

Automating the Hunt for Volcanoes on Venus

M.C. Burl^{†‡}, U.M. Fayyad[‡], P. Perona[†], P. Smyth[‡], and M.P. Burl[‡]

[†] California Institute of Technology
MS 116-81 — Pasadena, CA 91125
{burl,perona}@systems.caltech.edu

[‡] Jet Propulsion Laboratory
MS 525-3660 — Pasadena, CA 91109
{fayyad,smyth,maureen}@aig.jpl.nasa.gov

Abstract

Our long-term goal is to develop a trainable tool for locating patterns of interest in large image databases. Toward this goal we have developed a prototype system, based on classical filtering and statistical pattern recognition techniques, for automatically locating volcanoes in the Magellan SAR database of Venus. Training for the specific volcano-detection task is obtained by synthesizing feature templates (via normalization and principal components analysis) from a small number of examples provided by experts. Candidate regions identified by a focus of attention (FOA) algorithm are classified based on correlations with the feature templates. Preliminary tests show performance comparable to trained human observers.

1 Introduction

Many geological studies use surface features to deduce processes that have occurred on a planet. The recent JPL Magellan mission, which was successful in imaging over 95% of the surface of Venus with synthetic aperture radar (SAR), has provided planetary scientists with a virtual treasure-trove for the analysis of Venusian geology. This is especially true since all the scientific data from the mission has been released by NASA in digital form on CD-ROMS ensuring widespread, low-cost access.

Previous observations from Soviet Venera 15/16, U.S. Pioneer Venus, and ground-based radar have revealed that volcanism is a dominant geologic process on Venus. Indeed, preliminary global surveys of the Magellan data show there are approximately 1400 volcanic features larger than 20km in diameter [3], and scientists estimate the number of small volcanoes (diameter < 15km) on the planet to be $\sim 10^6$ [1]. Generating a comprehensive catalog including the size, location, and other relevant information about each volcano is clearly a prerequisite for more advanced studies such as cluster analysis of the volcano locations. Such analysis could provide insight into eruption mechanics, the relationship between volcanoes and local tectonic

structure, and the pattern of heat flow within the planet.

Automatically generating a volcano catalog from the Magellan image database presents a significant challenge to current pattern recognition and machine learning capabilities; however, the alternative approach of manually locating volcanoes is simply not feasible: the estimated one million small volcanoes visible in the Magellan dataset are widely scattered throughout 30,000 1Mbyte images.

Our long-term goal is to develop a general system for locating patterns of interest in image data. Toward this goal we have developed a prototype system, based on classical filtering and statistical pattern recognition techniques, for finding volcanoes. Training for the specific volcano-detection task is obtained from examples provided by experts. The absence of absolute "ground truth" leads to some practical problems for training and performance evaluation that are discussed in Section 3. The system is described in Section 4 with preliminary results on a small set of images given in Section 5.

2 Magellan imagery

A fundamental objective of the Magellan mission was to provide global mapping of the surface of Venus. The mapping was performed using synthetic aperture radar (SAR) because of its ability to penetrate the dense cloud cover surrounding Venus. A complete description of the Magellan SAR imaging system is given in [6], so here only the most important characteristics are summarized:

- Wavelength: 12.6cm
- Frequency: 2.385 GHz - S band
- Incidence Angle: 15° - 45° (nominal)
- Range resolution: 120m - 360m
- Azimuth resolution: 120m
- Pixel-spacing: 75m (full resolution)
- Number of looks: 5 - 16

Figure 1 shows a 30km \times 30km area imaged by Magellan (illumination from the left). This area located near



Figure 1: Magellan SAR sub-image: A $30\text{km} \times 30\text{km}$ region containing a number of small volcanoes. Illumination is from the left; incidence angle $\approx 40^\circ$.

(lat 30°N , lon 332°) contains many small volcanoes. Observe that the larger volcanoes in this figure have the classic radar signature one would expect based on the topography; that is, the upward sloping surface of the volcano in near-range (close to the radar) scatters more energy back to the sensor than the surrounding flat plains and therefore appears bright. The downward sloping surface of the volcano in far-range scatters energy away from the sensor and therefore appears dark. Together, these effects cause the volcano to appear as a left-to-right *bright-dark pair* within a circular planimetric outline. Near the center of the volcanoes, there is usually a summit pit that appears as a *dark-bright pair* because the radar energy backscatters strongly from the far-range rim. Small pits, however, may not appear or may appear as only a bright spot due to the image resolution.

