of identifying the parameters in the GC-C which characterize the situation of interest is in general terms the same as for any cross-correlation technique, i.e., a search is required for those parameters which maximize the value of the GC-C function. There are many different ways of tackling this problem but they are all computationally intensive. For this reason what is essentially a second-order small perturbation approach to GC-C was developed in Part I that produces a set of linear equations from which the parameters can be determined. This Part II article is mainly concerned with the small perturbation method implementation of GC-C.

The small perturbation approach is very attractive in terms of computing resources. However, the process of solving systems of equations based upon such approximation schemes is notoriously sensitive to the condition number of the system (Golub and Van Loan, 1989). Crudely speaking, for the parameter estimates to be reliable, the product of the condition number \( K \) and the coefficient error bound \( \Delta \) must be small compared to the smallest parameter value (Golub and Van Loan, 1989). As this article focuses upon the small perturbation method the examination of condition number and its effects on the applications forms an important part of Part II.

1.1 The Approach to Considering Discretization. It is unrealistic to attempt a general analytical application of the Nyquist Sampling Theorem directly to the GC-C problem due to the inherent need in GC-C to invert a substantial equation system. Thus the approach adopted here is to explore archetypal cases derived from continuous functions which are sampled so as to generate the required discrete data sets. Where required, controlled synthetic noise is added to such data to examine the noise sensitivity of the technique. Examples of the use of actual experimental data in GC-C will be reported on subsequently.

2 The GC-C Results Needed for Two-Dimensional Implementation

In its most general form GC-C as developed in Part I (Belmont and Hotchkiss, 1997) applies to arbitrary dimensional systems. However, the main interest here will be with two-dimensional cases as these reflect the important practical problem of extracting displacement/flow-field information from sequences of images. One-dimensional GC-C, which is valuable in applications such as the rapid-scan sonar or radar, is simply a special case of the present work and will therefore not be
3 The Spatial Sampling Frequency Required and Number of Parameters Needed

From a purely algebraic standpoint the bare minimum number of spatial samples needed is equal to the number of parameters. However, this view loses sight of physical issues and it is more fruitful to employ the Nyquist Sampling Theorem. This requires that the spatial sampling frequency of all the quantities concerned is at least twice the highest spatial frequency present in the GC-C process, otherwise aliasing errors will be present. Clearly the bandwidth $W_B$ of the displacement functions $S_1$ and $S_2$ must also match that of the displacement/velocity field. This is reflected in the changes brought about in Sections 3.14 and 3.1.4 and hence serves to define the spatial sampling rate as $4 \{W_s + W_f \}$ sample $^{-1}$.

3.2 Differentiation. The equation system coefficients require spatial derivatives of the initial image $f_1$ which are determined analytically in the simulation work. Applications using discrete noise corrupted data require a combination of minimum bandwidth and a band limited differentiating digital filter. The appropriate techniques are standard practice in two dimensional discrete linear systems work and the key elements involved are summarised in the appendix.

4 The Properties of the Equation System

4.1 Spectral Symmetries and Redundancies. The displacement/flow functions $S_1$ and $S_2$ are real, hence the well-known symmetries of Fourier Transform means that only half the $\mu_0$ values are unique. Consequently only half of the $(4/2) (\text{Max } + 1)^2$ equations comprising the system are required to solve for these unique parameters. The other half of the system must be removed otherwise it introduces linear dependencies which lead to singularity.

4.2 The Condition Number $\kappa$ and Sensitivity. The small perturbation GC-C implementation is second-order accurate, hence there are inevitably errors in the system coefficients. Denoting the scale of such coefficient errors by $\Delta$, the contribution of such errors to the solution is magnified by a factor of $\kappa$ (Golub and Van Loan, 1989). Thus it is necessary that the scale $\epsilon$ of $S_1$ and $S_2$ satisfies

$$\epsilon \gg \kappa \Delta.$$  (1)

However, $\epsilon$ must also remain small enough for $f_1$ to be legitimately treated as a perturbation of $f$.

