

The University Consortium for Geographic Information Science

Research Priorities



GEOSPATIAL DATA FUSION

THE PRIORITY

Fused sets of geospatial data, which are an increasingly important source of information to support geographic data analysis and modeling and methods to accomplish the fusion or synthesis of data with different geometries (raster, vector), data types (point, line, area), data scaling (nominal, numerical) and varying accuracies and resolutions, require a dedicated research effort over the next five or more years.

DESCRIPTION OF RESEARCH CHALLENGE

Data fusion or the ability to combine various sets of geographic data into a single composite dataset, is a basic tenet of modern geographic information processing. Unfortunately, there has been little research analyzing this process and no theoretical framework exists that can be used to determine the capabilities to combine data of differing types or accuracies. The problems compound when we attempt to combine data with different geometry, resolution, accuracy, type, and statistical scaling. These various combinations

are becoming necessary as data collection technologies and methods expand. Data fusion strategies and methodologies have not kept pace with recent advances in resolution of satellite sensors, radar and lidar technologies, *in situ* sensors using built-in Global Positioning System (GPS) capabilities, and other advances in geographic data collection and processing in geographic information systems (GIS) that are used in increasingly varied applications.

For data models, the difference in the raster and vector geometric models mean inherent integration problems and loss of accuracy between them. However, the problem is greater since within the vector model, topology may be explicitly stored in cellular or algebraic (simplicial) modes or not stored at all and the accuracy of vector lines varies significantly. Within the raster model, square cells (pixels) can be assumed (although new developments in discrete global grids use triangular, hexagonal, and other figures for cells), but the size (resolution) of those cells varies from extremely coarse to very fine. The types of values, categorical (land cover) and numeric or continuous (elevation and image) of the individual pixels

Author:

E. Lynn Usery

Department of Geography
University of Georgia
Athens, GA 30602

E-mail: usery@arches.uga.edu

University Consortium for GIS
Suzy Jampoler, Director
Arthur Getis, President

UCGIS

43351 Spinks Ferry Road
Leesburg, Virginia 20176-5631
TEL: (888) 850-8533
FAX: (703) 771-1635
Internet: <http://www.ucgis.org>

The UCGIS is a non-profit organization of universities and other research institutions dedicated to advancing the understanding of geographic processes and spatial relationships through improved theory, methods, technology, and data.

also affects methods and accuracy of data integration. Finally, the absolute and relative accuracy of the data and the resolution pose significant problems for data fusion. It appears logical that fusion of datasets of extreme differences in resolution and/or accuracy cannot be accomplished effectively, but the research to establish this and to establish the actual limits for fusion has not been performed. Given the plethora of problems and permutations, it is necessary to develop a theoretical framework as a step toward a comprehensive solution.

IMPORTANCE OF RESEARCH CHALLENGE

It has been demonstrated that fusion of geospatial data can provide insights and capabilities not possible with individual datasets. The expanding technology of GIS demonstrates this concept. However, without a conceptual and theoretical framework to provide a basis for data fusion, erroneous results are easily and commonly created. It is important to the United States national interest that this research question be addressed since significant statements concerning our natural resources, environment, urban settlements, and particularly internal or Homeland Security, are dependent on the results of fusion of geospatial datasets such as satellite images and vector-based geographic data including transportation, hydrography, and administrative boundaries.

EMINENT RESEARCH QUESTIONS

Can a theoretical model be developed and verified that provides a basis for fusing geospatial datasets of different geometry, resolution, and accuracy? Can such a model provide a basis for users to automatically combine data through access to metadata that includes resolution and accuracy? Can we develop appropriate methods of handling color and contrast issues when combining multiple raster and/ or raster and vector datasets? Can we develop methods to fuse data with widely varying characteristics and can we establish appropriate limits for the fusion of these datasets based on the methods and the basic data characteristics of geometry, data type and scaling, resolution, and accuracy?

REFERENCES

- Chavez, P.S., Jr. (1986) Digital Merging of Landsat TM and Digitized NHAP Data for 1:24,000-scale Image Mapping. *Photogram-metric Engineering and Remote Sensing* 52: 1637-46
- Cobb, M.A., Chung, M.J., Foley, H., Petry, F.E., Shaw, K.B., and Miller, H.V. (1998) A Rule-based Approach for the Conflation of Attributed Vector Data. *GeoInformatica* 2: 7-36
- Dare, P. and Downman, I. (2001) An Improved Model for Automatic Feature-based Registration of SAR and SPOT Images. *ISPRS Journal of Photogrammetry and Remote Sensing* 56: 13-28
- Jensen, J.R., Saalfeld, A., Broome, F., Cowen, D., Price, K., Ramsey, D., and Lapine, L. (1998) Spatial Data Acquisition and Integration. Washington, D.C., University Consortium for Geographic Science Research White Paper (available at http://www.ucgis.org/research_white/data.html.)
- Schetselaar, E.M. (2001) On Preserving Spectral Balance in Image Fusion and Its Advantages for Geological Interpretation. *Photogrammetric Engineering and Remote Sensing* 67: 925-34
- Welch, R. and Ehlers, M. (1987) Merging Multi-resolution SPOT HRV and Landsat TM Data. *Photogrammetric Engineering and Remote Sensing* 53: 301-3