The topography-induced features described above are the primary visual cues that geologists report using to locate volcanoes. However, there are a number of other, more subtle cues. The apparent brightness of an area in a radar image depends not only on the macroscopic topography but also on the surface roughness relative to the radar wavelength. Thus, if the flanks of a volcano have different roughness properties than the surrounding plains, the volcano may appear as a bright or dark circular area instead of as a bright-dark pair. Volcanoes may also appear as radial flow patterns, texture differences, or disruptions of graben. (Graben are ridges or grooves in the planet surface, which appear as bright lines in the radar imagery — see Figure 1.)

3 Uncertain ground truth

In the volcano location problem, as in many remote-sensing applications, real ground truth does not exist. No one has ever been to the surface of Venus (apart from a Russian robotic lander which melted within minutes), and trained scientists cannot determine from the imagery with 100% certainty whether an arbitrary feature is indeed a volcano (due to resolution, noise level, etc.). As a result, training and evaluation of an automatic system for finding volcanoes must be handled very carefully.

3.1 Collection of training examples

We have developed a graphical interface for harvesting training examples from scientists. The scientists use a mouse to identify image features that may be volcanoes and then specify size information by fitting circles. The scientists also provide a label indicating their subjective confidence p that the selected object is in fact a volcano. Based on discussions with the scientists, the confidence labels were quantized into four categories:

Category 1: $p \in [0.95, 1.0]$. Almost certainly a volcano, with all primary visual cues present.

Category 2: $p \in [0.75, 0.95]$. Probably a volcano, but a non-essential visual cue is missing.

Category 3: $p \in [0.5, 0.7]$. Possibly a volcano, but some primary visual cues are unclear or missing.

Category 4: $p \approx 0.5$. Only a pit is visible; could be a volcano, but more evidence is needed.

Some typical volcanoes from each category are shown in Figure 2.

The confidence labels may be incorporated into training as described in [8] (see also [5, 7]). However, we have not found a significant improvement in performance by doing so. The results presented in the remainder of this paper are, therefore, based on “label-free” training in which all examples are treated as definite volcanoes.

3.2 Evaluation of performance

Given that the scientists cannot classify each object with 100% confidence, how can we assess how well our algorithms are performing? The basic idea is to generate a “consensus ground truth” from several scientists working together and discussing the merits of each candidate volcano. Typical consensus data is shown in Figure 3 (without the confidence labels). An algorithm is considered satisfactory if its performance is comparable to that of an individual scientist.

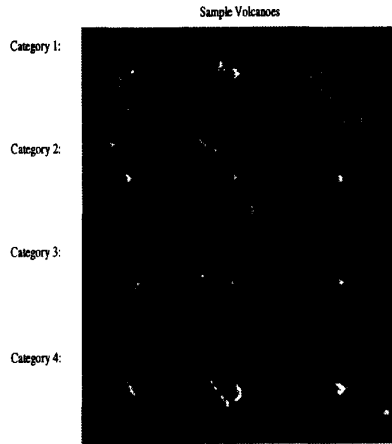


Figure 2: Volcano confidence categories.

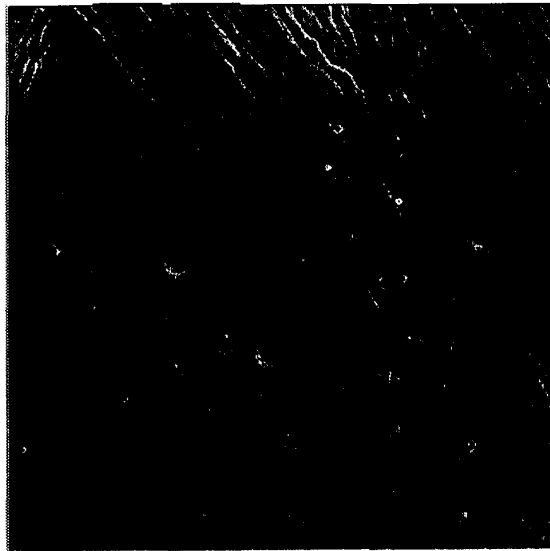


Figure 3: Magellan image with consensus ground truth showing size and locations of small volcanoes.

