CDI SSF Category: Computational Tools and Services

Development of Enhanced Feature Recognition Software for the Extraction of Mine Features from USGS Topographic Maps

Applicants/Principal Investigators:

Greg Fernette, USGS Central Mineral and Environmental Resources Science Center, Denver Federal Center MS 973, Denver, CO 80225. Ph (303) 236-5611 Email *gfernette@usgs.gov*

Peter Schweitzer, USGS National Center, MS 954, Reston, VA 20192. Ph (702) 648-6533 Email *pschweitzer@usgs.gov*

Abstract:

The objective of this proposal is to provide an automated feature recognition tool to transform non-integrable raster-based information – mine feature symbols on digital topographic maps -- into data that can be effectively integrated with other spatial data layers. An overarching purpose of the CDI includes integrating and facilitating access to scientific data and information, but inherent in the CDI strategic vision is the obvious first order requirement that data *be* integrable. There is an extensive trove of USGS spatial information that resides as raster images which cannot be queried or selectively identified, attributed, discriminated (except visually), edited or used in analysis. In other words, those data are not available for integration and therefore have limited use to the scientific community. The process of manually digitizing those data is, for large areas, exceptionally time consuming and extremely costly. Moreoever, there are historical (but non-integrable) data that could be used to compare temporally separate, but spatially overlapping features, but the effort required to capture those data is formidable if not prohibitive. This proposal to automate feature extraction from raster data to produce integrable geospatial data will serve the goals of the CDI charter, will leverage the extraordinarily valuable USGS topographic map archive, and will provided immediate benefit to a broad scientific community. There are far reaching effects anticipated by the proposed work and it is anticipated that it could serve to generate nationwide data that without development of these tools, will remain effectively "buried" in raster depictions, and thereby unavailable for integration.

Total funding amount requested: $104,030
Total in-kind funding: $146,015

Datasets:
USGS Historic Topographic Map Collection (HTMC), USGS Mineral Deposit Database

Geographic/Geologic/Ecosystem/Habitat/Taxonomic/Other Context:
US national**,** all geologic provinces, ecosystems and habitats **-** particular relevance to mineral resource assessment, mine land reclamation, land management

Type of Products(s) Generated:
Software, geospatial data

Summary

a) Introduction and background

USGS topographic maps are a significant repository of high quality geospatial information. Each map represents a "time slice" of information on such things as vegetation distribution, wetlands, springs, drainages and manmade features which is of particular value as it pre-dates satellite imagery. USGS recently completed the Historical Topographic Map Collection (HTMC) in which more than 175,000 topographic maps were scanned and georeferenced. While the HTMC greatly increased the accessibility of the maps, the information on the maps is available only by visual inspection and on a map by map basis.

The Mineral Deposit Database Project (USMIN) is developing a geospatial, mineral deposit database for the United States building upon the long term effort of the Mineral Resource Data System (MRDS). Currently, USMIN is capturing mine features from the topographic maps by manual digitization. Needless to say this is a massive, time consuming and labor intensive process. For example, in the state of Nevada alone there are more than 1,900 7.5-minute quadrangle maps with over 105,000 mine features. The ultimate objective of USMIN is to capture this data for the entire United States.

In this proposal we are requesting funding to develop an automated method of capturing mine features from scanned topographic maps via enhanced feature recognition software. The feasibility of this approach was demonstrated during the HTMC when software, known as QUAD-G was developed which automated the recognition of longitude-latitude tick marks and georeferencing of maps. Discussions with Greg Allord who led the HTMC and Dr. James Burt who developed QUAD-G indicate that this approach can be adapted to the recognition and extraction of other features from topographic maps. The methodology would not only address the existing data capture and data integration problem, but might also be adapted to other applications such as the extraction of wetland areas, springs, etc. As with the prior HTMC project, resulting software products will be made publically available.

b) CDI SSF Category 2 - Computational Tools and Services. The project aims to develop and test the software application and user procedures.

c) Project Title: Development of enhanced feature recognition software for the extraction of mine features from USGS topographic maps.

d) Project Leaders or Principal Investigators.

