Terrain Classification and Classifier Fusion for Planetary Exploration Rovers

Ibrahim Halatci, Christopher A. Brooks, Karl Iagnemma
Massachusetts Institute of Technology
Department of Mechanical Engineering
77 Massachusetts Avenue, Room 3-472m
Cambridge MA 02139
617-253-2334
ihalatci@alum.mit.edu ,{cabrooks; kdi}@mit.edu

Abstract—Knowledge of the physical properties of terrain surrounding a planetary exploration rover can be used to allow a rover system to fully exploit its mobility capabilities. Here a study of multi-sensor terrain classification for planetary rovers in Mars and Mars-like environments is presented. Two classification algorithms for color, texture, and range features are presented based on maximum likelihood estimation and support vector machines. In addition, a classification method based on vibration features derived from rover wheel-terrain interaction is briefly described. Two techniques for merging the results of these "low-level" classifiers are presented that rely on Bayesian fusion and meta-classifier fusion. The performance of these algorithms is studied using images from NASA's Mars Exploration Rover mission and through experiments on a four-wheeled test-bed rover operating in Mars-analog terrain. It is shown that accurate terrain classification can be achieved via classifier fusion from visual and tactile features[12].

TABLE OF CONTENTS

1. INTRODUCTION

Near-term scientific goals for Mars surface exploration are expected to focus on understanding the planet's climate history, surface geology, and potential for past or present life. To accomplish these goals, rovers will be required to safely access rough terrain with a significant degree of autonomy. Terrain areas of interest might include impact craters, rifted basins, and water-carved features such as gullies and outflow channels [1]. Such regions are in general highly uneven and sloped, and may be covered with loose drift material that causes rover wheel slippage and sinkage.

Terrain physical properties can strongly influence rover mobility, particularly on sloped, natural terrain [2]. For example, a rover might easily traverse a region of packed soil, but become entrenched in loose drift material. The effect of terrain properties on rover mobility was exemplified in April–June, 2005 and again in May–June, 2006 when NASA's Mars Exploration Rover (MER) Opportunity became entrenched in loose drift material and was immobilized for several weeks. Knowledge of terrain properties could allow a system to adapt its control and planning strategies to enhance performance, by maximizing wheel traction or minimizing power consumption.

Related Work

Terrain classification methods provide semantic descriptions of the physical nature of a given terrain region. These descriptions can be associated with nominal numerical physical parameters, and/or nominal traversability estimates, to improve traversability prediction accuracy. Numerous researchers have proposed terrain classification methods based on features derived from remote sensor data such as color, image texture, and range (i.e. surface geometry). Most of these algorithms have been developed in the context of terrestrial unmanned ground vehicles where the visual features have wide variance. It should be noted that a planetary surface presents a difficult challenge for classification since scenes are often near-monochromatic, terrain surface cover consists mainly of sands of varying composition and rocks of diverse shapes, and sandy "crusts" can form on (and therefore obscure) rocks.

Color-based methods for classification and segmentation of natural terrain have been developed that are accurate and computationally inexpensive. For these methods, researchers have utilized multi-spectral imaging [3], different color spaces and their distribution statistics [4] along with mixture of Gaussians modeling for classifying outdoor scenes [5] because many major terrain types such as soil, vegetation, and rock possess distinct color signatures. Color-based classification is also attractive for

[1] 1-4244-0525-4/07/$20.00 ©2007 IEEE.

[2] IEEEAC paper #1166, Version 2, Updated December 8, 2006

planetary exploration rover applications since most past, current, and planned rovers have included multi-spectral imagers as part of their sensor suites [6].

Texture is also an extensively used feature in this domain. Gabor filters [7], Fast Fourier Transform [4] and histogram-based methods [8] demonstrated effective results at segmenting natural scenes although they are generally computationally expensive.