		Consensus				
		1	2	3	4	0
A	1	28	11	6	1	0
	2	5	9	8	9	9
	3	1	2	20	8	31
	4	1	2	5	26	13
	0	0	6	11	4	0

		Consensus				
		1	2	3	4	0
B	1	19	6	6	3	1
	2	9	5	9	4	6
	3	4	13	18	6	37
	4	0	3	3	25	18
	0	3	3	14	10	0

Figure 4: The performance of two individual scientists (A and B) compared to ‘consensus’ ground-truth.

We evaluated the individual performance of two scientists (A and B) by asking each to *independently* label the volcanoes in a set of images using the subjective probability categories described above. Approximately one week later, the two scientists *jointly* generated a consensus labeling of the same images. (Note: they did not directly use their previous labelings.) Their confusion matrices are given in Figure 4, where the (i, j) entry is interpreted as the number of volcanoes labeled i by an individual that were labeled j in the consensus. Hence, off-diagonal elements show errors between the individual and consensus. The last row of the confusion matrix shows the number of misses (volcanoes not labeled by the individual), while the last column shows the number of false alarms. In Section 5 we will compare the performance of our algorithm with that of the scientists.

4 Algorithm description

In this section, we describe the prototype system we have developed for finding small volcanoes on Venus. As illustrated in Figure 5, the algorithm operates in two distinct phases: learning and production.

The purpose of the learning phase is to develop models of the objects of interest based on training examples (and possibly counter-examples). Ideally, the learning phase should be generic enough to enable the system to be applied to other detection tasks *without reprogramming* — the user merely supplies the system with a new set of examples. (Note that the number of examples required from the user should be as small as possible since labeling is

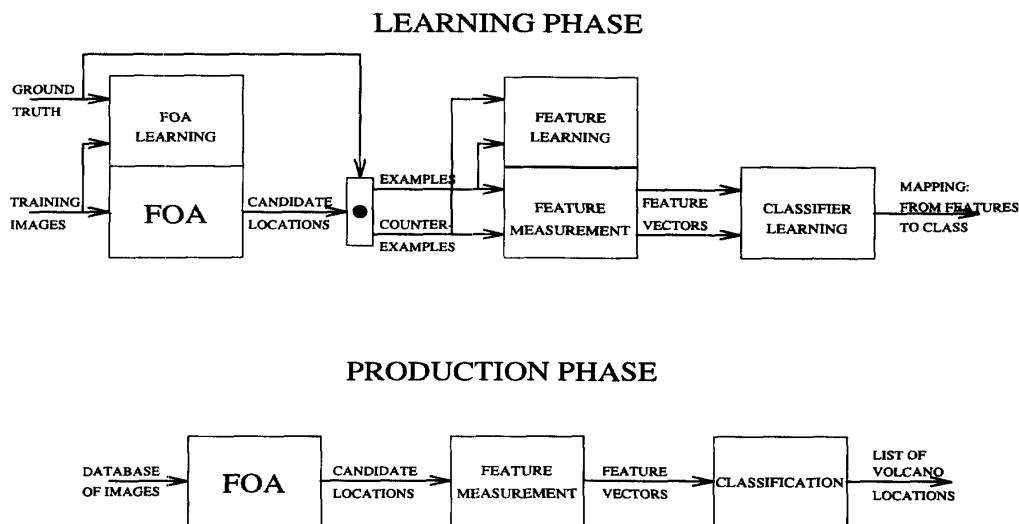


Figure 5: System block diagram.

very labor-intensive). During the production phase, the learned models are used to search through a database and find all objects of interest. Both the learning and production phases consist of three main components: focus of attention (FOA), feature extraction, and classification.

4.1 Focus of attention

The FOA is designed to quickly screen an image and output a list of candidate volcano locations. The FOA provides significant computational savings by excluding uninteresting areas of the planet (e.g., flat plains) from further consideration. A drawback, however, is that any volcanoes missed by the FOA will not be recovered.

Given the constraints of speed and low miss-rate, a reasonable approach to FOA is to use a matched filter, i.e., a linear filter that matches the signal one is trying to find. It can be shown that for detecting a known signal in white Gaussian noise, the matched filtering approach is optimal. Of course, the volcano problem does not satisfy these underlying assumptions. The set of observed volcanoes cannot be described as a known signal plus white noise, because there is structured variability due to size, type of volcano, surface roughness, etc. Likewise, the clutter background cannot be properly modeled as white noise. Nevertheless, we have empirically found that a modified matched filtering approach works well.

