



Intelligent Software Tools for Advanced Computing

Federal Manufacturing & Technologies

C. W. Baumgart

KCP-613-6424

Published April 2001

Final Report

Approved for public release; distribution is unlimited.



Prepared under prime contract DE-ACO4-01AL66850 for the

United States Department of Energy

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or any third party's use or the results of such use of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product,

process, or service by trade names, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof or its contractors or subcontractors. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Printed in the United States of America.

This report has been reproduced from the best available copy.

Available to DOE and DOE contractors from the Office of Scientific and Technical Information, P. O. Box 62, Oak Ridge, Tennessee 37831; prices available from (865) 576-8401, FTS 626-8401 Facsimile: (865) 576-5728, E-mail: reports@adonis.osti.gov

Available to the public from the National Technical Information Service, U. S. Department of Commerce, 5285 Port Royal Rd., Springfield, Virginia 22161, (800) 553-6847, Facsimile: (703) 605-6900, E-mail: orders@ntis.fedworld.gov

Honeywell

Federal Manufacturing
& Technologies

P. O. Box 419159

Kansas City, Missouri

64141-6159

A prime contractor with the United States

Department of Energy under prime contract

DE-ACO4-01AL66850.

Honeywell

KCP-613-6424

Distribution Category UC-706

Approved for public release; distribution is unlimited.

INTELLIGENT SOFTWARE TOOLS FOR ADVANCED COMPUTING

C. W. Baumgart

Published April 2001

Final Report

C. W. Baumgart, Project Leader

Project Team:

K. D. Linder

Contents

Section

Abstract

Summary

Discussion

Scope and Purpose

Activity

FET Software Design

Functional Descriptions

Accomplishments

Future Work

Illustrations

Figure

1 Flowchart of Feature Extraction Toolbox

2 Example of Feature Extraction Toolbox Pull-down Menu Interface

3 Feature Extraction and Feature Evaluation Pull-down Menu Options

Abstract

Feature extraction and evaluation are two procedures common to the development of any pattern recognition application. These features are the primary pieces of information which are used to train the pattern recognition tool, whether that tool is a neural network, a fuzzy logic rulebase, or a genetic algorithm. Careful selection of the features to be used by the pattern recognition tool can significantly streamline the overall development and training of the solution for the pattern recognition application. This report summarizes the development of an integrated, computer-based software package called the Feature Extraction Toolbox (FET), which can be used for the development and deployment of solutions to generic pattern recognition problems. This toolbox integrates a number of software techniques for signal processing, feature extraction and evaluation, and pattern recognition, all under a single, user-friendly development environment. The toolbox has been developed to run on a laptop computer, so that it may be taken to a site and used to develop pattern recognition applications in the field. A prototype version of this toolbox has been completed and is currently being used for applications development on several projects in support of the Department of Energy.

Summary

The processing, analysis, and interpretation of data collected "in the field" often requires feature extraction and evaluation to be performed as a precursor to the development of a pattern recognition solution. Additionally, the analysis must often be performed quickly, and it is beneficial to perform the actual development in the field. Since pattern recognition solutions are routinely developed in support of the DOE mission, this project was initiated to develop a prototype version of the Feature Extraction Toolbox (FET) for use in the development of solutions to pattern recognition problems. The basic focus of this toolbox is to provide the user (developer) with a tool to repeatedly and iteratively extract and evaluate features from any given set of data. This capability allows the user to identify and select the key pieces of information (features) from the raw data that will be most significant and useful to the development of the pattern recognition solution.

One of the most useful features of the FET is its portability. Its flexible design allows the program to be loaded on a laptop and taken to the data acquisition site. At the site, data may be input to the computer and evaluated in-situ. A flexible data input routine allows a wide variety of data formats to be interpreted and processed. By performing the feature extraction, evaluation, and pattern recognition development steps in the field, feedback on the performance of the prototype pattern recognition solution is immediately provided to the developer and the customer. Such feedback information allows the developer to identify undesirable systematic effects either in the data or in the customer's data acquisition equipment. Likewise, this feedback information can aid in the fine-tuning of system parameters which are beneficial to system performance.