5 Overview of Numerical Work

Section 1.2 indicated that it is unrealistic to develop a general treatment of the discretisation errors in GC-C and hence the approach adopted is to employ continuous functions with known $\mu_0$ parameters to model typical examples. These are then sampled to provide the discrete data. Using this methodology four key issues will be explored:

First, the state of conditioning of the system and its consequences. Closely linked to this is, how precisely can the $\mu_0$ be estimated by the perturbation method? Third, the robustness of the technique is assessed by determining its ability to estimate the $\mu_0$ in the presence of uncorrelated noise. Fourth, does discretisation introduce new boundary condition constraints and any other additional requirements?

5.1 Windowing the Images. To avoid leakage errors it is standard practice to window space or time series data (Brigham, 1988). The window functions employed (e.g., Hamming, 1977) typically have both zero value and derivative at boundaries. In terms of the present coordinates $\xi_1, \xi_2$ over the $L_1, L_2$ domain one of the simplest window functions $w(\xi_1, \xi_2, L_1, L_2)$ is

$$w(\xi_1, \xi_2, L_1, L_2) = \frac{1}{4} \left\{1 - \cos \left(\frac{2\pi \xi_1}{L_1}\right)\right\} \left\{1 - \cos \left(\frac{2\pi \xi_2}{L_2}\right)\right\}.  \quad (2)$$

Any uncorrelated differences between $f_1$ and $f_2$ are treated as noise. In two-dimensional applications such as image work these include out-of-plane displacements.
The distortions introduced by such a window can always be better compensated for, except very close to the boundaries.

Such a function automatically satisfies the boundary requirements on \( f_i \) and it appears that windowing offers a way of making arbitrary systems available to GCC. Unfortunately the corresponding window needed for \( f_i \) must be shifted by \( S \) which is clearly not known in advance. However, by extending the basic window concept, it is possible to employ a development of this approach because the small perturbation GCC also applies to systems where \( f_i \) and \( f_j \) are periodic over the domain. A simple way to make \( f_i \) and \( f_j \) periodic is to analytically continue both functions by mirroring them along the \( z_1 \) and \( z_2 \) boundaries. This does increase the domain size, doubling the fundamental spatial wavelength, but due to the symmetries does not increase the number of unique coefficients needed.

This technique is termed an extension of the windowing technique because the most general way of avoiding leakage and actually making a function periodic over the domain (Brigham, 1988). The traditional method of forcing it to be zero at the boundaries is simply a special case of this.

5.2 The Local Zoom Technique. The use of the analytical continuation form of windowing suggests the possibility of a Local Zoom technique to extract a \( \Delta L \) by \( \Delta L \) region of interest to which GCC-C can be legitimately applied with largest wavelengths scaled to \( \Delta L, \Delta L \). This avoids modeling the displacement/flow field on an unnecessarily fine scale over the whole domain.

This Local Zoom technique has obvious applications as a general purpose tool. However, it also has the potential to be a very powerful method when employed specifically to examine the self-similarity aspects (Mandelbrot, 1983) of complicated facility means that the flow fields can be analyzed at progressively smaller length scales using the same class of spectral techniques. The Zoom general purpose tool. However, it also has the potential to be a very powerful method when employed specifically to examine the small perturbation technique as implemented in discrete form in the absence of any other confounding factors.

The results of error versus size of displacement (the scale of \( S_1 \) and \( S_2 \)) are presented in Fig. 1 (a). The domain size \( L \) was set equal to \( 2\pi \) for convenience and 13 samples were used in each direction which is a little oversampled as the Nyquist requirement as defined in Section 3.1.4 is eight samples. The phase shifts \( \Psi_1 \rightarrow \Psi_2 \) in Eq. (5) were set to \( \pi/4 \) with the displacement scales set equal, i.e., \( a_1 = a_2 = \epsilon \). The ordinate in Fig. 1 (a) shows the percentage RMS error in the estimates of the model coefficients while the abscissa is the displacement scale parameter \( \epsilon \) as a fraction of the total domain size of \( 2\pi \).