Let \mathbf{v}_i denote a $k \times k$ pixel region around the i -th training volcano. Each region can be normalized with respect to the local DC level and contrast as follows:

$$\hat{\mathbf{v}}_i = \frac{\mathbf{v}_i - \mu_i \cdot \mathbf{1}}{\sigma_i} \quad (1)$$

where μ_i is the mean of the pixels in \mathbf{v}_i and σ_i is the standard deviation. We construct a modified matched filter \mathbf{f} by averaging the normalized examples.

Applying the matched filter to an image involves computing the normalized cross-correlation of \mathbf{f} with each $k \times k$ image patch. The cross-correlation can be computed efficiently using separable kernel methods to approximate the 2-D kernel \mathbf{f} as a sum of 1-D outer products [9]. High response values indicate that there is strong correlation between the filter and the image patch. Candidate volcano locations are placed where the matched filter response exceeds a threshold that is determined from training images. Any threshold crossings within a prescribed distance from each other are attributed to the same object and grouped together.

Although the matched filter can be justified on purely empirical grounds, we also offer the following arguments. First, the $k \times k$ windowing eliminates some of the inherent volcano variability, especially that due to size. Focusing on the central area makes the volcano problem more like that of finding a known signal. Second, normalizing each image patch with respect to the DC level and contrast causes non-descript clutter areas to resemble zero-mean, white noise. Hence the filter \mathbf{f} should be very suitable for discriminating these non-descript regions from volcanoes — the primary purpose of FOA. Of course, in places where the clutter has features such as graben, the matched filter is not ideal and will produce more false alarms.

4.2 Feature extraction

The ability to determine good features automatically from examples is essential for developing a trainable pattern recognition tool. Since the regions of interest (ROIs) identified by the focus of attention are embedded in a high dimensional pixel space, the set of possible features is immense. We have therefore restricted our search to the family of features defined by linear combinations of the ROI pixel values. This strategy is equivalent to projecting the n -dimensional pixel space onto a q -dimensional subspace (feature space).

Projecting onto a subspace is a classical problem in pattern recognition, from which two basic paradigms have emerged [4]: linear discriminant analysis (LDA) and principal components analysis (PCA). LDA attempts to maximize the separation of classes in the projected space, while PCA attempts to best represent the original data. For the volcano problem, the class \bar{v} , consisting of all ROIs that are not volcanoes, is extremely broad. This class remains broad even after conditioning the ROIs upon acceptance by the FOA. Hence, the usual LDA scatter criterion will not provide a useful measure of the class separability, and the resulting features will not be good. A better approach is to use PCA to define features that encode the volcano class and then accept or reject an ROI based on how similar it is to the volcanoes.

The method of principal components has been used extensively in statistics, signal processing (Karhunen-Loeve transform), and pattern recognition (Turk and Pentland [10]). The basic problem formulation is to find a q -dimensional subspace such that the projected data is closest in L_2 norm to the original data. The subspace we seek is spanned by the highest-eigenvalue eigenvectors of the data covariance matrix. Although the full covariance matrix cannot be computed reliably from the number of examples we typically have available, the approximate basis vectors can be computed using the singular value decomposition (SVD) as described below.

Each normalized training volcano is reshaped into a vector and placed as a column in an $n \times m$ matrix X , where n is the number of pixels in an ROI and m is the number of ROIs. With the SVD, X can be factored as follows:

$$X = USV^T \quad (2)$$

For notational convenience, we will assume m is less than n . Then in Equation 2, U is an $n \times m$ matrix such that $U^T U = I_{m \times m}$, S is $m \times m$ and diagonal with the elements on the diagonal (the singular values) in descending order, and V is $m \times m$ with $V^T V = V V^T = I_{m \times m}$. Notice that any column of X (equivalently, any ROI) can be written exactly as a linear combination of the columns of U . Furthermore, if the singular values decay quickly enough, then

the columns of X can be closely approximated using linear combinations of only the first few columns of U . That is, the first few columns of U serve as an approximate basis for the entire set of examples in X .

