LOCATING HISTORICAL FEATURES ON RECENT IMAGERY USING AN EVOLUTIONARY ALGORITHM

Rachel Hixson
Steven Brumby
Nancy David
Kim Edlund
ISR-2
Ed Van Eeckhout
D-4
Los Alamos National Laboratory
Los Alamos, NM 87545
rhixson@lanl.gov
brumby@lanl.gov
ndavid@lanl.gov
kedlund@lanl.gov
emvan@lanl.gov

ABSTRACT

Using the automatic feature extraction software package GENIE (GENetic Imagery Exploitation) developed at Los Alamos National Laboratory (LANL), an algorithm has been generated that identifies ground features associated with historical burial sites. GENIE uses a genetic algorithm approach to evolve new feature extraction algorithms from a library of commonly-used spectral and spatial/textural image processing operators. For this particular application, GENIE used Daedalus airborne 5m visible/infrared multispectral imagery (MSI), and was trained using known trench locations from the 1940s. As part of our algorithm validation process, additional similar vintage trenches were located off LANL property on county, federal, and Native American lands. Available ground truth reports and visits to these sites confirmed the presence of historical disturbances. Hence, our initial results suggest that it may be possible to automate the mapping of these and other burial sites. This paper will briefly describe GENIE and its application in identifying these sites.

INTRODUCTION

The Los Alamos National Laboratory (LANL), located in northern New Mexico (Figure 1), lies on the Pajarito Plateau on the eastern side of the Jemez Mountains, at an average elevation of 7500 ft. The environment of the region is characterized as semi-arid (high desert). The plateau is comprised of a series of narrow mesas and canyons. The mesas of the Pajarito Plateau are composed primarily of Bandelier Tuff, a series of ash-flow deposits that were formed 1.1 to 1.5 million years ago. The tuff is soft and easily workable with power equipment. Soil cover on the mesa tops is typically less than 4 ft deep (Pope, 1996a).

Figure 1. Location map of LANL.

Since its inception, LANL has been involved in the use and disposal of different types of materials. Several in-house disposal areas were created early in the history of the institution to deposit such materials as radionuclides and hazardous organic and inorganic chemicals. According to David et al. (David, 1996), not all disposal sites are documented and located. Reasons for this include lost or destroyed records, undocumented locations, and lost memories. The search of large areas for buried waste and buried waste containers is a difficult and expensive problem when using conventional, ground-based methods.

In more recent years, LANL has been engaged in the remediation of its old hazardous waste sites. Past records, current and historical aerial photographs, satellite and airborne remote sensing, as well as geophysical surveys have all been used to characterize known waste sites. The effective combination of these data have provided clearer insight into defining several problem areas, as well as indicating where more detailed information might be required (Schowngerdt, 1997). As has been previously noted, these methods can be timely and costly. Accuracy of finding well-documented sites is relatively high, and has been useful in the clean-up process. However, in the case of poorly documented or even undocumented sites, clean-up has been near non-existent.

We investigate the use of the GENetic Imagery Exploitation (GENIE) software developed at LANL to help solve this problem. Using the well-documented Materials Disposal Area F (MDA-F), we train our software to detect image characteristics of a disturbed earth site from the late 1940s. Once the software has found the appropriate algorithm we are then able to exploit this on further imagery. This technique will not only reduce the amount of man hours required in further remediation efforts, but will also help identify unknown locations of previous disturbance.

Focusing on LANL property and the surrounding area within Los Alamos County in this study, we search for historic earth disturbance locations. While others have used aerial photography, thermal and infrared data, as well as satellite images to study previously known waste sites (Schowengert, 1997), our approach focuses on finding those trenches that have not yet been identified. This is especially important in LANL's efforts to reduce environmental damage caused by a half-century of waste contamination.