A standard approach for detecting obstacles relies on stereo cameras or range finders. Algorithms that use such sensors generally exploit elevation points [5], [9]; statistical distributions of 3D data points [10]; or disparity maps [11]. Note that such methods allow for detection of "geometric" hazards or terrain features such as rocks, however they cannot easily detect "non-geometric" hazards or terrain classes that are not characterized by geometric variation.

Although nearly all terrain classification methods rely on features derived from remote sensor data, recently methods have been proposed to classify terrain based on "tactile" features. A method for terrain classification based on analysis of vibrations arising from robot wheel-terrain interaction was first proposed in [2] and developed by [12]. Similar work was presented in [13] and [14]. It was shown that data from various sensor modalities can be fused to produce reliable class estimates.

Classifier fusion methods attempt to combine the results from "low-level" classifiers into class assignments that are (ideally) of higher accuracy than those attainable from any individual classifier. Recent work in classifier fusion includes algorithms that fuse intensity and elevation data to identify scientifically interesting targets [15], [16]; color, texture, spatial dependence, and elevation data for rock detection [17]; and color and texture histograms for geological target detection [18]. Note that several methods exist that employ a larger set of visual features such as texture and infrared imaging in addition to range data; however, their focus is detecting relatively structured roads and obstacle detection rather than terrain classification [7], [19].

This paper presents a study of multi-sensor terrain classification for planetary rovers in Mars and Mars-like environments. Two "low-level" classification algorithms for color, texture, and range features are presented based on maximum likelihood estimation and support vector machines. In addition, classification of terrain based on features derived from rover wheel-terrain interaction is briefly described. Two techniques for merging the results of these low level classifiers are presented that rely on Bayesian fusion and meta-classifier fusion. The performance of these algorithms is studied using images from NASA's Mars Exploration Rover mission and through experiments on a four-wheeled test-bed rover operating in Mars-analog terrain. It is shown that accurate terrain classification can be achieved via classifier fusion from visual and tactile features.

2. DESCRIPTION OF LOW LEVEL CLASSIFIERS

Classifier Architectures

Two low-level classifiers are defined that rely solely on a single feature type. As noted in Section 1, such classifiers have been studied extensively for terrain classification. Here we study the performance of two distinct classification methods: a maximum likelihood classifier based on mixture of Gaussians modeling (MoG), and a support vector machine (SVM) classifier.

MoG Method—The MoG method models the distribution of data points in the feature space as a mixture of Gaussians (MoG) [20]. The likelihood of the observed feature y given the terrain class x is computed as a weighted sum of k Gaussian distributions:

$$f(y \mid x_i) = \sum_{j=1}^{k} \alpha_j G\left(y, \mu_j, \Sigma_j\right) \tag{1}$$

Here, α is the weight of the Gaussian component whose mean and variance is defined by μ and Σ, respectively. Parameters of the model are learned through off-line training using the Expectation Maximization algorithm [20], [21]. Similar to [5] good results were obtained using three to five Gaussian modes, with a greater number of modes often leading to over-fitting.

SVM Method—The second classification method was based on a Support Vector Machine (SVM) framework [22]. This approach builds a binary classifier for each pair of classes and is constructed as a linear combination of similarity measures between the point to be classified y and the training points x_j:

$$f(y) = \sum_{j=1}^{n} \alpha_j K\left(y, x_j\right). \tag{2}$$

The similarity measure, K, is the kernel function. For this work linear, polynomial, and Gaussian kernels were evaluated. Values for the α_j are calculated during training by minimizing a loss function over the training data set. Complexity of the function $f(y)$ is limited by restricting the values of α_j to lie in the range $[0,C]$, and for the Gaussian kernel by controlling the width of the Gaussian using a parameter γ. Cross-validation over a training data set was used to determine an appropriate choice of kernel and reasonable values for the regularization parameters C and γ.

The SVM algorithms used in this work were implemented with the LIBSVM library with additional optimization for

linear classification [23]. Binary classifiers were combined into multi-class classifiers using a voting scheme.

