

Vector-Ordering Filter Procedure for Data Reduction

The essential characteristics of original large sets of data are preserved.

Marshall Space Flight Center, Alabama

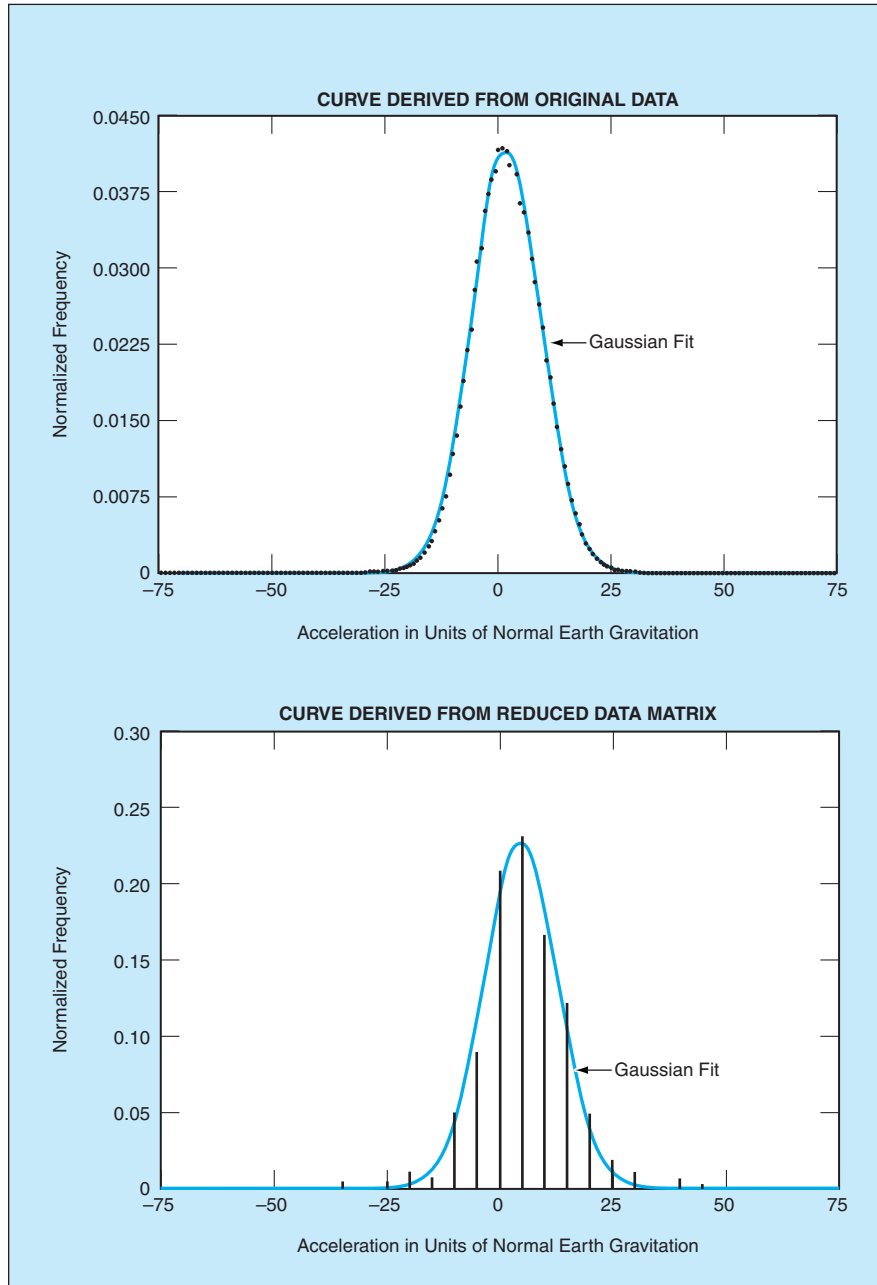
The vector-ordering filter (VOF) technique involves a procedure for sampling a large population of data vectors to select a subset of data vectors that fully characterize the state space of the large population. The VOF technique enables a large reduction of the volume of data that must be handled in the automated-

monitoring system and method discussed in the two immediately preceding articles. In so doing, the VOF technique enables the development of data-driven mathematical models of a monitored asset from sets of data that would otherwise exceed the memory capacities of conventional engineering computers.

Data-driven mathematical models have been shown to offer high fidelity for purposes of control and monitoring of assets. In practice, a collection of asset-operating observations is acquired with the intention that the collection contain observations characteristic of the full dynamic range of operation of the asset. Often, such a collection contains an extremely large number of observations, many of which are redundant. The VOF technique fills the need for a means to extract, from the original collection of observational data, a reduced data matrix that excludes redundant data while maintaining the full statistical character and dynamic range of the original data. The reduced data matrix can then be used as the input data for development of a mathematical model of the monitored asset, or as training data for a neural-network substitute for an explicit mathematical model of the asset. Alternatively, the reduced data matrix can, itself, be used directly as a mathematical model of the monitored asset, as is commonly done in multivariate state-estimation techniques.