APPROACH

The remotely-sensed images used in this paper are Daedalus 1268 5m Multi-spectral data. Data were acquired May 31, and June 1, 2000, encompassing the greater Los Alamos, New Mexico area. The data were collected during good flying conditions following the main thrust of the Cerro Grande fire, a recent event that caused further environmental data collection. Imagery was geometrically corrected to RSL GCS Level 3 and has been rectified to a UTM projection Zone 13 using a WGS 84 spheroid. Imagery was terrain corrected using digital elevation models that corresponded to USGS 7 ½ minute quarter quadrant coordinates. The five imagery strips contain all 11 multispectral bands.

We did not have any atmospheric measurements available for the scene, so we did not attempt to carry out any corrections for haze or atmosphere. The topography of Los Alamos is complex, consisting of a dormant volcano (the Jemez Mountains) rising to approximately 10,000 feet (3.3km), surrounded by a radiating network of mesas at 7,000-8,000 feet, falling off to the Rio Grande river valley at approximately 6,500 feet elevation. Traditionally, illumination effects due to complex topography can be approximately factored out by using band ratios, or removed using principal components analysis. Here, we are interested in the GENIE software's ability to derive results based on the raw imagery, and do not add any additional band ratio or band difference planes.

While the only completely accurate way to locate the waste trenches we seek would be to find documentation of such areas from the time of their inception, this is an ideal which, unfortunately, can not be obtained. Thus we turned to GENIE to help us.

Genetic Imagery Exploitation

GENIE is a software system for rapidly evolving feature extraction algorithms for image analysis (Brumby, 1999; Harvey, 2000a; Holland, 1975; Perkins, 2000; Theiler, 1999). GENIE uses a genetic algorithm approach (Fogel, 1966; Harvey, 2000b; Rechenberg, 1973) to assemble image processing algorithms from a collection of low-level image processing operators (e.g., edge detectors, texture measures, spectral operations, and various morphological filters). This system has been shown to be effective in identifying complex terrain features, such as golf courses and forest fire burn severity (Brumby, 2000; 2001). GENIE can sequentially extract multiple features for the same scene to produce terrain classifications (Albers, 1993).

As the software has been discussed in-depth in other papers we will only give a brief overview here. GENIE follows the classic evolutionary paradigm: a population of candidate image-processing algorithms is randomly generated, and the fitness of each individual assessed from its performance in its environment, which for our case is a user-provided training scene. After fitness has been assigned, reproduction with modification follows via the evolutionary operators of selection, crossover, and mutation, applied to the most fit members of the population. The process of fitness evaluation and reproduction with modification is iterated until some stopping condition is satisfied. The algorithms assembled by GENIE will generally combine spatial and spectral processing, and the system was in fact designed to enable spatio-spectral image processing experimentation. Each individual chromosome in the population consists of a fixed-length string of genes. Each gene in GENIE corresponds to a primitive image processing operation, and so the whole chromosome describes an algorithm consisting of a sequence of primitive image processing steps. Once generated, this algorithm can then be exploited on further imagery and datasets.

Training Data

Our training data is based on data from LANL's Information, Records and Media Services group, as well as the Energy and Infrastructure Analysis group. Figure 2 shows MDA-F, the source for our training data.