Another key component to the design of the FET is its ability to accommodate a wide range of generic engineering, pattern recognition applications. To meet this need, the software architecture is very modular in design, which allows new techniques and methods to be easily integrated into the prototype. The prototype was developed on a PC-based 32-bit Windows platform with Microsoft Visual Basic 5.0 used to develop the graphical user interface (GUI) and with Microsoft Visual C/C++ and Digital Visual FORTRAN used for the development of the signal processing and analysis modules. A number of third

party DLLs have also been incorporated to aid graphical data display.

The prototype FET may be applied to a wide variety of signal processing, image processing, feature extraction, and pattern recognition applications. Such applications may include

- Signal and Image Analysis
 - analysis of gas chromatography and mass spectrometer data
 - edge detection and image feature extraction
 - morphologically-based object recognition
- Pattern Recognition
 1. anomaly detection
 2. analysis of geophysical sensor data (ground penetrating radar, magnetometer, electromagnetic induction)
 3. analysis of chemical sensor data
 4. data fusion (multi-sensor and multi-source)
- Process Control
 1. nuclear material control and accountability
 2. manufacturing
 3. quality
- Process Monitoring
 1. nuclear material control and accountability
 2. data fusion (multi-sensor and multi-source)

The remainder of this report describes the functionality of the FET in more detail.

Discussion

Scope and Purpose

Feature extraction and evaluation are two procedures common to the development of any pattern recognition application. These features are the primary pieces of information which are used to train the pattern recognition tool, whether that tool is a neural network, a fuzzy logic rulebase, or a genetic algorithm. Careful selection of the features to be used by the pattern recognition tool can significantly streamline the overall development and training of the solution for the pattern recognition application. This report summarizes the development of an integrated, computer-based software package, called the Feature Extraction Toolbox (FET), which can be used for the development and deployment of solutions to generic pattern recognition problems. This toolbox integrates a number of software techniques for signal processing, feature extraction and evaluation, and pattern recognition, all under a single, user-friendly development environment. The toolbox has been developed to run on a laptop computer, so that it may be taken to a site and used to develop pattern recognition applications in the field. A

prototype version of this toolbox has been completed and is currently being used for applications development on several projects in support of the Department of Energy.

Activity

The FET is designed to aid the software engineer in the development of a pattern recognition solution for a given application. Once the basic data to be processed is identified, and the goal of the recognition task is defined, the FET may be used to design, test, and implement solutions to the problem.

FET Software Design

The FET design is broken down into six main modules that mirror the process used to develop a pattern recognition application. These steps are

1. Input and format the raw acquired data
2. View the raw data
3. Extract features
4. Evaluate features
5. Select the final feature set
6. Apply features using pattern recognition techniques to develop training and test data sets for development of the pattern recognition solution.

After the raw data is read into the system, a menu of feature extraction techniques is presented to assist the user in making a determination of which features will be used in the pattern recognition tool. A number of different feature extraction techniques are provided to the user, who may iterate the feature selection and evaluation steps many times in an effort to discover the optimal feature set. The quality of a selected set of features may be evaluated at each iteration with one or more statistical analysis tools. Using these tools, the user may assess which features in the selected test set appear to correlate significantly across the raw data and which features provide no discriminatory power at all. Once a "best" set of features is chosen, the selected features are used to extract training and test data sets to be used to train and evaluate the performance of the pattern recognition tool chosen for the application (neural network, fuzzy logic reasoning system, expert system, etc.). Currently a back propagation neural network is the only pattern recognition tool installed in the prototype. A user interface is provided to aid the user in defining and optimizing the neural network parameters.

A flowchart of the various functions included within the FET is shown in Figure 1. While each box in the flowchart shows the major functions, the bullets within each function are the modular components. Additional modular components (feature extraction and evaluation techniques) may be easily added to fit the needs of a particular application or a particular user.