The behavior of the condition number \( \kappa \) for the equation system described in Section 4 is shown in Figs. 1 (b) and 1 (c). These plots indicate the variation of \( \kappa \) with displacement scale and also the manner in which the RMS error depends on \( \kappa \).

5.3 Noise-Free Simulations. The simplest and most natural nontrivial example of an image satisfying the boundary conditions for small perturbation GCC-C is where \( f_i \) has the form of the window function in Eq. (2), i.e.,

\[
f_i(z_1, z_2) = \frac{1}{4} \left\{ 1 - \cos \left( \frac{2\pi z_1}{L} \right) \right\} \left\{ 1 - \cos \left( \frac{2\pi z_2}{L} \right) \right\}. \tag{3}
\]

In the examples considered subsequently the domain is taken to be square so \( L_1 = L_2 = L \). Sensible choices for the displacement/flow fields are the simplest forms which exhibit the full range of different types of boundary conditions. For each dimension three distinguishable cases occur: (i) zero value, finite derivative (ii) finite value, zero derivative, i.e., a maximum (iii) finite value and finite derivative. The case where both value and slope are zero is a combination of (ii) with a constant and thus is not treated separately.

The following are the simplest forms that describe the various zero value/slope cases in two dimensions:

\[
S_1 = \epsilon_1 \frac{1}{4} \sin \left( \frac{2\pi z_1}{L} \right) \sin \left( \frac{2\pi z_2}{L} \right) \tag{4}
\]

\[
S_2 = \epsilon_2 \frac{1}{4} \sin \left( \frac{2\pi z_1}{L} \right) \sin \left( \frac{2\pi z_2}{L} \right)
\]

where \( \sin/\cos \) means the different combinations of \( \sin \) or \( \cos \) which give different boundary conditions. There are only five such possibilities as \( \sin \left( \frac{2\pi z_1}{L} \right) \cos \left( \frac{2\pi z_2}{L} \right) \) does not exhibit a different type of behavior to \( \cos \left( \frac{2\pi z_1}{L} \right) \sin \left( \frac{2\pi z_2}{L} \right) \).

The finite value and slope situation is representable by

\[
S_1 = \epsilon_1 \frac{1}{4} \sin \left( \frac{2\pi z_1}{L} + \Psi_1 \right) \sin \left( \frac{2\pi z_2}{L} + \Psi_2 \right) \tag{5}
\]

\[
S_2 = \epsilon_2 \frac{1}{4} \sin \left( \frac{2\pi z_1}{L} + \Psi_3 \right) \sin \left( \frac{2\pi z_2}{L} + \Psi_4 \right)
\]

The attraction of these forms for \( S_1 \) and \( S_2 \) is that they can all be described exactly by the trigonometric models given in Eq. (21) and thus are ideally suited for testing the precision of the small perturbation technique as implemented in discrete form in the absence of any other confounding factors.

The results of error versus size of displacement (the scale of \( S_1 \) and \( S_2 \)) are presented in Fig. 1 (a). The domain size \( L \) was set equal to \( 2\pi \) for convenience and 13 samples were used in each direction which is a little oversampled as the Nyquist requirement as defined in Section 3.1.4 is eight samples. The phase shifts \( \Psi_1 \rightarrow \Psi_2 \) in Eq. (5) were set to \( \pi/4 \) with the displacement scales set equal, i.e., \( a_1 = a_2 = \epsilon \). The ordinate in Fig. 1 (a) shows the percentage RMS error in the estimates of the model coefficients while the abscissa is the displacement scale parameter \( \epsilon \) as a fraction of the total domain size of \( 2\pi \).

The behavior of the condition number \( \kappa \) for the equation system described in Section 4 is shown in Figs. 1 (b) and 1 (c). These plots indicate the variation of \( \kappa \) with displacement scale and also the manner in which the RMS error depends on \( \kappa \).