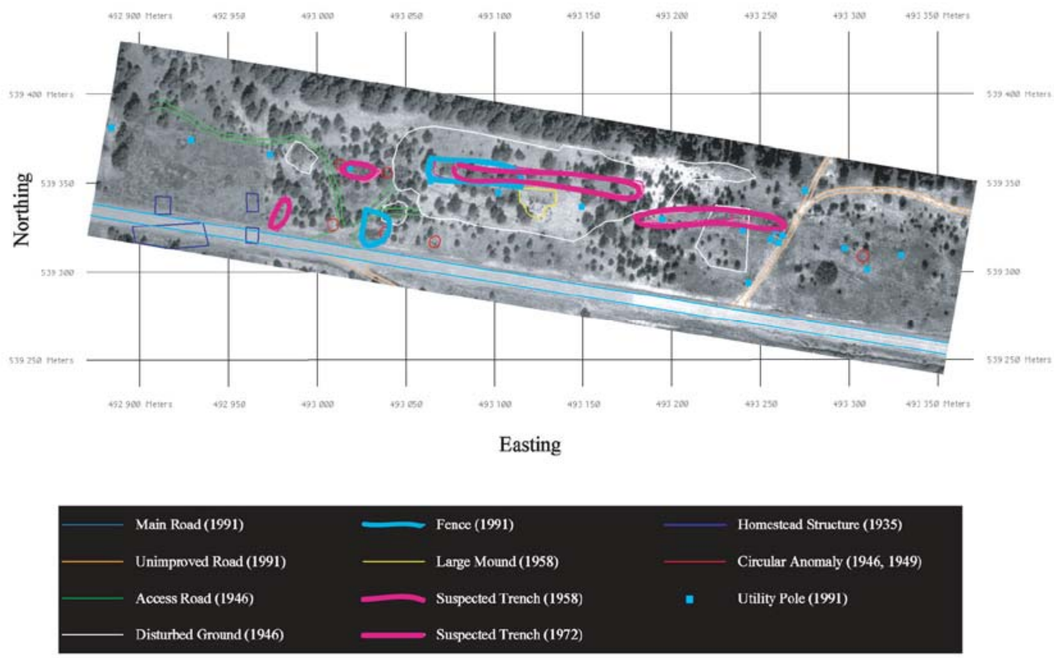

Figure 2. Overlay of past data at Materials Disposal Area F.

Figure 3. GENIE training overlay at MDA-F.

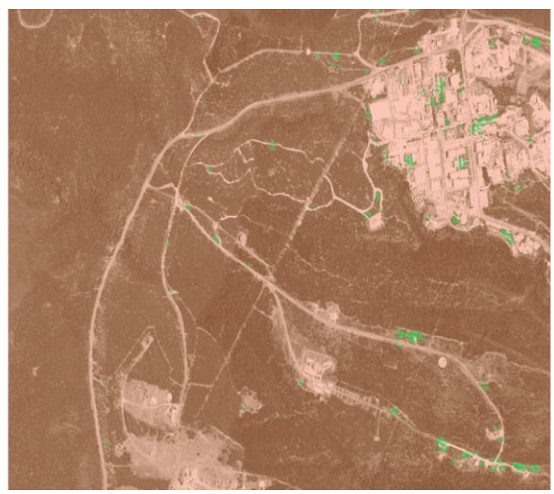

Figure 4. GENIE results overlay at MDA-F.

MDA-F was used during World War II for explosives testing and detonation of high explosive lenses. After the war, the area was used for burial of test-related, obsolete materials. Trenches were excavated through the topsoil and into the tuff, and then backfilled and capped with crushed tuff and soil. The area has been reclaimed by Ponderosa Pine and herbaceous vegetation. It is possible that high explosives may have been buried in the trenches. A long, rectangular trench boundary lies within a large area of disturbed soil at MDA-F, as can be seen in Figure 2. The main burial trench at MDA-F was chosen as the training data for this study because it is one of the most well studied (Pope, 1996a, 1996b; David, 1996).

Figures 3 and 4 show Daedalus images of the MDA-F area with GENIE training and results overlays, respectively. In Figure 3, areas of known disturbance are marked in green and areas that are known to be undisturbed are marked in red. As can be seen in Figure 3, the amount of positive data given to the software is minimal, but GENIE is still able to process it.

In Figure 4, GENIE produces results that closely match the pixel characteristics from the training data given in Figure 3 which verifies our training information. In addition, we see that in Figure 4 GENIE has detected other areas of disturbance farther south of MDA-F. The algorithm evolved from the training step is then applied it to other images from the same sensor. Thus, we found several different locations around Los Alamos County with similar pixel characteristics.

RESULTS

The results of using the evolved algorithm from GENIE on other images within the Los Alamos region are shown in Figures 5 through 7. The photos include areas around the airport, North Mesa, and Bandelier National Monument. In Figures 5 through 7, green areas indicate areas of detected ground disturbance.