Figure 1. Flowchart of Feature Extraction Toolbox

Functional Descriptions

The operation of each of the main functional modules in the FET is described below. Due to the modular

nature of the FET design, the functionality of each of the main modules may be easily extended by adding application-specific analysis techniques and tools.

Data Input Function

The data input to the program may be either the raw data (as acquired for the particular application), features previously extracted from the raw data (either from earlier toolbox execution or other methods), or neural network training data. While the raw data may be in a variety of file formats, the feature files and training files must be in a specific format. These selections are made via a pull-down menu as shown in Figure 2.

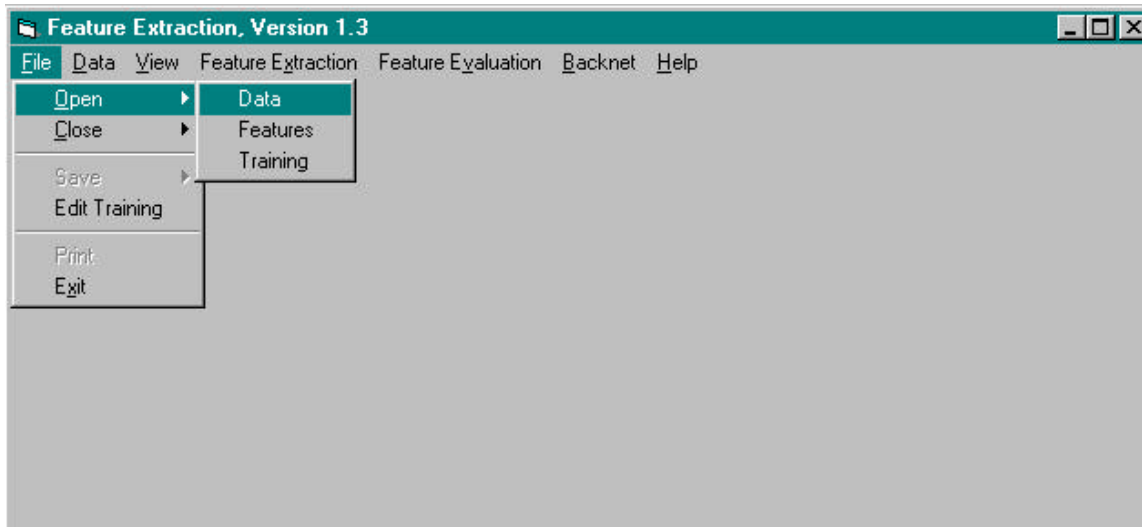


Figure 2. Example of Feature Extraction Toolbox Pull-down Menu Interface

View Data Function

Data display formats range from simple text display to two-dimensional image plots. Additionally, after the various feature analysis methods have been applied to the raw data, the feature extraction results may also be viewed in various formats. Currently available viewing capabilities include text, line plot, two-dimensional image, and spreadsheets.

Feature Extraction Function

At the heart of this package are the feature extraction and evaluation capabilities. Features are first selected and extracted from the raw data, and then evaluated by the user. Currently available features for extraction include moments, minimum and maximum amplitude values, Fast Fourier Transforms, and the Hilbert Transform. Other techniques are also available and any additional technique may be easily incorporated into the FET via a DLL (or other such library) and corresponding menu modification. After selecting the techniques of interest via the *Feature Extraction* pull-down menu, the *Extract* option instructs the software to perform the selected feature extraction operations. Figure 3 shows the *Feature Extraction* pull-down menu.

