Application of Wavelets to the Modeling, Analysis, Visualization, and Data Handling of Complex Geological Processes

David Yuen[1], Gordon Erlebacher[2], Oleg Vasilyev[3]
Montserrat Fuentes[4] and Dan Goldstein[3;5]

(1) Minnesota Supercomputer Institute, Department of Geophysics, University of Minnesota, Minneapolis, MN, USA, (e-mail: dave yuen@krissy.msi.umn.edu, phone: (612) 624-9801, fax: (612) 625-3819). (2) School of Computational Science and Information Technology, Department of Mathematics/Computer Science, Florida State University, Tallahassee, FL (e-mail: erlebach@csit.fsu.edu, phone: (850) 644-0186, fax: (850) 644- 0098). (3) Department of Mechanical and Aerospace Engineering, Univ. Missouri, Columbia, MO, USA (email: vasilyevo@missouri.edu, (deg67d@missouri.edu) phone: (573) 882-3755, fax: (573) 884-5090). (4) Department of Statistics, North Carolina State University, Raleigh, NC 27695-8203 (email: fuentes@stat.ncsu.edu, phone: (919) 515-1921, fax: (919) 515-1169). (5) (email: deg67d@missouri.edu, phone: (573) 884-2177, fax: (573) 884-5090).

Abstract

Wavelets have a wide range of useful functions that permit them to effectively treat problems such as data compression, scale-localization analysis, feature extraction, visualization, statistics, numerical simulation, and communication. We discuss their features and how they can be used in an integrated manner to handle large-scale problems in earthquake physics and other nonlinear problems in the geosciences.

Introduction

A veritable explosion of data from many fronts in geophysics has precipitated an urgent need for novel tools to process data efficiently and to help better understand the complex dynamical processes buried inside the data. Geophysical processes are intrinsically nonlinear and have a multiple-scale character in both space and time. Wavelets, a recent mathematical innovation, are basis functions localized in both space and frequency, which are thus ideally suited for the modelling and simulation of these types of nonlinear phenomena. They are also very useful for the visualization, and statistical analysis of the data. Our research team, with members from different disciplines, is able to address some of these diverse problems in a unified fashion.

General Properties of Wavelets

Wavelet analysis is a recent numerical concept, which allows one to represent a function in terms of basis functions, localized in both location and scale [1]. One may think of a wavelet decomposition as a multilevel or multiresolution representation of a function, where each level of resolution *j* (except the coarsest one) consists of wavelets ψ_1^j or family of wavelets $\psi_1^{\mu,j}$ having the same scale but located at different positions. The most general wavelet decomposition of a function *f*(x) can be written as

$$f(\mathbf{x}) = \sum_{\mathbf{k}\in\mathcal{K}^0} c_{\mathbf{k}}^0 \phi_{\mathbf{k}}^0(\mathbf{x}) + \sum_{j=0}^{+\infty} \sum_{\mu} \sum_{\mathbf{l}\in\mathcal{L}^{\mu,j}} d_{\mathbf{l}}^{\mu,j} \psi_{\mathbf{l}}^{\mu,j}(\mathbf{x}) \quad (1)$$

where $\phi_k^0(\mathrm{x})$ and $\psi_1^{\mu,j}$ are respectively n-dimensional scaling functions and wavelets of different families and levels of resolution. The major strength of wavelet analysis, i.e. their ability to both compress and de-noise signals, now appears. For functions that contain isolated small scales on a large-scale background, most wavelet coefficients will be small. A good approximation can be retained even after discarding a large number of wavelets with small coefficients. This attractive property of wavelets allows one either to compress or to de-noise the function.

Traditionally, wavelets are constructed by the discrete (typically dyadic) dilation and translation of a single mother wavelet $\psi(x)$. This results in the construction of first generation wavelets [1] that are defined either in infinite or periodic domains. It is desirable in many geophysical applications to have a larger class of wavelets that can be defined in general domains and on irregular sampling intervals. We must now abandon translation/dilation relations of the first generation wavelets and instead construct the wavelets in the physical domain rather than in Fourier space. Recently, a whole new class of wavelets, currently referred to as second generation *wavelets* [7] have come to the fore. The main advantages of the second generation wavelets over the first generation wavelets include the following:

1. Second generation wavelets are constructed in a spatial domain and can be customized for complex multi-dimensional domains and irregular sampling intervals.
2. No auxiliary memory is required and the original signal can be replaced with its wavelet transform.