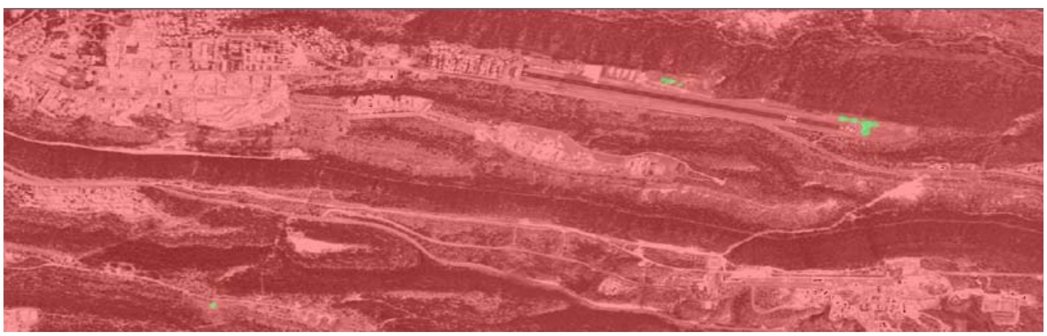

Figure 5. Results overlay from GENIE showing detected ground disturbance at the Los Alamos airport.

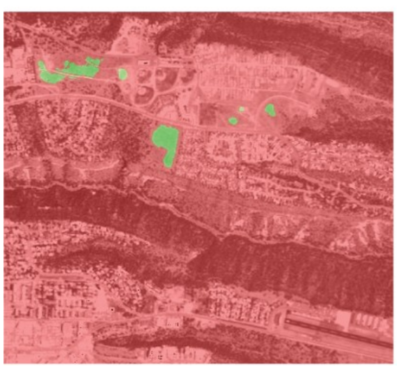

Figure 6. Results overlay from GENIE showing detected ground disturbance on North Mesa.

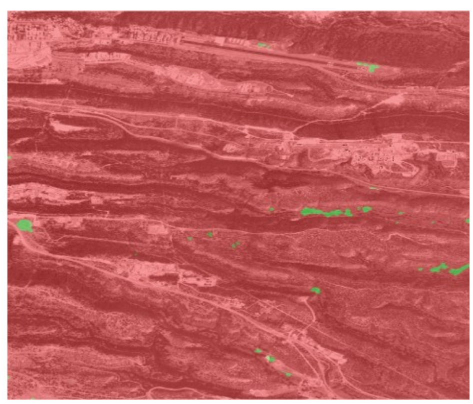

Figure 7. Results overlay from GENIE showing detected ground disturbance at Bandelier National Monument (in lower left).

Figure 8. Oblique aerial view of MDA-F, facing northeast.

Based on the GENIE-detected disturbances in a variety of locations around Los Alamos County, we conducted ground referencing at each of our selected sites, as seen in Figures 5-7. The first site we investigated was the Los Alamos Airport. We compare all of our ground truth to the ground characteristics seen in this aerial photograph of MDA-F (Figure 8).

Los Alamos Airport

The first region of interest (ROI) at the airport is past the Eastern-most part of the airstrip, as seen in Figure 5. Ground evidence at this site shows strong indication of disturbance, through presence of rough dirt and sparse vegetation.

Figure 9. Vegetation near airport runway.

Figure 10. Characteristic “disturbed” ground.

The second ROI at the airport lies just north of the midpoint of the runway. This area showed similar attributes to the first ROI. Only low ground-cover vegetation exists, excluding sage and brush that exists in other places around the airport. Most interesting to note at this site is the clear delineation between the sparser vegetation we are describing and the lusher, greener vegetation immediately adjacent to our study area (Figures 9 and 10).

North Mesa

A particularly interesting aspect of the North Mesa site is the existence of a major anomaly in the midst of residential development. The large reverse “P” in Figure 6 is this anomaly. Information from the Los Alamos Historical Society indicates the vacancy of this lot since the 1970s. Inspection of the ROI reveals similar ground characteristics to the airport site noted above. This ROI can be seen in Figure 11. Figure 12 shows a close-up of the ground.