Feature Selection

Color—Color is an obvious distinguishing characteristic of many terrain types and color-based classification has yielded accurate results in natural terrain [5], [9]. It should be noted, however, that color variation is somewhat limited for the surface of Mars. Mars' lack of moisture (and, therefore, vegetation) leads to a narrow distribution of colors for distinct terrain types. In this work red, green and blue channel intensity values were selected as the 3D color feature vector for every image pixel. Construction of this feature vector for MER imagery was slightly different due to the nature of the rover imaging system, and is detailed in Section 4.

Texture—Texture is a measure of the local spatial variation in image intensity. For our present work, the texture length scale of interest is on the order of tens of centimeters. This scale allows us to observe textural appearances of surfaces in the range of four to thirty meters, which corresponds to the range of interest for local planetary rover navigation [24]. In this work we employ a wavelet-based fractal dimension signature method, which yields robust results in natural texture segmentation as demonstrated by [25]. For this work, three levels of transformation were applied using the Haar wavelet kernel and neighborhood windows of 7, 9, and 11 pixels. This feature extraction method yields a 3D feature vector for every pixel.

Range—Surface geometry information can be used to distinguish between terrain classes that possess inherent geometric dissimilarity. An example of two such classes is rock and cohesionless sand. Since cohesionless sand can never attain a slope greater than its angle of repose (whereas rock, of course, can), features related to terrain slope were applied for range feature selection. In this work, range data was acquired from stereo imaging techniques. To compute range features in a scene, a 20 cm x 20 cm grid-based patch representation of the terrain surface was constructed. This patch size was selected to be similar to one rover wheel diameter. Best-fit planes were found within every patch using least-squares estimation, and the surface normal vector was extracted. The 3D range feature vector was then composed of the surface normal vector, along with the step height within the patch.

Vibration—Analysis of vibrations propagating through a rover's wheel/suspension structure can be used to distinguish between various types of terrain the rover is traversing [12]. This classification mode is unique among the low-level classifiers described here in that it relies on a "tactile" sensor signal that is modulated by physical rover-terrain interaction. The performance of such a classifier is not degraded by illumination variation, making it a potentially attractive complement to vision-based classification techniques. The general classification framework employed here is identical to that in [12]. Vibration signals were processed as the log power spectral density for every one-second time step at 557 frequencies in the frequency range 20.5 Hz to 12 kHz. For this work, a support vector machine with a linear kernel was used as the classifier.

3. DESCRIPTION OF HIGH LEVEL CLASSIFIERS

Low-level classifiers can yield poor results when applied individually in certain problem domains. Due to sensitivity to environmental changes (i.e. illumination) and measurement specifications (i.e. feature distance) poor classification performance is possible for low-level classifiers in some scenarios. Classifier fusion attempts to yield a robust class estimate despite the shortcomings of individual low level classifiers.

It should also be noted that since certain class distinctions are unobservable by individual low level classifiers, classifier fusion aims to overcome this problem by combining different sensing modes. Although this difference makes it more difficult to directly compare classifier performance, such increase in the number of detectable classes is a performance boost in itself.

Bayesian Classifier Fusion

Bayesian fusion was applied to merge the results of low-level classifiers. This technique has been proposed for classification of natural scenes with promising results [26]. Here, the low level MoG classifiers' outputs yield conditional class likelihoods. Posterior distributions of conditional class assignments are computed by Bayes' Rule, using the assumption that prior likelihoods are equal. Assuming that the visual features are conditionally independent, simple classifier fusion is applied as in Equation 3. Here $P(x_i|y_j)$ is the posterior probability of terrain class (x_j) given the sensing mode (y_j).