The best q -D subspace on which to project is the one spanned by the first q columns of U . The columns of U are shown in Figure 6-b reshaped into ROIs; we refer to these as features or templates. Notice that the first ten templates exhibit structure while the remainder appear very random. This suggests projecting onto a subspace of dimension ≤ 10 . The singular value decay shown in Figure 6-c also indicates that 6–10 features encode most of the information in the examples.

Having determined q , we project an ROI into feature space as follows:

$$\mathbf{y} = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \dots \quad \mathbf{u}_q]^T \mathbf{x} \quad (3)$$

where \mathbf{x} is the ROI reshaped as an n -dimensional vector of pixels, \mathbf{u}_i is the i -th column of U , and \mathbf{y} is the q -dimensional vector of measured features. These feature vectors will serve as input to the classification algorithm.

Despite its intuitive appeal, there are a number of arguments against using such a simple template-based approach for recognition: most notably it is not invariant with respect to translation, rotation, scaling, and direction of illumination. A certain (hopefully small) number of templates will be required in order to represent the inherent variability of an object; any additional variability due to spatial shifting, rotation, scaling, or noise will dramatically increase the number of templates required to encode the object. Thus, the template-based approach may not be feasible unless appropriate normalization steps are taken prior to feature learning. These invariance issues need to be resolved in order to develop a general system; however, for the volcano problem they are not critical since (1) the FOA algorithm “centers” the volcanoes well, (2) the volcanoes have significant rotational symmetry, (3) the central area of the volcanoes (on which the templates are based) are relatively insensitive to overall scale, and (4) the direction of illumination is known.

4.3 Classification

The goal of classification is to determine a mapping from feature space to class label (v or \bar{v}). Up to this point in the processing, we have eschewed using counter-examples for training. (The FOA filter and PCA features were determined solely based on volcanoes.) The classifier could also be designed this way, but Fukunaga shows that such an algorithm is subject to considerable error even in relatively low dimensions because the location of the “other” distribution is unknown [4]. To overcome this problem,

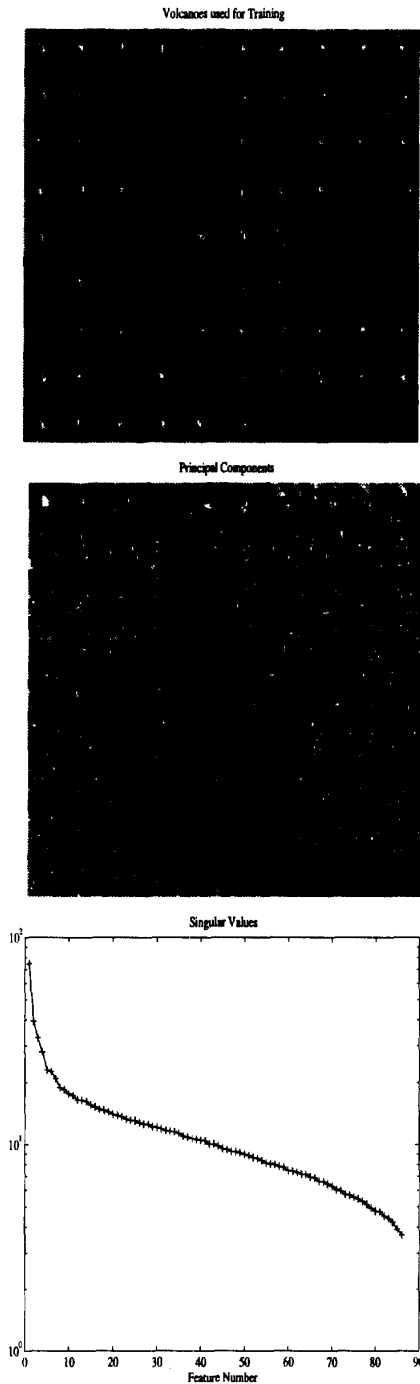


Figure 6: (a) example volcanoes, (b) principal components, (c) singular value decay

we have experimented with various supervised two-class methods including quadratic classifiers, decision trees, and kernel density estimation. Good results were obtained with all these methods, but we currently favor the quadratic classifier due to its simplicity.

