N82 28705

3.6 REPORT OF THE SUBPANEL ON RESAMPLING FUNCTIONS

The resampling functions subpanel met in two four-hour sessions on November 18 and 19 to define research needs in this area for earth resource observation systems. The panel consisted of:

Richard Juday NASA/JSC	Lee Bender USGS
Arun Prakash General Electric	Scott Cox NASA/GSFC
Allan Mord Ball Aerospace	Robert H. Dye ERIM
Henry Muse E-Systems Inc.	Roger A. Holmes General Motors Institute
Paul Heffner NASA/GSFC	David Dow NASA/NSTL
Carter Glass Goodyear Aerospace	

3.6.1 State of the Art

The state of current practice in resampling appears limited to the use of linear filters, usually cubic convolution on nearest neighbor. Less frequently used alternatives are bilinear and least squares filters. Nearest neighbor is usually the choice when low cost is important or when it is believed that other methods degrade the data. Resampling functions are established on incomplete theory. Cubic convolutions, for example, is based on a sine function but does not employ statistical noise characteristics. There are several different bases for determining the cubic convolution parameters. It is unsettled as to the "optimum" set of parameters.

Several important points related to resampling were brought up. The term "resampling" is not clearly defined. A universally agreed upon "figure of merit" does not exist. Communications between disciplines which enter the resampling problem from different directions and with different terminology is confused. Cubic convolution can not be used to fill image pixel gaps. The TM uses spine weights for this purpose. Nonlinear filter techniques are reported in the literature but do not appear to be used extensively in the civilian remote sensing community.

3.6.2 Anticipated Requirements

Improvements in the use of resampling in applications of remote sensing will require development of a sound theoretical and philosophical basis predicated on argueable cost, error, or benefit criteria. Two reasonably distinct areas of resampling usage are identifiable, with different levels of difficulty in theoretical development. Spatially periodic and complete input lattices of sampled points to be resampled to periodic output lattices of point values describes the more tidy area of resampling. Input lattices which are irregular, including such phenomena as missing points, missing lines, skew from one line subset to another (TM extended line problem or MLA sudden yaw or pitch problem), attitude warp, and bow-tie or other mapping distortions pose a more difficult area with significant overlap with registration and rectification techniques. When such sound bases are established, the applications field will require a portable set of standard resampling codes featuring various known levels of affordability in time and money and well-designed linkages with archival data annotation blocks.

The wide range of potential applications and expanded needs for resampling versatility in new imaging sensors leads to a concern for creating standardized test data sets and ground training sites. The following steps need to be taken: a) We need to develop synthetic data sets related to specific disciplines. b) Data sets based on actual measurements should be tightly controlled both geometrically and radiometrically. Information concerning the creation of the data set should be made universally obtainable. For aircraft acquired scanner data, extensive pre- and postflight performance evaluations should be undertaken. c) Ground training sites used for evaluation of various sensor types and processing functions should be extensively mapped and photographed. Updates at reasonable intervals are mandatory. d) Experiments involving registration and rectification require precise geometric and geodetic control, sites should be surveyed to present cadastral standards. Ancillary recorded data such as DTM, soil type, etc. should be made available at a grid size smaller than the expected IFOV or spaceborne sensors. One order of magnitude smaller is desirable.

With respect to resampling algorithms, the possibility exists to construct a library of tested algorithms available from NASA/Cosmic or an alternate source. In pursuit of this goal, standards need to be developed with respect to: a) transportability, b) documentation (should include the mathematical basis of the algorithm as well as a thorough description of the algorithm itself), c) test data sets, and d) standardized performance specifications.

3.6.3 Recommended Research

Seven research areas were identified by the panel and are described below.

I. Perform a study research task to develop a theory for defining an input psf of an instrument such that upon resampling of the processed data into the desired output grid the radiometric and geometric properties are a best estimate of the true upwelling radiances. Such a theory would include noise, desired output psf, and application specific best estimate parameter.

II. Develop algorithms for estimating missing data or for ignoring missing data. Such missing data sources to be considered are: a) dead detectors in multiscanners (including MLSs), b) scanner sweep gaps, c) transmission dropouts. Such algorithms should be characterized by their effect on the output image such that the choice of procedure can be made by the user for the application being considered.

III. Perform a study on optimal resampling functions. These optimal functions may be application specific or determined for: a) radiometric accuracy,

b) geometric fidelity, c) spatial feature enhancement and extraction. Questions to be addressed include the nature of data characteristics which permit such optimality and what kind of trade-offs are made between the different kinds of optimal filter and under what conditions can an optimal filter be designed to be optimal for both radiometric accuracy and geometric fidelity.

IV. Standard Data Sets. The choice of a resampling function is problem dependent. For a complex problem, the specific effects of a given resampling function cannot be predicted by pencil-and-paper methods. As a result, a two-stage simulation with standard data sets is recommended. Both real and synthetic data sets are required because there are really two functions to be performed. The synthetic data set can be used to study the behavior of various resamplers on controlled data. We can derive various figures of merit for comparing filter performance and add controlled amounts of noise to study noise-rejection performance. Synthetic data exhibits the following properties: (a) ability to control dimensionality, (b) power spectral density, (c) type and magnitude of noise, and (d) features, etc. An authentic data set is needed to give the final level of credibility to the resampler performance. Real data is essential to uncover phenomena which cannot be anticipated by artifically contrived data. Real data should represent the problem ("Typical" scene or even "worst-cast" scene). The synthetic data set lends itself well to determining quantitative, controlled performance measures. The real-data set lends itself to qualitative (visual) comparison and discovery of gross processing artifacts.

V. The least squares resampler developed by Dye should be further developed and tested. Cost benefit analyses must also be done and clear definitions of of applicability and limitations.

VI. There is a need for merit functions for assessing resampling techniques. Various numbers and/or functions have been proposed as measures of goodness of resampling operations. Confusion exists regarding the relationships between them and what they measure. People expect that different end users will need to optimize different figures of merit, but aren't quite sure which. The research approach would: a) List and define all candidate merit functions. Perform theoretical analysis to find the relationship between them and what information they convey. b) Survey users in various disciplines to see which merit functions look most useful to each. c) Compute all promising merit functions for one or two examples of resampling. Focus on exploring the merit functions, not the resampling.

VII. Develop a benign resampling algorithm. Present "A to P" conversion for Landsat D imagery introduces problems in the "P" data that makes some users undertake the conversion themselves, or wish they could. Modern pipeline processors may make some sort of A to P conversion strongly attractive, provided that a kernal can be found which does not damage the data unacceptably. Research should be undertaken to analyze resampling althorithms to find ones which retain nearly all of the quality of the data originally collected while doing rubber sheet resampling plus filling small fixed gaps. Characterize it in terms of suitable merit functions. Design algorithms to be implemented in fast hardwired processor. Demonstrate feasibility on real data.