5.3.1 Discussion of Noise-Free Results. The form of the RMS error is as expected for the small perturbation method, i.e., zero error for infinitely small displacements with an increasing error as the displacement scale increases. There is some variation in precision depending upon the type of displacement boundary conditions but no evidence of any pathological cases. The manner with which the condition number \( \kappa \) varies with displacement scale in Fig. 1 (b) together with the error dependence on \( \kappa \) in Fig. 1 (c) confirms this.

5.3.2 A Pathological Image Form. Unlike the zero slope and value requirement for leakage error reduction, continuous variable GCC-C only needs a zero image value at the boundary. Hence, for completeness it was decided to investigate the behavior of such image forms. The simplest type of \( f_i \) function that exhibits this behavior is

\[
f_i(z) = \sin \left( \frac{2\pi z}{L} \right) \sin \left( \frac{2\pi z}{L} \right). \tag{6}
\]

Clearly this function cannot be a physical image because of the negative intensities present, nonetheless its simplicity justifies its use.

The RMS error and condition number plots paralleling those for the image in Eq. (3) are presented in Figs. 2 (a) and 2 (b). These show markedly different condition number behavior, both in value and form, which are also reflected in very erratic error behavior. Given the approximate nature of the method condition number behavior of this type generally leads to unreliable solution estimates (Golub and Van Loan, 1989).

That condition number is responsible for this pathology, rather than any other type of error, is confirmed by substituting the \( \mu_0 \) parameters into the system and evaluating the right-hand side vector. The equations are then found to be adequately satisfied. Further exploration of this case reveals that the coefficient matrix becomes singular in the limit of small displacements.
Clearly this form of image is pathological and it would be of academic interest to pursue the specific reasons for such behavior in more depth. However, in view of the nonphysical nature of this type of function, such effort is not considered to be justifiable and it is sufficient to highlight the special nature of this type of example.

5.4 Noise Rejection Capabilities of GC-C. A fundamental feature of cross-correlation is its ability to reject noise that is uncorrelated with the data of interest. In order to examine this aspect of GC-C, calculations were performed with wideband noise added to the displacement functions \( S_1 \) and \( S_2 \). This simulated noise in the displacement/velocity field.

The computational procedure was to add wideband random noise with a uniform probability density to the displacement functions \( S_1 \) and \( S_2 \) then to estimate the model parameters and hence the error as in the noise-free case. This procedure was then repeated 100 times and the RMS error over the set of 100 runs was determined. As with the noise-free case, 13 spatial samples were used in each dimension. The results are presented for the most general type of boundary condition i.e., both finite value and derivative. Figure 3(a) shows the results for various sized displacement scale values. The behavior is as expected with the error asymptoting to the noise-free limit in each case. Even for the modest number of samples used the noise rejection is good.

To assess the effects of data corruption during the image capture process noise was added directly to the image \( f' \). Using the same conditions as in Fig. 3(a) with a specific displacement scale fraction of 0.005 the results presented in Fig. 3(b) were obtained. These show that there is very little difference between the effects of noise addition to either the flow field or the image.

Finally, in order to illustrate the effect of sample number size on noise rejection, the 0.005 displacement scale case in Fig. 3(a) was repeated for a modest range of sample values. Figure 3(c) shows that for moderate signal-to-noise ratios, as expected the noise rejection improves as the number of samples increases. The results with a signal to noise ratio of 1 show a large scatter.
which reflects the poor statistical quality obtained at high signal-to-noise ratios when only 100 averages are used.

5.5 Larger Systems and Conditioning Effects. The examples presented above use 16 real parameters and show that reasonable precision can be achieved with condition numbers up to the order of $10^7$. The next obvious question is how the condition number behaves as the required number of parameters increases.

Doubling the spatial resolution for the above systems means that the number of real parameters needed is increased to 32 parameters. This causes the condition number to rise dramatically, typically to several thousand, depending in detail as would be expected on the exact form of $S_1$ and $S_2$. The criterion given in Section 4.2 predicts that this is likely to produce unacceptable errors. Numerical experimentation with a range of displacement functions confirms this to be the case.