Second generation wavelets have been utilized recently to construct a dynamically adaptive wavelet collocation method [8] for the solution of both time evolution and elliptic problems. The method employs wavelet compression as an integral part of the solution. The adaptation is achieved by retaining only those wavelets whose coefficients are greater than an a priory prescribed threshold. This property of the multi-level wavelet approximation allows local grid refinement up to an arbitrary small scale without a drastic increase in the number of grid points; thus high resolution computations are carried out only in those regions where sharp transitions occur. With this adaptation strategy, a solution is obtained on a near optimal grid, i.e. the compression of the solution is performed at each timestep.

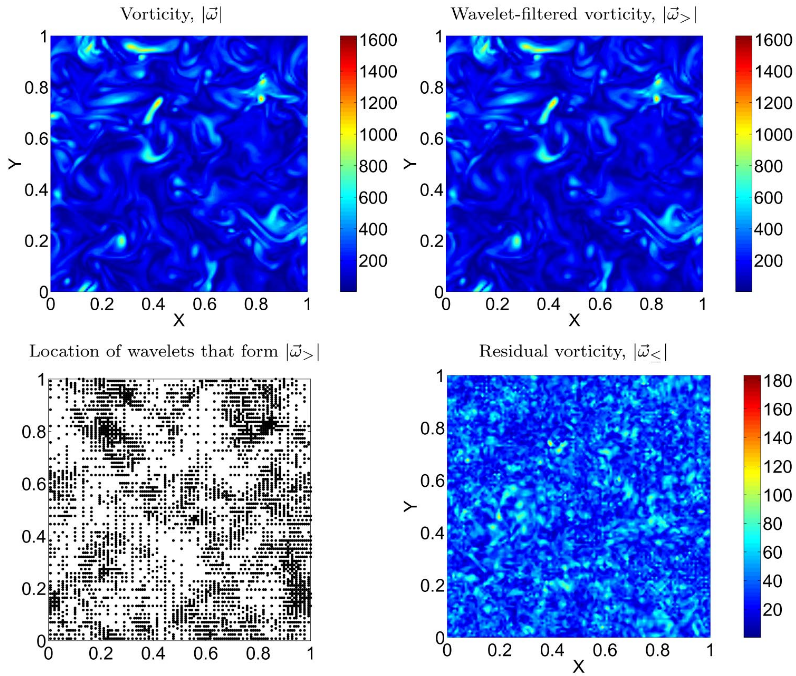

Figure 1: Example of vorticity field decomposition (Eq. (2)) using wavelet thresholding filter for three-dimensional forced isotropic turbulence (two-dimensional slices are shown). The locations of wavelets corresponding to coherent field are also shown. $Re_\lambda = 168$.

Simulation and Modelling

Wavelet De-Noising

The wavelet de-noising procedure, also called wavelet-shrinkage, was originally introduced by Donoho [2]. It can be briefly described as follows: given a function that is the superposition of a smooth function and noise, one performs a forward wavelet transform, and sets to zero the "noisy" wavelet coefficients if the square of the wavelet coefficient is less than the noise variance σ^2. This procedure, known as hard or linear thresholding, is optimal for denoising signals in the presence of Gaussian white noise because wavelet-based estimators minimize the maximal L^2–error for functions with inhomogeneous regularity. In many geophysical applications, the assumption of Gaussian noise is no longer true. In this case, alternative nonlinear thresholding strategies, called soft thresholding, can be utilized. In soft thresholding, the threshold values for wavelet coefficients are scale-dependent.