$$P(x_i \mid y_1, \ldots, y_n) = \prod_{j=1}^{j=n} P(x_i \mid y_j) \qquad (3)$$

However, this formulation implicitly requires that all classifiers function in the same class space (i.e the set x_j is same for all sensing modes). In the absence of this assumption, the class space of the final fusion is formed as the Cartesian product of the low-level class spaces, which yields a high number of non-physical terrain classes. Although previous researchers have addressed this problem with an unsupervised dimensionality reduction algorithm [26], this method did not exploit physical class knowledge that could be inherited from supervised classifiers. In this work the fusion class space was manually grouped into a

lower-dimensional space of physically meaningful terrain classes based on physical class knowledge of the Mars surface. Such a grouping explicitly encodes physical knowledge in the final class decisions.

Meta-classifier Fusion

A second approach to high-level classifier fusion is meta-classifier fusion. Meta-classifier fusion is a patch-wise classifier with features extracted from the outputs of low level classifiers. Specifically, it employs as features the continuous class likelihood outputs of the low-level classifiers

Meta-classifier fusion is very similar to stacked generalization (SG) presented by [27] and applied for road detection in [4]. In the method described here, low level classifiers described in Section 2 correspond to the "level-0 generalizer" where meta-classifier corresponds to "level-1 generalizer" of SG architecture. However, in the current work, the data points may not have the same resolution for all low-level classifiers. As described in Section 2, color- and texture-based classification was performed on a pixel-wise basis while range-based classifier was performed on a patch-wise basis. A trivial solution to this data association problem is addressed by a pixel to patch conversion. This conversion computes the continuous class likelihood of a patch by averaging the class likelihood values of every pixel in a particular patch. This high-level classifier is also a supervised classifier which requires training with a distinct set of training data than that employed by the low-level classifiers.

Data Fusion

A simple data fusion method was employed as a baseline to compare the performance of the Bayesian and meta-classifier fusion techniques and as a method for combining wheel vibration and vision data. Feature vectors from the various visual sensing modes are combined to form a single feature vector, which are then mapped to a probability distribution function using a MoG model. An SVM classifier was also applied to the data fusion framework. Note that the class space for data fusion included all observable classes, and SVM was implemented as a multi-class classifier.

Data fusion was also applied as an approach to combine vibration and vision data for improved local terrain classification accuracy. Here, images captured using a camera pointed at a rover wheel provided visual data corresponding to the terrain being sensed by a wheel-mounted vibration sensor, as seen in Figure 1. Visual data was represented as the mean RGB value of the pixels in a small region below the wheel. This 3-element vector was appended to the 557-element vibration vector using the data fusion framework, producing a 560-element combined vision/vibration vector. An SVM classifier was used to identify the local terrain class.

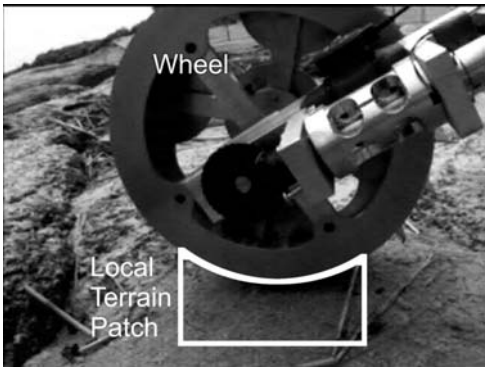

Figure 1: Image of wheel and terrain from belly-mounted camera

4. EXPERIMENTAL RESULTS

The performance of the low- and high-level classifiers was studied using images from NASA's Mars Exploration Rover mission and through experiments on a four-wheeled test-bed rover operating in Mars-analog terrain. These results are described below.

MER Imagery

Publicly available images from the MER mission's Spirit and Opportunity rovers were used to assess the performance of the low-level and high-level classifiers. Fifty-five images from the rovers' panoramic camera stereo pairs were selected from the Mars Analysts' Notebook database [28]. Ten images were used for classifier training and identifying meta-parameters. An additional five images were used for meta-classifier fusion and data fusion in addition to the training set to overcome data scaling problem. The remaining forty images were used to evaluate algorithm accuracy and computation time. For MER imagery, the vibration-based classification approach was not employed since only image data was available.