Greg Fernette - Research Geologist
US Geological Survey
Denver Federal Center Box 25046 Mail Stop 973
Denver, CO 80225-0046
Voice (303) 236-5611
Email: gfernette@usgs.gov

Peter N. Schweitzer - Geologist
US Geological Survey
National Center Mail Stop 954
Reston, VA 20192
Voice (702) 648-6533
Email: pschweitzer@usgs.gov

e) Collaborating organizations

Dr. James E. Burt

Department of Geography University of Wisconsin-Madison
550 North Park Street Room 425 Science Hall
Madison, Wisconsin 53706-1404 USA
Phone: 608-263-4460
Email: jeburt@wisc.edu

f) Other relevant personnel

Greg Lee - Physical Scientist
US Geological Survey
Denver Federal Center Box 25046 MS 973
Denver, CO 80225
Phone: (303) 236-2491
Email: glee@usgs.gov

Carma San Juan - Physical Scientist
US Geological Survey
Denver Federal Center Box 25046 MS 973
Denver, CO 80225
Phone: (303) 236-2450
Email: csanjuan@usgs.gov

g) Context of the project and its importance or value

The USGS Mineral Resources Program provides scientific information for objective resource assessments and unbiased research results on mineral potential, production, consumption, and the environmental effects of mineral development. The MRP is the sole Federal source for this information. The development and maintenance of a mineral deposit database is one of the programs critical functions.

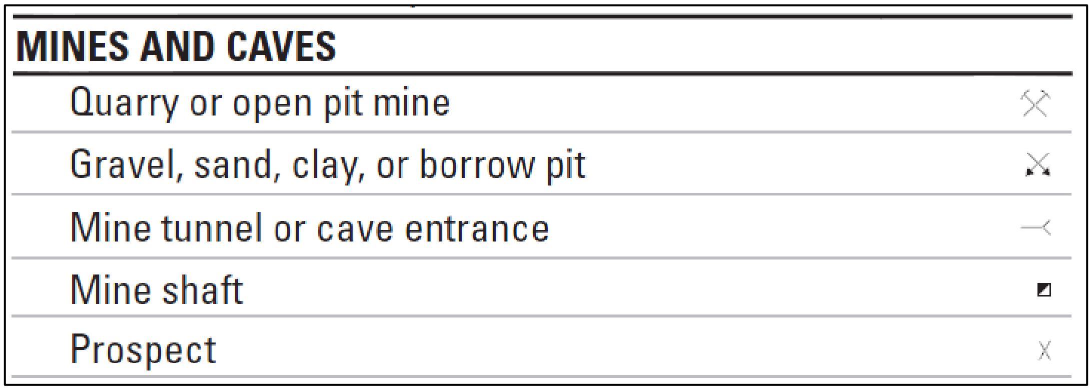
The proposed product will aid in the development of the national-scale Mineral Deposit Database. The data contained therein are located in all ecosystems and habitats in the United States. The data will be of immediate and substantial value to the USGS Mineral Resource Program and other Federal Agencies such as BLM, and Forest Service, as well as State land and environmental agencies, educational and scientific institutions and private industry.

Scope

The project will develop software to use enhanced feature recognition to extract the location and nature of mine symbols from scanned USGS topographic maps. Input files will be georeferenced TIFF files from the HTMC which were created from the largest known collection of USGS topographic maps and the conversion process, valued at approximately $1M.

The new software will build on experience gained with HTMC in which software was developed that converts an image from a scanner's coordinate system to the geographic (latitude, longitude) coordinate system implicit in the original map. The software, known as QUAD-G, operates in much the same way as other georeferencing tools: a small number of points are located in both image coordinates and geographic coordinates and are used to georectify the image. QUAD-G differs from standard tools in that the control points are found automatically rather than provided by a user through on-screen digitizing. Thus the software can process an arbitrarily large batch of scanned images without operator supervision.

In the proposed project the software will search an image for appearances of pre-defined features chosen from the USGS topographic symbol set including quarry or open pit mine, gravel or borrow pit, mine tunnel entrance, mine shaft and prospect shown below. The software will process the symbols and report all instances to a log file that will be used to generate geospatial output. In addition, there will be an option to extract and store a symbol’s surrounding sub window as a separate image file for visual inspection and checking.