Figure 11. North Mesa site **Figure 12.** Type of ground found at N. Mesa.

Bandelier

Within the Bandelier site there were several detected anomalies. We chose to focus on those ROIs that were close to accessible trails, as not to disturb the national monument grounds. The areas chosen can be seen as the southern-most green areas in Figure 7. As in the previous sites, the Bandelier ground referencing produced ground with sparse, established vegetation. The ground also has a different quality than in areas surrounding it, namely a dry, cracked look. Figure 13 shows the site, with a visible delineation in vegetation following the larger shrubs behind our scale. The ground type can be seen in Figure 14.

Figure 13. Bandelier site. **Figure 14.** Ground at Bandelier site.

DISCUSSION

The history of the airport site has positive implications for our study. From 1943-1975 the land north and east of the runway, as well as the canyon to the north was used as a solid waste disposal area for Los Alamos County (Snodgrass, 2003). Remediation efforts were conducted at the airport in 2003. Portage Environmental, Inc. was responsible for collecting solid waste deposited in the canyon northeast of the runway. Further remediation efforts are to be conducted in the near future, including soil sampling of both the canyon and the area on the mesa previously used as the landfill. The markers used to delineate the areas of remediation were in place at the time of our ground referencing. A point of interest concerning these markers is that the boundaries for the process coincide with the boundaries GENIE identified as areas of historic disturbance. Research into the site in the 1940s produced the archival photograph seen in Figure 15. Smoke from a fire can be seen in the center of the photograph at the end of the runway.

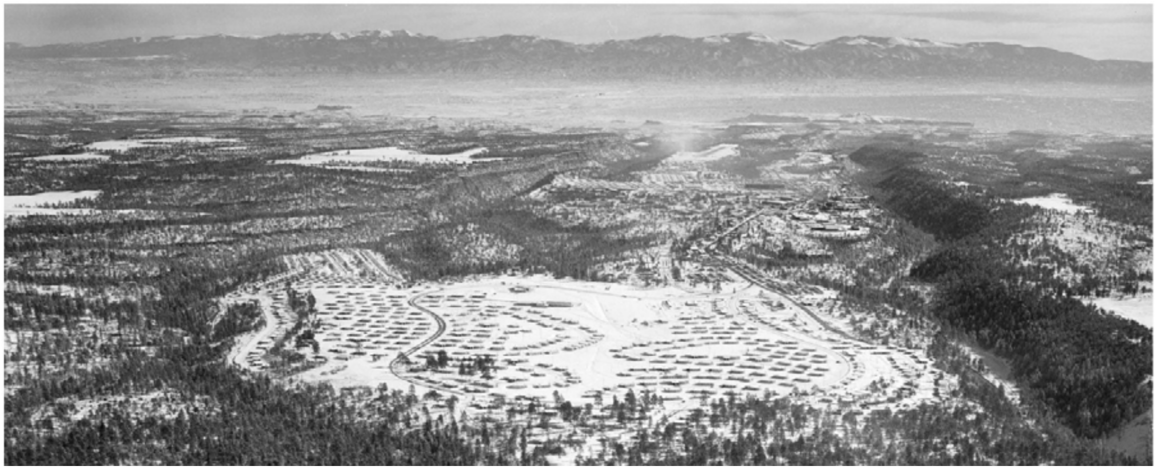

Figure 15. Mid-1940s photograph of Western Area in Los Alamos. The airport runway can be seen in the distance with a plume of smoke from a trash fire at the end of the runway.

Research into the history of the North Mesa site revealed the land had been owned by the government until the 1970s. Since that time the current owner has been in possession of the land. The nature of how the land was used by the government is not clear at this time. What is interesting about this site is how clear the boundaries are. The shape that GENIE detected is in its exact form still. Surrounded by piñon trees and residences, the land itself is similar in its appearance to that of the airport sites.