The MER panoramic camera pair has eight filters per camera; left filters mostly in the visible spectrum and right filters in the infrared region (with the exception of filter R1 at 430 nm). For color feature extraction, the combination of 4^{th} filter at 601 nm, 5^{th} filter at 535 nm, and 6^{th} filter at 482 nm intensities were chosen since they are near to the red, green and blue wavelengths, respectively. Texture feature extraction was performed on the intensity image from the 2^{nd} filter of the left camera at 753 nm. Range data was extracted by processing stereo pair images using stereo libraries developed at JPL [29].

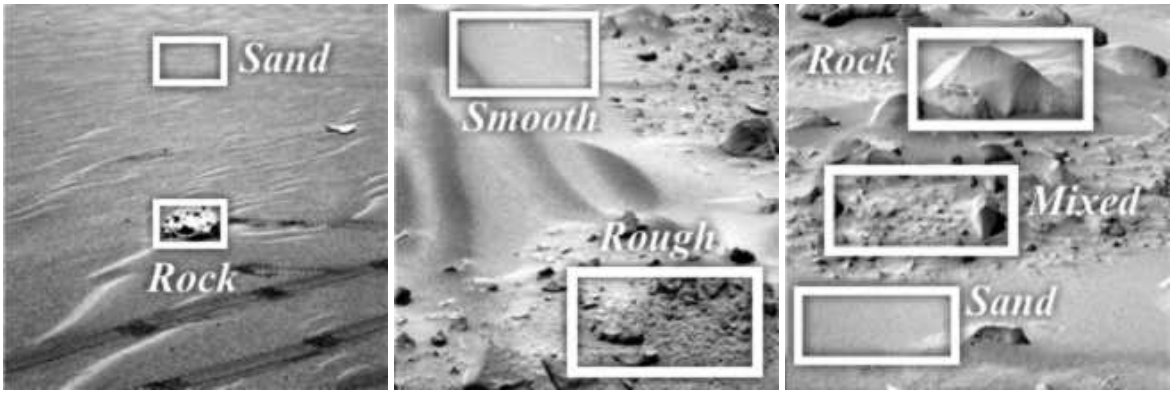

Figure 2: Class distinctions: color- and geometry based classes (left), texture-based classes (middle), fusion classes (right)