MINES AND CAVES
Quarry or open pit mine
Gravel, sand, clay, or borrow pit
Mine tunnel or cave entrance
Mine shaft
Prospect

Several aspects of the problem are unusual and will require research. For example, the HTMC resolution of about 600 pixels per inch was chosen to faithfully reproduce the original map, but it is higher than needed for symbol detection. Research will address the efficacy of down sampling prior to attempting pattern recognition. Such testing will determine the tradeoff between execution time and the possibility of erroneous feature recognition. Research is also needed to explore optimum search strategies. Some symbols (e.g., quarry, pit or prospect) are more properly considered as linear features than point symbols. Others symbols can appear in arbitrary orientations (e.g., adit/mine tunnel). To address these (and other) research questions, existing GIS data for the State of Nevada will be used for testing to ensure the software meets the needs of end users.

The project will deliver an immediate benefit in the development of a tool to leverage the recently completed USGS HTMC and makes data available in geospatial formats that can be used in a wide range of applications. While the software would be developed to extract mine features from topographic maps it will lay the foundation for development of procedures to extract other features such as springs, wetlands, vegetation, and shorelines, and for feature extraction from other map data sets such as the National Geologic Map Database. The data generated by the project will be of immediate benefit to USGS MRP scientists and research programs as well as a wide range of state and federal land management agencies and research groups.

Technical Approach

Work will proceed as four tasks including software development, large area testing, testing on new areas, and final product delivery.

Task 1: Software and Procedures Development. In the first task the symbols sets will be defined and feature recognition software routines developed. Research will include an evaluation of the efficacy of down-scaling the 600 pixel per inch scans to reduce processing time, and optimizing search strategies for the desired pattern recognitions.

The software will be developed and tested using a selected area in Nevada which exhibits all of the desired features at a variety of densities. The results of test runs will be evaluated and revisions made as indicated.

Task 2: Large Area Testing. Software will be tested on a larger area in Nevada and evaluated against the existing GIS data. Operating procedures and output files will be documented and draft manuals prepared.

Task 3: Testing on New Areas. Software will be tested on an area lacking existing GIS data for mine features, most likely the State of Arizona. The test process will include data acquisition by USMIN personnel to test operating procedures. The resulting data will be evaluated using USMIN quality control procedures.

Task 4: Product Delivery. After final checks and revisions to procedures, manuals, and software the program will be delivered to USMIN.

Project Experience

Mr. Greg Fernette (USGS Denver) is Project Lead for USMIN. He has 40 years of experience as a geologist in the minerals industry and in the USGS. His experience includes the management of a number of large data collection and compilation projects. In addition Mr. Fernette has an in depth knowledge of mining, mineral deposits and GIS applications which result in an understanding of the needs of users of mineral deposit data.

Dr. Peter Schweitzer (USGS Reston) is Co-Project Lead for USMIN. Dr. Schweitzer carries out research and technological development to improve the public's ability to use scientific research results from the USGS. He worked on the MRDS project for a number of years and is the developer of the MRP spatial data portal (http://mrdata.usgs.gov/).

Dr. James Burt, professor of geography at the University of Wisconsin-Madison worked on the HTMC project and development of QUAD-G software (http://geography.wisc.edu/research/projects/QUAD-G/). This experience is foundational to the proposed project.