Decomposition into Coherence and Incoherence Components

The wavelet de-noising property was recently used by Farge et al. [4] to suggest an approach, called Coherent Vortex Simulation (CVS). In CVS the turbulent vorticity field is decomposed into coherent (organized), $\vec{\omega}_{>}$, and incoherent (random, Gaussian), $\vec{\omega}_{\leq}$, fields:

$$\vec{\omega} = \vec{\omega}_{>} + \vec{\omega}_{\leq}\,, \qquad (2)$$

This decomposition is achieved by performing a forward wavelet transform, setting to zero wavelet coefficients whose L^2 or L^∞ norm is below an a priori prescribed threshold parameter ε, which can vary for different levels of resolution, followed by an inverse wavelet transform. An example of the vorticity field decomposition for the three-dimensional forced homogeneous turbulence is shown in Fig. 1, where two- dimensional slices are shown. Figure 1 also shows the locations of wavelets that form $|\vec{\omega}_{>}|$, i.e. whose coefficients are above ε When a non-linear wavelet thresholding filter is applied to a moderately high Reynolds number isotropic turbulence field, the residual field is close to being statistically Gaussian. This has been shown in Farge et al. [4] for two-dimensional turbulent flow and by Goldstein et al. [5] for three dimensional homogeneous turbulent flow.

Statistics over an heterogeneous grid

Classical statistics is based on processes that are stationary and isotropic in the sense that the spatial structure of the flow is independent of location. However, geological processes, such as earthquake dynamics, are inherently time-dependent and spatially heterogeneous. Statistical modelling using wavelets can be used to address this lack of stationarity and spatial homogeneity by expanding the semivariogram of a process in terms of wavelet basis functions [6]. This wavelet representation is motivated by nonstationarity. They also emphasize the computational applicability of the approach for very large systems. We can expand the corresponding covariance function in a wavelet basis, which permits the standard deviation to depend on position. The leading term corresponds to a stationary isotropic model. We study the statistics of the incoherent region to develop models that are input into the evolution equations of the coherent structures.

Feature Extraction, Visualization, Querying

Collocation wavelets set up a one to one correspondence between the points in physical space and the set of wavelet coefficients. Proper identification of the wavelet basis functions that cover the features of interest, both in position and scale, will automatically provide a reduced representation of these features. We have used simple thresholding technique to identify plumes in thermal convection problems. Let I_u be the set of wavelet coefficients with absolute value greater than ε, and let I_p be the corresponding set of grid points in physical space. We have constructed a user interface to Amira, a visualization software package, to allow a user to dynamically construct I_p and plot these points as small cubes. They serve to localize the thermal plumes. Aside from providing new algorithms to compute physical quantitites such as heat flux through the plumes, this identification permits the creation of algorithms to automatically select color maps that generate appropriate volume rendered images. Similar techniques will allow the identification and extraction of faults in earthquake deformation modelling.

As explained earlier, thresholding is an effective means of data compression. Starting from I_u, let $\tilde{I}$ be the full set of wavelet coefficients after setting to zero all those not in I_u. The compression ratio is defined as ord(I_u)= ord(I_∞), where ord is the counting operator. The inverse wavelet transform on $\tilde{I}$ generates a modified image in physical space. Figure 2 compares the plume structure using all the coefficients (left) and one percent of the coefficients (right).

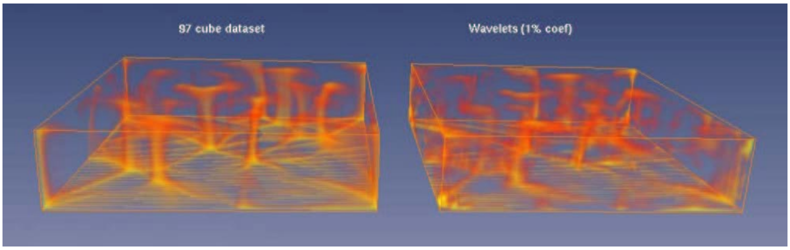

Figure 2: Volume-rendered thermal plumes at $Ra = 10^6$ on 97^3 grid. Original dataset with no compresion (left), 99 percent compression (right).