5.6 Condition Number Versus System Size. The present findings indicate that the class of linear equations with spectral coefficients that are of interest here have poor conditioning properties as the size of the equation system increases. To probe their behavior more closely some of the coefficients were set to zero allowing an arbitrary variation in the size of the equation system. To probe their behavior more closely some of the coefficients were set to zero allowing an arbitrary variation in the size of the equation system. To probe their behavior more closely some of the coefficients were set to zero allowing an arbitrary variation in the size of the equation system.

5.7 Numerical Techniques. The results presented here employed direct matrix inversion to solve the parameter equation system. Given that Singular Value Decomposition is often found to be valuable where systems have poor condition number behavior, this method was also explored. However, it showed no advantage over direct inversion.

In analytical terms an equation system with $\kappa \gg 1$ behaves as an expensive rather than contracting mapping of the parameter vector, thus it would be expected that substituting the known values of the parameters into the system should produce a good estimate of the right-hand side vector. This was found to be the case confirming that the condition number behavior was directly responsible for the unreliability of the parameter estimates. Given this type of condition number behavior and the fact that evaluation of the RHS vector using the known solutions was acceptably precise, it appears that the most effective numerical technique would be a recursive approximation method. This would exploit the fact that the equation system behaves as a contractive mapping for the forward substitution process involved in such techniques, in contrast to the expansive behavior manifest in attempts to solve, i.e., invert, the system. As in essence the present numerical task is an optimization problem, i.e., maximizing the GC-C function, this suggests using either traditional linear programming methods or the more recently developed genetic algorithms (Goldberg, 1989). The results of such a study will form the basis of a future report.

6 Consequences of Condition Number Behavior on Applications

The main thrust of this work has been to discover the consequences of discretizing CG-C, mainly in its small perturbation form, and not to explicitly consider implementations for applications. However, given the limits which the condition number behavior imposes on the number of parameters which can be reliably calculated, it is clearly necessary to at least consider the consequences of this particular constraint on potential applications.

The main finding of the condition number work is that typically at most 20 parameters in the $S_1$ and $S_2$ displacement functions can be reliably determined with the numerical techniques employed here. For one-dimensional cases such as rapid scanning sonar this restriction is unlikely to present a serious problem. However, until numerical techniques can be determined which circumvent the condition number behavior, the two-dimensional applications of GC-C to experimental data will have to be approached in a much more circumspect manner. The basic requirement is to make the best possible use of the parameters available.
6.1 Systems With Zero Flow/Displacement on the Boundary. The number of length scales that the available number of parameters can represent will obviously depend upon the context. The most sensible situations to examine first are those whose properties motivated the choice of the trigonometric polynomials employed to model \( S_1 \) and \( S_2 \). Such systems are characterized by zero displacement/flow velocity on the boundary and represent problems such as flows in closed regions, cyclonic weather systems, strain in regions clamped at the boundary, etc.

Analytically this boundary condition requires that only the sin terms in \( S_1 \) and \( S_2 \) are nonzero. Thus Eq. (4) in vector form becomes

\[
S(\xi, \mu) = \sum_{r=1}^{\text{Max}} \sum_{\zeta=1}^{\text{Max}} \mu_{\zeta} \sin \left( \frac{2\pi \xi \zeta}{L_{\mu}} \right) \sin \left( \frac{2\pi \xi \zeta}{L_{\nu}} \right),
\]

i.e., the size of the system has fallen by a factor of four. A common feature of such cases is the presence of global rotation which means that the longest lengthscale present is twice the size of the domain. As an example of this case a simulation was made of a very simple model of a cyclonic weather system. The defining equations are

\[
S_1(\xi, \eta) = V_0 \delta \sin \left( \frac{\pi \zeta}{L_{\mu}} \right) \sin \left( \frac{2\pi \xi \eta}{L_{\nu}} \right),
\]

(8)

and

\[
S_2(\xi, \eta) = -V_0 \delta \sin \left( \frac{2\pi \xi \eta}{L_{\nu}} \right) \sin \left( \frac{\pi \eta}{L_{\nu}} \right),
\]

(9)

where \( V_0 \) and \( V_0 \) are the maximum flow speeds in each coordinate and \( \delta \) is the time between images. A vector plot of the flow field produced by Eqs. (8) and (9) is shown in Fig. 4.