The original data are collected from the asset over a range of operating states and are put in matrix form. Each column vector in the original data matrix represents the signal values acquired at a particular operational state of the asset. Thus, the number of columns of the original data matrix equals the number of observed states and the number of rows in this matrix equals the number of signals acquired at each observation. In the VOF technique, one extracts the reduced data matrix from the original data matrix through the selection of a representative subset of the column (state) vectors.

In its simplest form, the VOF procedure is a two-stage procedure, in the first stage of which one selects those data vectors that characterize the extrema present in the original data. At the beginning of the first stage, one finds the extrema of each signal as represented by the minimum and maximum values in the corresponding row of the original data matrix. The column vectors that contain these values are selected as candidates for inclusion in the reduced data matrix. Before a candidate column vector is added to the reduced data matrix,



These **Plots Characterize Data** from an original data matrix and a reduced data matrix derived from it in two VOF passes. The original set of data represented 62,914,252 observations, each observation comprising readings of 6 accelerometers on the space shuttle main engine. The first VOF pass, made over 394 subsets of the original data, yielded a reduced data matrix of 160,358 vectors. The second pass extracted a final reduced data matrix containing 250 vectors.

it is compared with those column vectors already in the reduced data matrix to ensure that only one copy of that vector ends up in the reduced data matrix. In the second stage, one orders the column vectors of the original data matrix by their Euclidean norms and then selects a subset of the vectors according to a spacing criterion. All column vectors selected in this way are compared to those vectors selected during the first stage. Only those column vectors that were not already included in the reduced data

matrix during the first stage are added to the reduced data matrix.

Practical sets of data tend to be so large that excessive computer memory would be necessary for a single pass of the two-stage VOF procedure over all the data. In such a case, the VOF procedure can be applied recursively to successive subsets of the original data that are small enough to fit in the available memory. The figure presents plots from an example of a two-pass application of the VOF technique to

some space-shuttle engine vibration data.

This work was done by Randall L. Bickford of Expert Microsystems, Inc., and James P. Herzog of Argonne National Laboratory for Marshall Space Flight Center. This technology is immediately available using the SureSense™ Signal Validation System software produced by Expert Microsystems, Inc. For more information, contact the company at (916) 989-2018 or at expert@expmicrosys.com. MFS-31588

Remote Sensing and Information Technology for Large Farms

Timely data on spatial and temporal variations in fields help farmers manage crops.

Stennis Space Center, Mississippi

A method of applying remote sensing (RS) and information-management technology to help large farms produce at maximum efficiency is undergoing development. The novelty of the method does not lie in the concept of "precision agriculture," which involves variation of seeding, of application of chemicals, and of irrigation according to the spatially and temporally local variations in the growth stages and health of crops and in the chemical and physical conditions of soils. The novelty also does not lie in the use of RS data registered with other data in a geographic information system (GIS) to guide the use of precise agricultural techniques. Instead, the novelty lies in a systematic approach to overcoming obstacles that, heretofore, have impeded the timely distribution of reliable, relevant, and sufficient GIS data to support day-to-day, acre-to-acre decisions concerning the application of precise agricultural techniques to in-

crease production and decrease cost.

The development and promotion of the method are inspired in part by a vision of equipping farm machinery to accept GIS (including RS) data and using the data for automated or semi-automated implementation of precise agricultural techniques. Primary examples of relevant GIS data include information on plant stress, soil moisture, and effects of applied chemicals, all derived by automated computational analysis of measurements taken by one or more airborne spectroradiometers.

Proper management and timeliness of the large amount of GIS information are of paramount concern in agriculture. Information on stresses and changes in crops is especially perishable and important to farmers. The need for timeliness and management of information is satisfied by use of computing hardware and software capable of (1) rapid geo-rectification and other pro-

cessing of RS data, (2) packaging the output data in the form of GIS plots, and (3) making the data available to farmers and other subscribers by Internet password access. It is a goal of this development program to make RS data available no later than the data after an aerial survey. In addition, data from prior surveys are kept in the data base. Farmers can, for example, use current and prior data to analyze changes.

This work was done by John E. Williams of Global Positioning Solutions, Inc., and Jimmie A. Ramsay of DataStar, Inc. for Stennis Space Center.

In accordance with Public Law 96-517, the contractor has elected to retain title to this invention. Inquiries concerning rights for its commercial use should be addressed to:

*Global Positioning Solutions, Inc.
P.O. Box 89*

Inverness, MS 38753

Refer to SSC-00150, volume and number of this NASA Tech Briefs issue, and the page number.