Well-designed web services, provided by the Grid, should provide users with the ability to query large datasets that reside on geographically distributed installations. These servers act as a remote database for raw and metainformation. Additional features should include the interactive manipulation of this data, either via computations or visualization. We have constructed a prototype system that presents high resolution convection data as a series of “maps"[9]. The data can be queried either in its entirety or in user-seleted subregions. The statistical data is presented in the form of histograms. The software is optimized to send information efficiently over a 802.11b wireless interface. Wavelet technology should bring about more efficient compression and better analysis facilities. This kind of data-query setup can be extremely useful for geoscientists who are working out in the field.

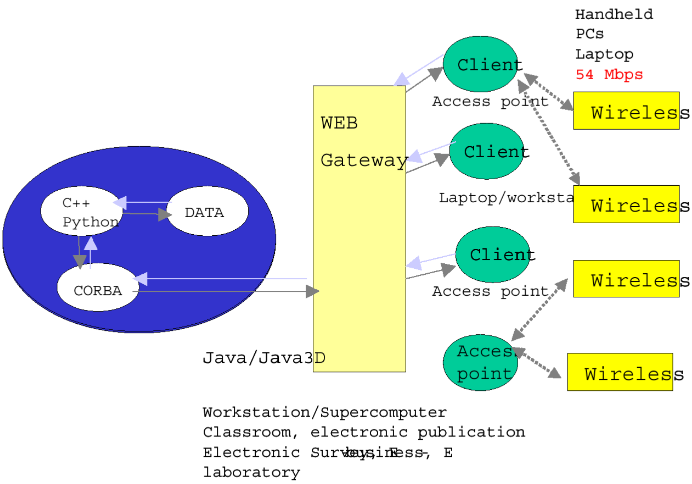

Figure 3: Wireless Image Querying System [9].

Concluding remarks and perspectives

Our experience indicates that second generation wavelets will permit varied applications, such as solution of nonlinear PDE's, visualization and the generation of wavelet-based toolkits. We must also be on the alert for ongoing developments in beamlets and curvelets[3] that are new concepts in multiresolution analysis specialized to the detection and extraction of features with discontinuities lie along curves or surfaces. They might serve to identify as faults and InSar patterns from satellites, and in covering the temporal evolution earthquake deformation.

Acknowledgments

This research has been supported by the Geophysics and CISE programs of the National Science Foundation, the Department of Energy, and the Environmental Protection Agency

References

[1] Daubechies, I., 1988, *Orthonormal bases of compactly supported wavelets*, Comm. Pure and Appl. Math., **41** , 909-996.

[2] Donoho, D.L., 1993, *Unconditional bases are optimal bases for data compression and for statistical estimation*, Appl. Comput. Harmon. Anal., **1**, 100-115.

[3] Donoho, D.L. and Huo, X., 2002, *Beamlets and Multiscale Image Analysis*, in Multiscale and Multiresolution Methods: Theory and Applications, ed. By T.J. Barth et al., 149-196, Springer Verlag, New York.

[4] Farge, M., Schneider, K., and Kevlahan, N., 1999, *Non-Gaussianity and coherent vortex simulation for two-dimensional turbulence using an adaptive orthogonal wavelet basis*, Phys. Fluids, **11**, No. 8, 2187-2201.

[5] Goldstein, D.A., Vasilyev, O.V., Wray, A.A., and Rogallo, R.S., 2000, *Evaluation of the use of second generation wavelets in the coherent vortex simulation approach*, In Proceedings of the 2000 Summer Program, Center for Turbulence Research, 293-304.

[6] Nychka, D., Wikle, C., and Royle, J.A., 1999, *Large spatial prediction problems and nonstationary random fields*, Geophysical Statistical Program, National Center for Atmospheric Research, preprint.

[7] Sweldens, W., 1998, *The lifting scheme: A construction of second generation wavelets*, SIAM J. Math. Anal., **29**, No. 2, 511-546.

[8] Vasilyev, O.V. and Bowman, C., 2000, *Second generation wavelet collocation method for the solution of partial differential equations*, J. Comp. Phys., **165**, 660-693.

[9] Garbow, Z.A., Olson, N.R., Yuen, D., Boggs, J.M., 2001, *Interactive Web-Based Map: Applications to Large Data Sets in the Geosciences*, Electronic Geosciences, **6**, http://link.springer.de/link/service/journals/10069/free/technic/webbased/index.htm