The initial image \( f_i(\xi, \eta) \) used is described by Eq. (3) and varying amounts of wideband noise were added to the flow.
Fig. 3(a) The RMS percentage errors in the parameter estimates plotted against signal-to-noise ratio. The displacement field was given by Eq. 5 with \( \psi = \pi / 4 \). Wideband noise was added to the displacement values. The legend indicates the displacement scale. The initial image \( f \) was as defined by Eq. (3) and 13 spatial samples were used per dimension. Results were the average of 100 separate runs.

Fig. 3(b) A comparison of the effects of adding noise to the displacement field with adding noise to the image for the case of a displacement scale of 0.5 percent. The conditions were as for Fig. 3(a).

field. The fractional RMS errors in the estimated coefficients 10 are presented in Table 1.

6.3 Use of GC-C in Conjunction With Cellular Correlation. If approximate values are available for the dominant components of the displacement/flow field then it is possible to select just these in GC-C. This suggests that a combination of conventional cellular cross correlation and currently available numerical implementation of GC-C could provide a powerful tool.

The approach would be as follows:

1. Make an assessment of the displacement/flow field with cellular technique correlation (Kamachi, 1989; Leese et al., 1971; Utami et al., 1991) at a cell size \( L_c \).
2. Determine the spatial spectrum of this estimate and identify the key terms in the displacement/flow field as represented by \( S_1 \) and \( S_2 \) formulation.

If step 2 leads to less than roughly 20 GC-C parameters, then these two steps are repeated with a smaller value of \( L_c \).
3. Perform a GC-C estimation of the chosen parameters.

This results in a global displacement/flow-field model with all the advantages of the GC-C technique.

6.4 Strip Correlation. A halfway house between full GC-C and conventional cellular correlation is to segment the domain into strips of width \( \delta r \) and perform one-dimensional GC-C on each separate strip. This clearly loses the fully two-dimensional coherence of GC-C but does provide long-range coherence along each strip and allows all the parameters to be employed in modeling the behavior along every individual strip.

Clearly \( \delta r \) must be small enough to ensure negligible change in the coordinate normal to the strip and thus \( \delta r \ll \lambda_{\text{max}} \) where \( \lambda_{\text{max}} \) is the smallest spatial wavelength of interest. A sensible check on precision is to perform a second strip GC-C calculation...
along the coordinate normal to that in the first analysis. This provides an overall measure of self-consistency and will also highlight any specific local anomalies.

The strip technique is the simplest current numerical GC-C option to implement and it is felt that until appropriate numerical methods are found for the full two-dimensional GC-C this strip method will probably find the most use in application. As this approach is very obvious and does not require any additional new results, explicit illustrations are not presented here.

7 Conclusions

Discrete implementation has been performed of the small perturbation form of Two-Dimensional Generalized Cross-Correlation (Belmont and Hotchkiss, 1997), where the intention was to recover a spatially varying transformation between successive image-like functions. Typical applications are expected to be the recover of displacement/flow-field data from successive images. For a moderate number of parameters the technique achieves its desired aim and rejects additive noise in the anticipated manner.

For a large number of parameters the conditioning of the equation system used to extract the parameters becomes very poor. The properties of the equation system used to calculate the parameters suggests that this restriction can be overcome by using optimisation rather than inversion-based solution methods. Genetic Algorithms (Goldberg, 1989) appear to be attractive in this respect. The consequences for applications are that most one-dimensional problems such as rapid scan sonar can tackled in a relatively routine way, but that at present two-dimensional applications GC-C must be used in a more thoughtful manner. A direct approach is viable for cases with a few dominant length scales or in conjunction with other methods which provide a preliminary exploration of the problem. A compromise form of GC-C can be set up which does yield a very large number of model parameters in a routine manner. This preserves long-range coherence in one rather than two dimensions.