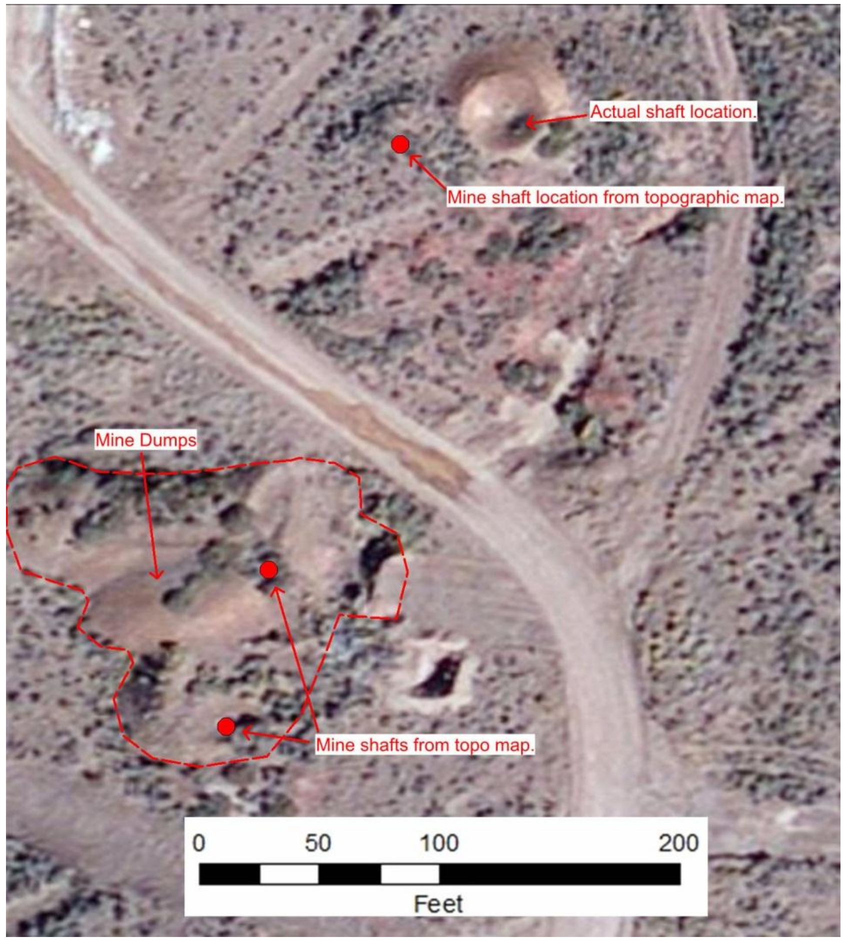
Other project personnel include Greg Lee and Carma San Juan, both of USGS Denver. Mr. Lee and Ms. San Juan have over 50 combined years of geospatial project experience at the USGS. Most recently, they served as project leads in the preparation of national GIS data sets for use in an assessment of the mineral resources of the United States.

Commitment to Effort

USMIN is an ongoing project of the MRP with the goal of preparing a comprehensive geospatial mineral deposit database for the United States. Data extraction using the software and procedures developed in this project will continue after the completion of the software development. The data extracted by the software will be reviewed and checked before incorporation into the USMIN database. This work will be done by USMIN personnel. Consideration is being given to releasing mine feature data sets as stand-alone products on a state by state basis as well as being incorporated into the national database.

After incorporation of the mine feature data into USMIN the data will move to the “Image Check” phase. In this phase the data points are checked against modern high-resolution areal imagery such as NAIP or Bing. The locations are then corrected if needed and additional features captured such as pit outlines and mine dumps as shown at right.

At this stage additional attributes are added to the data such as current status. The overall data structure is illustrated in Appendix A. The capture of the original feature location and type from the topographic maps is foundational to this process.



Budget

The total project budget is $250,045. Of this $104,030 are requested funds and $146,015 are matching funds from the Central Mineral and Environmental Resources Science Center. Funding for Dr. James Burt will be used for a Cooperative Educations Service Unit (CESU) grant.

Travel expenses are included for two trips to Madison, Wisconsin during the course of the project, one trip to a scientific meeting for Dr. Burt and one trip to the ESRI User Conference to present the results of the project. Funds are also included for two personal computers which will be set up to run the automated feature extraction software in Denver on an ongoing basis.