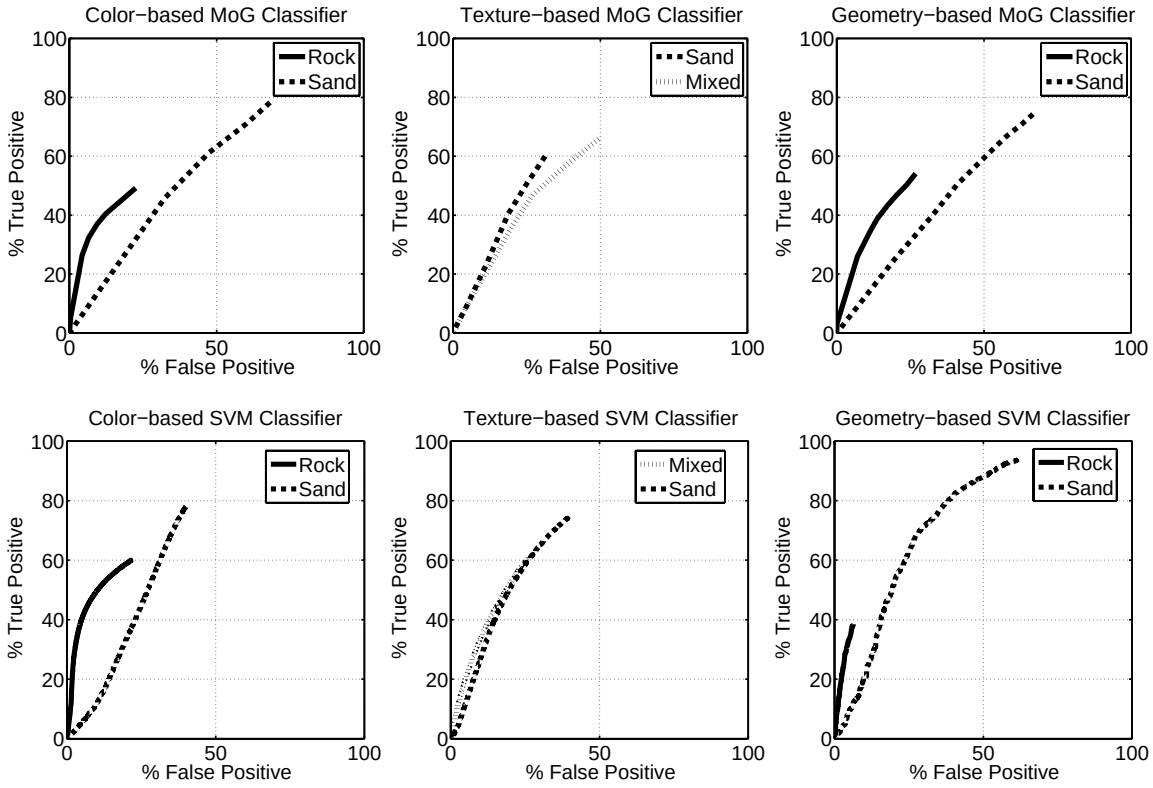

Figure 3: ROC curves of the low level classifier, MoG (top row), SVM (bottom row).

For Mars surface scenes, three primary terrain types that are believed to possess distinct traversability characteristics were defined: rocky terrain, composed of outcrop or large rocks; sandy terrain, composed of loose drift material and possibly crusty material; and mixed regions, composed of small loose rocks partially buried or lying atop a layer of sand. Examples of these terrains are shown in Figure 2 (right). High-level classifiers are expected to distinguish these three terrain classes; however, low-level classifiers can distinguish only a subset of them (Figure 2 left, middle). For instance, the color space of mixed terrain class, since it is composed of small rocks scattered on sand, overlaps with the color spaces of rock and sand terrain classes, so a color-based classifier cannot identify a distinct "mixed" terrain. Similarly, texture on the rock surfaces are not observable given the range of observation is 4 to 20 meters, so rock and sand both fall in the "smooth" class.

Low-level Classifier Results—Quantitative results of low level classifier are presented in Table 1 as average performances over the test set. The color-based classifiers produced results close to expectation of random choice between two classes on average. This might be expected due to the monochromatic nature of Martian surface. Texture-based classifier performed better than color since the discrimination between mixed and sandy terrain is more apparent. However, the performance for texture-based classification is still not sufficiently robust since texture classification accuracy is sensitive to the scaling of the image. Poor performance was observed in classifying

terrain outside a 4 to 20 meter range. The range-based classifier demonstrated the best performance, with 75% average classification accuracy, although variance was quite high. Failures in range-based classification were observed when sand was steeply sloped, forming ridges and dunes.

Table 1: Low-level classifier performance

		Average Accuracy (%)	95% Confidence Interval	Standard Deviation (%)
Color-based	MoG	57.2	[52.4 62.1]	15.6
	SVM	68.1	[63.4 72.7]	15.0
Texture-based	MoG	60.9	[56.1 65.7]	15.6
	SVM	66.7	[61.4 71.9]	16.8
Geometry-based	MoG	75.5	[69.0 82.1]	21.2
	SVM	70.2	[63.0 77.3]	23.0

Figure 3 shows ROC curves for each low-level classifier, illustrating the accuracy of the MoG and SVM classifiers across a range of confidence thresholds. These results demonstrate the weaknesses of the low-level classifiers. Besides being unable to distinguish between the three terrain classes of interest, low classification accuracy is exhibited due to the challenging nature of the classes. It should be observed that SVM and MoG classifiers demonstrated similar performance for each of the low-level sensing modes.