So far only a trigonometric representation of the displacement/flow-field model version of discretized GC-C has been examined in detail in terms of its condition number behavior. There are a limitless number of other possible representations (Belmont and Hotchkiss, 1997) to explore and given the analytic difficulties it will almost certainly be necessary to investigate each of these in turn.

8 Results Update

The above work employed the lowest possible bandwidth image, i.e., one cycle over the domain. This has recently been
shown to be very much a worst-case situation and findings for wider band images, much more typical of real experimental data, show that the condition number problems are far less for wider band images, much more typical of real experimental than 100 nonzero parameters. The reasons for this are almost certainly because the width of any type of correlation object is a reducing function of the bandwidth of the data. Hence: (i) in this case, a slowly changing image can be approximated by a proportionately wider range of "wrong parameters" than a wider bandwidth image, and (ii) for a given size of shift a much larger change will be introduced in the wideband case. While a reducing function of the bandwidth of the data. Hence: (i) in this case, a slowly changing image can be approximated by a proportionately wider range of "wrong parameters" than a wider bandwidth image, and (ii) for a given size of shift a much larger change will be introduced in the wideband case. While the formal position concerning the current state of development of the Discrete Implementation of GC-C must still remain that described for the worst case those potential users wishing to implement the technique in the nonanalytic form on typical experimentally derived images can expect with some degree of confidence to press the method far beyond the most pessimistic position.

References


APPENDIX

This appendix lists the discrete convolutions needed to compute the Θ[p, q] functions in the equation system coefficients, (28')–(33')

Defining the following Fourier Series coefficients,

\[ P_1[p, q] = f_1(\zeta_1, \zeta_2) \] \hspace{1cm} (A1)

\[ P_2[p, q] = f_2(\zeta_1, \zeta_2) \] \hspace{1cm} (A2)

the terms Θ, to Θn are given by Θ[p, q] as previously stated in (28'), and

\[ Θ_1[p, q] = \frac{2}{N^2} \sum_{r=0}^{N-1} \sum_{s=0}^{N-1} r \Phi[r, s] \times P_1[r, s] P_1[(r-p), (s-q)] \] \hspace{1cm} (A3)

\[ Θ_2[p, q] = \frac{2}{N^2} \sum_{r=0}^{N-1} \sum_{s=0}^{N-1} s \Phi[r, s] \times P_1[r, s] P_2[(r-p), (s-q)] \] \hspace{1cm} (A4)

\[ Θ_3[p, q] = \frac{2}{N^2} \sum_{r=0}^{N-1} \sum_{s=0}^{N-1} s^2 \Phi[r, s] \times P_1[r, s] P_2[(r-p), (s-q)] \] \hspace{1cm} (A5)

\[ Θ_4[p, q] = \frac{2}{N^2} \sum_{r=0}^{N-1} \sum_{s=0}^{N-1} r s \Phi[r, s] \times P_1[r, s] P_2[(r-p), (s-q)] \] \hspace{1cm} (A6)

\[ Θ_5[p, q] = \frac{2}{N^2} \sum_{r=0}^{N-1} \sum_{s=0}^{N-1} r^2 \Phi[r, s] \times P_1[r, s] P_2[(r-p), (s-q)] \] \hspace{1cm} (A7)

where the index N must be optimized in a given application in order to minimize discrete convolution and effect errors, and \( \Phi[r, s] \) is a low pass filter function whose rolloff is set just outside the bandwidth of the Θ[p, q] coefficients. The rolloff of Φ[r, s] must be chosen so as a compromise between minimizing that portion of the differentiator bandwidth beyond that of the Θ[p, q] coefficients while also minimizing Gibbs Phenomenon errors.

This approach obviously entails more computational effort than direct differencing and should only be invoked when significant uncorrelated high-frequency differences, considered as noise, exist between \( f_1 \) and \( f_2 \).