The quadratic classifier is optimal if the class-conditional probability densities of the feature vector \mathbf{y} are multivariate Gaussian. Assuming \mathbf{y} has the postulated class-conditional densities, the posterior probability that an ROI is a volcano can be estimated using Bayes rule:

$$p(v|\mathbf{y}) = \frac{p(\mathbf{y}|v)p(v)}{p(\mathbf{y}|v)p(v) + p(\mathbf{y}|\bar{v})p(\bar{v})} \quad (4)$$

where $p(v)$ and $p(\bar{v})$ are the respective prior probabilities, and

$$\begin{aligned} p(\mathbf{y}|v) &= N(\mathbf{y}, \mu_v, \Sigma_v) \\ p(\mathbf{y}|\bar{v}) &= N(\mathbf{y}, \mu_{\bar{v}}, \Sigma_{\bar{v}}) \end{aligned} \quad (5)$$

with the notation $N(\mathbf{y}, \mu, \Sigma)$ denoting the multivariate Gaussian density with mean μ and covariance Σ . One can show that thresholding the posterior probability in Equation 4 is equivalent to partitioning the feature space with a quadratic hypersurface.

5 Experimental performance results

Preliminary experiments were conducted using a cross-validation paradigm on four images that contained 163 small volcanoes and covered a $150\text{km} \times 150\text{km}$ area of the planet. All results were scored relative to the scientists' consensus labeling with confidence categories 1-4 treated as true volcanoes. The figure of merit we measure is the number of true volcanoes detected versus the number of false alarms. Both quantities are expressed as percentages relative to the total number of volcanoes. Hence, the detection rate is bounded above by 100%, while the false alarm rate can be arbitrarily large.

The performance of the matched filter focus of attention algorithm is shown in Figure 7. The algorithm has a free parameter (a threshold) that controls its aggressiveness in declaring volcanoes, i.e., the trade-off between misses and false alarms. Varying this parameter generates a curve comparable to a standard ROC (receiver operating characteristic) curve. Since the purpose of the FOA is to provide a quick list of candidates to later stages of the algorithm with a low miss rate, this threshold parameter was fixed to provide a 90% detection rate; the resulting false alarm rate was about 300%. At this operating point the matched filter detected all the Category 1 and 2 volcanoes; the 10% missed were from Category 3 and 4.

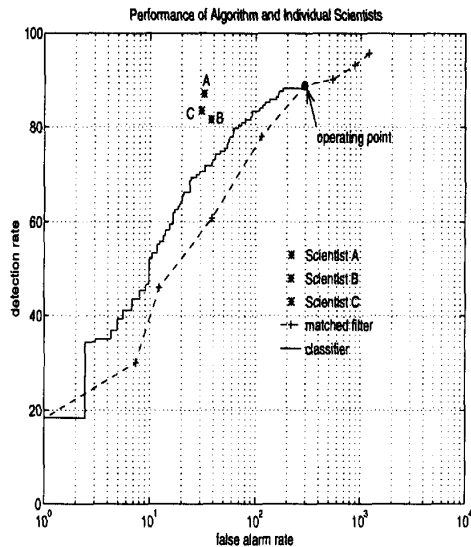


Figure 7: Performance of our algorithm compared to individual scientists.

Following the FOA, candidates are mapped from pixel space to feature space and then classified using the quadratic classifier. The goal is to reject any false alarms generated by the FOA while retaining as many true detections as possible. The performance of the end-to-end algorithm using a six dimensional feature space ($q = 6$) is shown in Figure 7. As with the FOA, the classifier has a parameter that can be varied to select aggressiveness (namely, the threshold applied to Equation 4). At maximum aggressiveness, every candidate from the FOA is declared to be a volcano; hence, the classifier performance curve is constrained to start from the FOA operating point. Observe that the combination of FOA and classification yields better performance than using only a matched filter (as proposed in [11]).

As a basis for comparison, we evaluated the individual labeling performance of three planetary scientists who are all familiar with the Magellan data and with the appearance of volcanoes in the data (see Section 3.2). Each scientist's performance is plotted as an asterisk in Figure 7. Note that the algorithm detection rate is clearly within 10% of Scientist B's detection rate at the same false alarm rate.

We also performed an empirical study to investigate the sensitivity of the algorithm to the number n of SVD features used. Figure 8 shows the measured detection rate versus n at a few selected false alarm rates. Since the detection curves are quite flat versus n , we conclude that the performance is relatively insensitive to the exact number of features, provided at least four are used.