High-level Classifier Results—As described in Section 3, classifier fusion methods combine the data from multiple sensing modes to compute a class label. By merging the results of color- and range-based classifiers, fusion algorithms aim to compensate the weaknesses of low-level classifiers (e.g., to decrease the false positives of rock vs. sand detection). Moreover, inclusion of texture data enabled the observation of roughness and allows the definition of a "mixed" class.

Error! Reference source not found. shows ROC curves for the data fusion method applied with SVM and MoG as a multi-class classifier. As expected, data fusion performed poorly. This may be due to the difficulty of modeling in a high-dimensional feature space. In each case, it was observed that the classifier tend to have a bias towards a certain terrain class which yields poor average performance. These results also demonstrate the need for high-level classifier fusion for robust classification performance. Table 2 shows the comparison between the data fusion and classifier fusion methods in terms of global performance results.

Regarding the comparison between low- and high-level classifiers, note that high-level classifiers distinguish between three classes, whereas the low-level classifiers each distinguish between only two. Therefore the performance in terms of average accuracy is not directly comparable. However, it should me remembered that color- and texture-based classifiers perform close to the expectation of random choice whereas classifier fusion performance is much more robust.

Table 2: High-level classifier performances

		Average Accuracy (%)	95% Confidence Interval for	Standard Deviation (%)
Data Fusion	MoG	38.0	[32.5 43.5]	17.8
	SVM	47.0	[41.6 52.3]	17.3
Bayesian Fusion		64.7	[59.9 69.5]	15.5
Meta-classifier Fusion		59.6	[55.3 63.7]	13.6

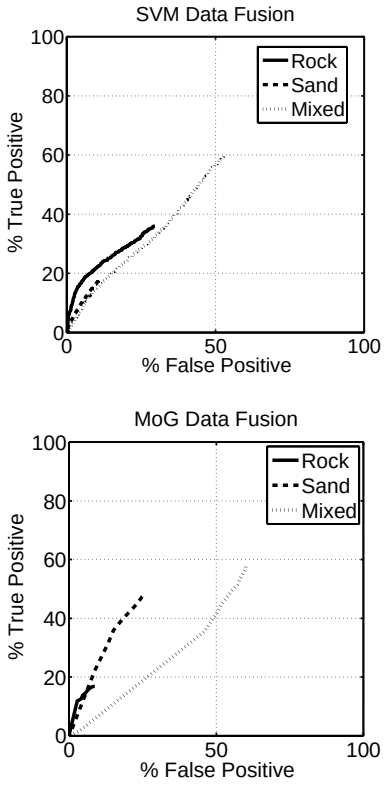

Figure 4: Data fusion ROC curves using SVM classifier (upper) and MoG classifier (lower)

Comparing high-level classifiers based on the ROC curves presented in Figure 5, it can be observed that Bayesian and meta-classifier fusion were much more accurate than data fusion. Although scaling of data (from pixel to patch)

potentially affects both data fusion and meta-classifier fusion, classifier fusion demonstrates better results than data fusion given the same amount of training data. For this data set, Bayesian fusion demonstrated similar accuracy to meta-classifier fusion. However, meta-classifier fusion requires more training data for the second level of classifier, besides the training set of low level classifiers. Bayesian fusion, on the contrary, does not require extra training for the second level, but the relationship between low-level classes and high-level classes has to be manually defined based on the environment setting. In short, there is a trade-off between predefining the class space and supplying additional training data for these fusion methods.