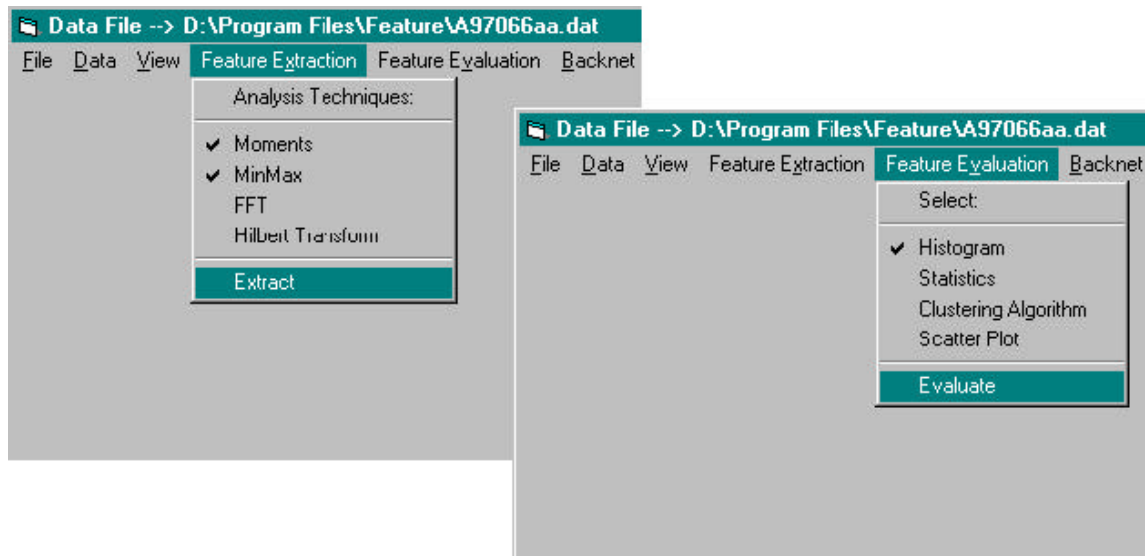


Figure 3 . *Feature Extraction* and *Feature Evaluation* Pull-down Menu Options

Feature Evaluation Function

Once the desired feature extraction methods are selected, the *Feature Evaluation* option instructs the software to perform the selected evaluation operations. Feature evaluation is an important aspect of selecting an optimal set of features for a given application. The feature evaluation tools provided in the FET allow the user to test each extracted feature for statistical significance over the entire data set. In this way, features with greater discriminatory capability may be identified and selected for use in the application. The feature evaluation tools currently included in the FET are histogram analysis, statistical analysis, clustering, and scatter plots.

Feature Selection Function

This process of extracting features and evaluating them may be performed multiple times in a given toolbox session. This is often the process required to identify an optimal set of features. In doing so, the user is able to easily extract a wide variety of features and visually inspect them to see how useful they are in obtaining the most effective set of features. Once the optimal features have been identified, they may be written to a feature file or fed directly into a neural network.

Pattern Recognition Function

Currently the only pattern recognition tool integrated into the toolbox is a back propagation neural network. After selecting *Backnet* from the pull-down menu, there are two options: *Train* and *Classify*. Training the network takes place with the current features (just read in or just calculated), and classifying the data takes place using previously trained (and stored) network parameters. Additional pattern recognition tools may be added to include Statistical Regression, Functional Link Networks, Genetic Algorithms, and Adaptive Resonance Theory.

Accomplishments

The goal of this project was to design and develop a prototype of a portable software system for pattern recognition application development which can easily be used on a wide variety of engineering applications. A prototype software package has been developed and tested. Both in-house and commercial packages may be integrated into the existing FET processing architecture. This tool has been designed and successfully demonstrated on numerous projects. In addition, the FET may be easily customized to any particular application so that additional projects should benefit from this work in the future.

Future Work

The Feature Extraction Toolbox provides the software engineer with a variety of feature extraction and evaluation tools for use in the development of solutions to pattern recognition problems. Future work with the FET will primarily consist of the use of the FET for application development efforts. Additional analysis capability and functionality will be added as the FET is applied to develop solutions to a variety of pattern recognition problems. Several projects are currently making use of the FET to aid the application development process.