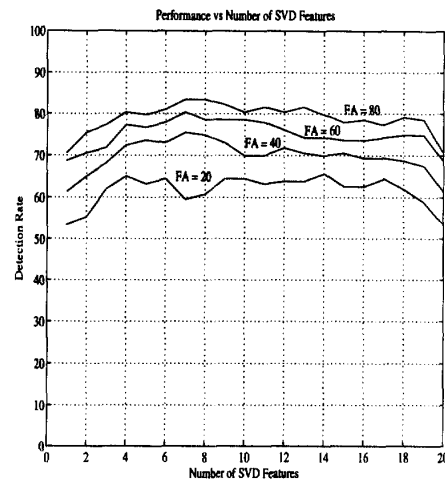


Figure 8: Empirical performance vs number of SVD features.

6 Discussion and conclusion

We have developed a trainable system, which is based on classical filtering and pattern recognition techniques, for locating small-scale patterns in large image databases. The system is being applied to the problem of locating small volcanoes in the Magellan SAR imagery of Venus. Tests conducted on four images (containing 163 small volcanoes and covering an area $150\text{km} \times 150\text{km}$) show that our system is performing nearly as well as trained human observers. These results should, of course, be considered preliminary being based on just four images (out of 30,000); further experiments are underway.

As we mentioned at the outset, our long-term goal is to develop a trainable pattern recognition tool that can be applied to various remote-sensing and visual inspection problems without reprogramming. The prototype system described in this paper is trained completely from examples and appears to work well for finding volcanoes. However, there are a number of technical issues that remain to be addressed. Achieving invariance to translation, scaling, rotation, and illumination without renouncing the advantages of filter-based processing is foremost. (Recall that these issues were not critical for the volcano problem.) Effectively using counter-examples and allowing the scientists to enter "hints" such as "find this object at any scale" are important open issues. Finally, we note that a general system will likely incorporate a variety of pattern recognition techniques, each more or less suited to particular types of problems. Automatically deciding which technique(s) to apply in a specific situation is a problem we are currently

investigating [2].

Acknowledgements

The research described in this report has been carried out in part by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. Support was provided primarily by NASA Office of Advanced Concepts and Technology (OACT - Code CD), a JPL DDF award, and NSF research initiation grant IRI 9211651. We would like to thank geologists Jayne Aubele and Larry Crumpler of Brown University for their assistance in labeling and analyzing the Magellan data. We would also like to thank John Loch, Jennifer Yu, and Joe Roden for help in developing the software and user-interfaces.

References

- [1] J.C. Aubele and E.N. Slyuta. Small domes on venus: characteristics and origins. *Earth, Moon, and Planets*, 50/51:493–532, 1990.
- [2] S. Chien, J. Gratch, and M.C. Burl. A statistical approach to adaptive problem solving for large-scale scheduling and resource allocation problems. In *Decision-Theoretic Planning Spring Symposium*, to appear, spring 1994.
- [3] J. W. Head et al. Venus volcanic centers and their environmental settings: Recent data from magellan. *American Geophysical Union Spring meeting abstracts*, EOS 72:175, 1991.
- [4] K. Fukunaga. *Introduction to Statistical Pattern Recognition*. Academic Press, 2nd edition, 1990.
- [5] G. Lugosi. Learning with an unreliable teacher. *Pattern Recognition*, 25(1):79–87, 1992.
- [6] G.H. Pettengill, P.G. Ford, W.T.K. Johnson, R.K. Raney, and L.A. Soderblom. Magellan: radar performance and data products. *Science*, 252:260–265, 1991.
- [7] B. Silverman. Some asymptotic properties of the probabilistic teacher. *IEEE Trans. Info Theory*, IT-26(2):246–249, 1980.
- [8] P. Smyth. Learning with probabilistic supervision. In *Computational Learning Theory and Natural Learning Systems 3*, To appear 1994.
- [9] S. Treitel and J. Shanks. The design of multistage separable planar filters. *IEEE trans. Geosci. Electron.*, GE-9(1):10–27, 1971.
- [10] M. Turk and A. Pentland. Eigenfaces for recognition. *J. of Cognitive Neurosci.*, 3:71–86, 1991.
- [11] C.R. Wiles and M.R.B. Forshaw. Recognition of volcanoes on venus using correlation methods. *Image and Vision Computing*, 11(4):188–196, 1993.