Information from the National Park Service (NPS) indicates that the land on which our Bandelier results were located was previously LANL property. The area was used as a buffer of sorts when the Manhattan Project was brought to Los Alamos in the early 1940s. The land was transferred to the NPS as part of Bandelier National Monument expansion in the 1960s (Rothman, 1988).

Our findings among all four sites (the training site and three other sites) seem consistent. Short, ground-covering vegetation can be found in the absence of large trees and bushes. From the variety of locations in which there were detected disturbances and the consistency among ground truth, we feel that GENIE has evolved an algorithm that can detect historic disturbances on recent imagery. No intrusive ground sampling has been done to confirm any type of chemical consistency in the ground; however, our visual confirmations of the consistency in the disturbed areas are homogenous within the group.

While not necessarily as prominent as in other cases, we believe the hard-packed dirt, sparse ground cover, only low-growing vegetation indicative of our sites correspond to traits similar of those found in archeology. For decades now archaeologists have been using aerial photographs, and in recent years, satellite images to detect ancient constructs that are otherwise undetectable to human eyes. Dr. Scott Madry describes one such instance of this on his website. After a drought in 1976, a Galla-Roman villa was discovered in a French farmer's after an aerial fly-by. The villa was detected because of the crop marks that occur when variations in crop vigor, color, or height when crops or natural vegetation grows over buried walls or other cultural remains (Madry, 2006). For our purposes, crop marks are not our distinguishing factors, but the principles are the same. Differences in vegetation vigor, as well as dirt clumping patterns seem to be pointing out historical disturbances in the Los Alamos area.

Especially interesting in our study is the fact that all sites were previously government owned. Each site has similar ownership history and dates that the land was given up for different uses.

SUMMARY AND CONCLUSIONS

The results for this study indicate that GENIE is detecting areas where the disturbances occurred in the same era. We are confident that with further ground truth it can be determined what GENIE is detecting. Initial applications for this technology could include identifying buried waste trenches, archaeological sites and perhaps an application for minefield remediation.

In the area of detecting buried objects by means of remote sensing, relatively little has been explored in the way of using evolutionary software to detect the sites. We propose GENIE will reduce time and cost in finding unknown buried objects through an evolved algorithm. While others have used remote sensing in detecting buried trenches, in most cases these were already well-documented cases. Our goal is to be able to find further trenches that have yet to be uncovered.

To further determine the accuracy of our detection, we hope to take soil samples in each site. Finding more concrete similarities will further the notion that GENIE is detecting a definite type of disturbance, which could lead to further research. Our data was trained using a site that had a specific type of disturbance. If we discover that it is detecting only those sites with similar disturbances, this could mean being able to find several types of historical disturbances, with only varying the training data used.

ACKNOWLEDGEMENTS

The authors would like to thank the US Park Service for access to field sites in Bandelier National Monument, and the Los Alamos Historical Society for access to historical imagery of the town of Los Alamos, New Mexico. For assistance in obtaining historic aerial photos of the Los Alamos Airport we thank Pete Trujillo from LANL's group IM-9. We would also like to thank Robert Enz from the DOE for access to the Los Alamos Airport. We thank Nancy Cerutti from the Los Alamos County Planning and Zoning Commission for her valuable information on the North Mesa site. Further thanks are given to Edward Collins for helping prepare the ground truth map of Bandelier, and to Damian Eads and Rohan Loveland for their help in obtaining ground truth from the Bandelier site.

REFERENCES

Albers, B. J., C. Purdy, and D. F. Roelant (1993). Geomatics for Environmental Characterization and Monitoring within the Department of Energy (DOE). Presented at *The Ninth Thematic Conference on Geologic Remote Sensing* Pasadena, CA-USA

Brumby, S. P., J. Theiler, S.J. Perkins, N.R. Harvey, J.J. Szymanski, J.J. Bloch, and M. Mitchell (1999). Investigation of feature extraction by a genetic algorithm, *Proc. SPIE*, Vol. 3812, pp. 24-31

Brumby, S.P., N. R. Harvey, J. J. Bloch, J. Theiler, S. Perkins, A. C. Young, and J. J. Szymanski (2001). Evolving forest fire burn severity classification algorithms for multi-spectral imagery. Presented at *Aerosense 2001: SPIE's 15th Annual International Symposium on Aerospace/Defense Sensing and Controls* Orlando, FL-USA.