Budget Category	**Federal Funding**	**Matching Funds**
1. Salary		
USGS Personnel		
Greg Fernette (4 pay periods)		$24,720
Peter Schweitzer (2 pay periods)		$12,090
Greg Lee (2 pay periods)		$10,900
Carma San Juan (3 pay periods)		$14,280
Contract Personnel		
James Burt	$31,622	
Student contractors (2 contracts @ $17.31/hr@1040 hrs)	$36,005	$36,005
Total Salaries	$67,627	$97,995
2. Fringe Benefits		
James Burt	$11,733	
Total Fringe Benefits	$11,733	
3. Travel Expenses		
Per Diem and lodging		
G. Fernette – Madison, WI (10 days)	$1,450	
C. San Juan – San Diego, CA (5 days)	$1,025	
C. San Juan – Denver, CO (CDI meeting)	$0	
Airfare		
Denver-Madison RT (2)	$1,000	
Denver-San Diego (1)	$500	
Vehicle		
Madison, WI (10 days)	$400	
Other travel expenses		
J. Burt – Technical meeting travel and per diem	$2,000	
Total Travel Expenses	$6,375	
4. Other Direct Costs		
Computer Hardware – 2 PC+Monitor+solid state drive	$10,000	
Supplies (J. Burt)	$500	
Total Other Direct Costs	$10,500	
Total Direct Costs	$96,235	
Indirect Costs		
James Burt (17%)	$7,795	
Central Mineral and Environmental Science Center (49%)		$48,020
GRAND TOTAL	$104,030	$146,015

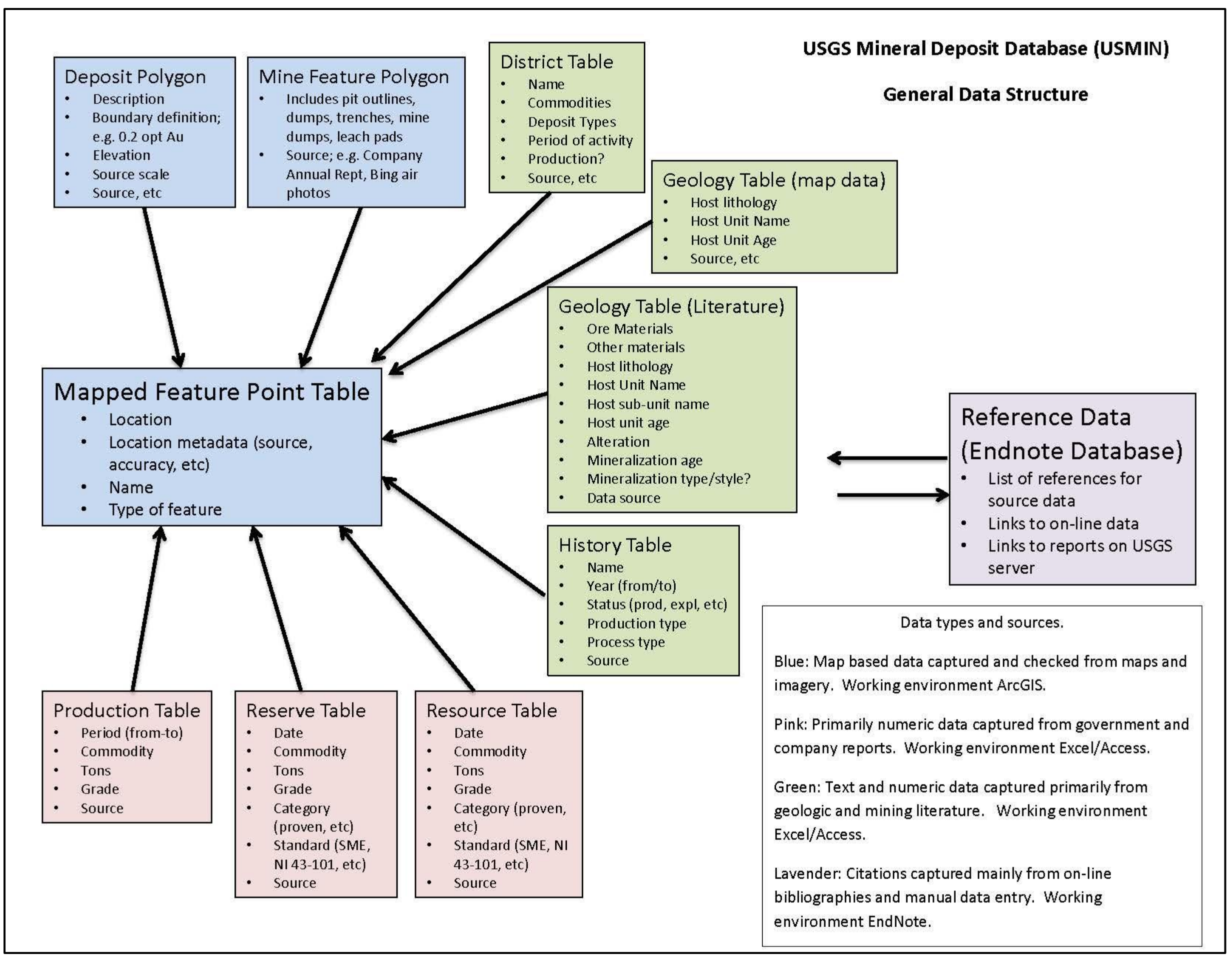
Timeline

Date(s)	Work Description
March 31	Date of Award
April 1-May 30	Task 1: Software and Procedures Development • Submission of feature requirements. • Preliminary coding. • Testing of feature recognition routines. • Development and testing of search strategy. • Tests on selected areas.
May 30	Submission of alpha version of software
June 1-June 30	Task 2: Large Area Testing • Evaluation of selected area test results. • Software modifications. • Large area testing on Nevada maps. • Draft documentation and procedures. • Purchase and set up of PC's in Denver.
June 30	Submission of beta version of software.
July 1-July31	Task 3: Testing in New Areas • Evaluation of beta version of software and large area test data. • Testing on new areas. • Evaluation of new area test results. • Revisions of documentation and procedures. • Training of Denver personnel.
August 1	Submission of final version of software
August 1-31	Task 4: Product Delivery • Establish ongoing data collection routines in Denver. • Preparation of final report.
August 31	Submission of final report and product to CDI
September 1 onward	Ongoing data collection under USMIN Project.

Appendix A: Background to the Mineral Deposit Database Project

In October 2012 the USGS Mineral Resources Program (MRP) started a new project to develop a comprehensive geospatial database of the mineral deposits of the United States (USMIN). The project follows and builds upon the Mineral Resource Data System (MRDS. The primary objectives of USMIN are to move the database into the geospatial realm and to expand the user base.

As the foundation for the database two feature types: mine features and mineral occurrences, are being captured. Mine features include things like open pits, quarries, shafts, adits and prospect pits. Mineral occurrences comprise mineral deposits (those with a grade tonnage estimate and/or production), prospects and mineral occurrences. In all cases the aim is to capture a high quality location and the area (polygonal outline) of the feature if it is available. Various other data attributes are then linked to the point/polygon feature as shown in the diagram below.



The capture of mine feature and mineral deposit data is proceeding in parallel as the sources of the data and the methodology used in its capture differ as do the qualifications of the personnel required. Mine features are captured primarily from USGS topographic maps and areal imagery whereas mineral deposit data is captured mainly from geologic reports and maps. GIS personnel can accomplish much of the mine feature data capture however geologists are needed for the capture of mineral deposit data.

The capture of mine feature data from topographic maps is an enormous task. For example the State of Nevada has 1955 7.5 minute quadrangle maps which show over 104,000 mine features. An example of an area with a high density of features is shown below. Manual digitizing of these features, even "heads up" digitizing from scanned and georeferenced maps is a time consuming process. Development of the work flows, data capture and quality control procedures for this process have been the focus of much of the work on USMIN in FY12.

Work started with a data set captured by the Nevada Department of Minerals for their abandoned mine lands program. The data was reviewed and corrections made where needed. Polygonal features, which had not been a part of the original data set were captured. Procedures for data capture, metadata and quality control were developed and tested while checking the Nevada data and are now being used to capture data from other states.

The location and nature of mine features is of value to a number of user groups. One use that receives a great deal of attention is Abandoned Mine Lands Reclamation. In a report issued in July 2011 the General Accounting Office (GAO) stated that "GAO's past work has shown that there are no definitive estimates of the number of abandoned hard rock mines on federal and other lands." The report also notes that "From 1997 to 2008, four federal agencies—BLM, the Forest Service, EPA, and OSM—had spent at least a total of $2.6 billion to reclaim abandoned hard rock mines on federal, state, private, and Indian lands." For this use alone mine feature data captured from USGS topographic maps would be of great value as illustrated by the fact that, in addition to Nevada the State of California has digitized mine features from more than 600 USGS topographic maps.