Brumby, S. P., N.R. Harvey, S. Perkins, R.B. Porter, J.J. Szymanski, J. Theiler, and J.J. Bloch (2000). *A genetic algorithm for combining new and existing image processing tools for multispectral imagery*. *Proc. SPIE*, Vol. 4049, pp. 480-490.

David, N. A., I.W. Ginsberg, E.M. Van Eeckhout, L.K. Balick, A.A. Lewis, J.B. Odenweller, G.A. Stahl, W.A. Tyler, and R.M. Weber (1996). Remote Sensing Characterization of Selected Waste Sites at the Los Alamos National Laboratory. *Environmental Geosciences*, Vol. 3, No. 1.

Fogel, L., A. Owens and M. Walsh (1966). *Artificial Intelligence through Simulated Evolution*, Wiley, New York.

Harvey, N. R., S.P. Brumby, S.J. Perkins, R.B. Porter, J. Theiler, A.C. Young, J.J. Szymanski, and J.J. Bloch (2000a). Parallel evolution of image processing tools for multispectral imagery , *Proc. SPIE*, Vol. 4132, pp. 72-82.

Harvey, N.R., S. Perkins, S.P. Brumby, J. Theiler, R.B. Porter, A.C. Young, A.K. Varghese, J.J. Szymanski, and J.J. Bloch (2000b). Finding golf courses: The ultra high tech approach , *Proc. Second European Workshop on Evolutionary Computation in Image Analysis and Signal Processing (EvoIASP2000)*, Edinburgh, UK, pp 54-64.

Holland, J. H. (1975). Adaptation in Natural and Artificial Systems, University of Michigan, Ann Arbor.

Madry, S., C. Crumley (2006). Applications of Geomatics for Long Term Regional Archaeological Settlement Pattern Analysis, http://www.informatics.org/france/aerial.html

Perkins, S., J. Theiler, S.P. Brumby, N.R. Harvey, R.B. Porter, J.J.Szymanski, and J. J. Bloch, *GENIE: A hybrid genetic algorithm for feature classification in multi-spectral images* , Proc. SPIE, Vol. 4120, pp 52-62, 2000.

Pope, P., E. Van Eeckhout, M., C. Rofer, S. Baldridge, J. Ferguson, L. Balick, N. Josten, and M. Carpenter (1996a) Environmental Waste Site Characterization Utilizing Aerial Photographs, Remote Sensing and Surface Geophysics, presented at *SPECTRUM '96*, Seattle, WA-USA.

Pope, P., N. Becker, E. Van Eeckhout, C. Rofer, N. David and J. Irvine (1996b) Environmental Site Characterization Utilizing Aerial Photographs and Satellite Imagery: Three Sites at Los Alamos National Laboratory, New Mexico.

Rechenberg, I. (1973). Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution, *Fromman-Holzboog*, Stuttgart, Germany.
Rothman, H. (1998). Bandalier National Monument: An Administrative History. National Park Service Division of History Southwest Cultural Resources Center Professional Papers No. 14.
Schowengerdt, R. A. (1997). *Remote Sensing, 2nd ed.*, San Diego.
Snodgrass, R.. (2003). Airport landfill work finishes ahead of schedule and budget *Los Alamos Monitor*.
Theiler, J., N.R. Harvey, S.P. Brumby, J.J. Szymanski, S. Alferink, S.J. Perkins, R. Porter, and J.J. Bloch (1999). Evolving retrieval algorithms with a genetic programming scheme, *Proc. SPIE*, Vol. 3753, pp.416-425.