Wingaersheek Beach Experiments

Experimental Setup—Additional experiments were performed using a four wheeled mobile robot developed at MIT, named TORTOISE (all-Terrain Outdoor Rover Test-bed for Integrated Sensing Experiments), shown in Figure 6. TORTOISE is an 80-cm-long x 50-cm-wide x 90-cm-tall robot with 20 cm diameter wheels. The TORTOISE sensor suite includes the following: a forward looking mast-mounted Videre Design "dual DCAM" stereo pair with 640 x 480 resolution; a belly-mounted color monocular camera with 320 x 240 resolution to observe local terrain; and a Signal Flex SF-20 contact microphone mounted on the rover suspension near the front right wheel assembly to sense vibrations. During experiments, TORTOISE traveled at an average speed of 6 cm/sec. It captured monocular images at 2Hz, and vibration data at 44.1 kHz. Stereo images were captured every 1.5 seconds.

Experiments were performed at Wingaersheek Beach in Gloucester, MA. This is an oceanfront environment dominated by large (i.e. meter-scale) rock outcrops and distributions of rover-sized and smaller rocks over sand. Neighboring areas exhibit sloped sand dunes and sandy flats mixed with beach grass. Figure 7 shows a typical scene from the experiment site. This scene shows a large rock in the foreground and scattered, partially buried rocks in the middle range. Sand appears grayish in color while rock features vary from gray to light brown and dark brown. This test site was chosen because of its visual and topographical similarities to Mars surface scenes.

For the following experiments, the terrain classes of interest were "rock," "sand," and "beach grass." The "mixed" class was not defined due to lack of scattered small sized rocks; dry beach grass was used to reflect a distinct texture signature in an effort to maintain a consistent number of classes with MER results.

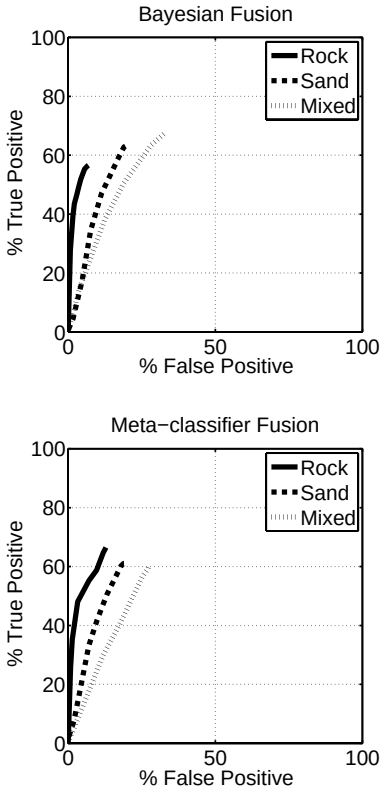

Figure 5: ROC curves for Bayesian fusion (upper) and meta-classifier fusion (lower)

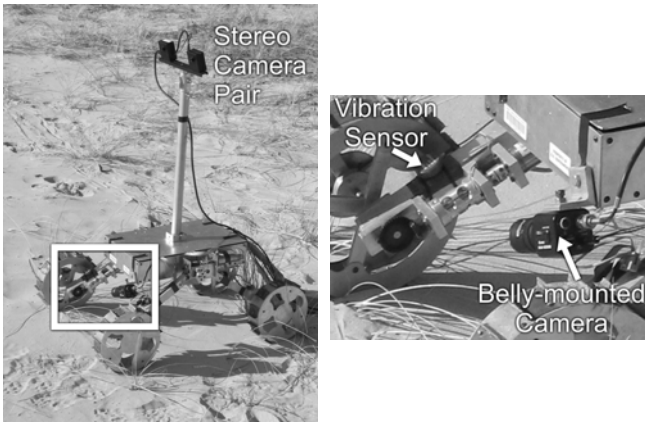

Figure 6: TORTOISE experimental rover (left), local sensing suite (right)

Figure 7: Sample scene from Wingaersheek Beach

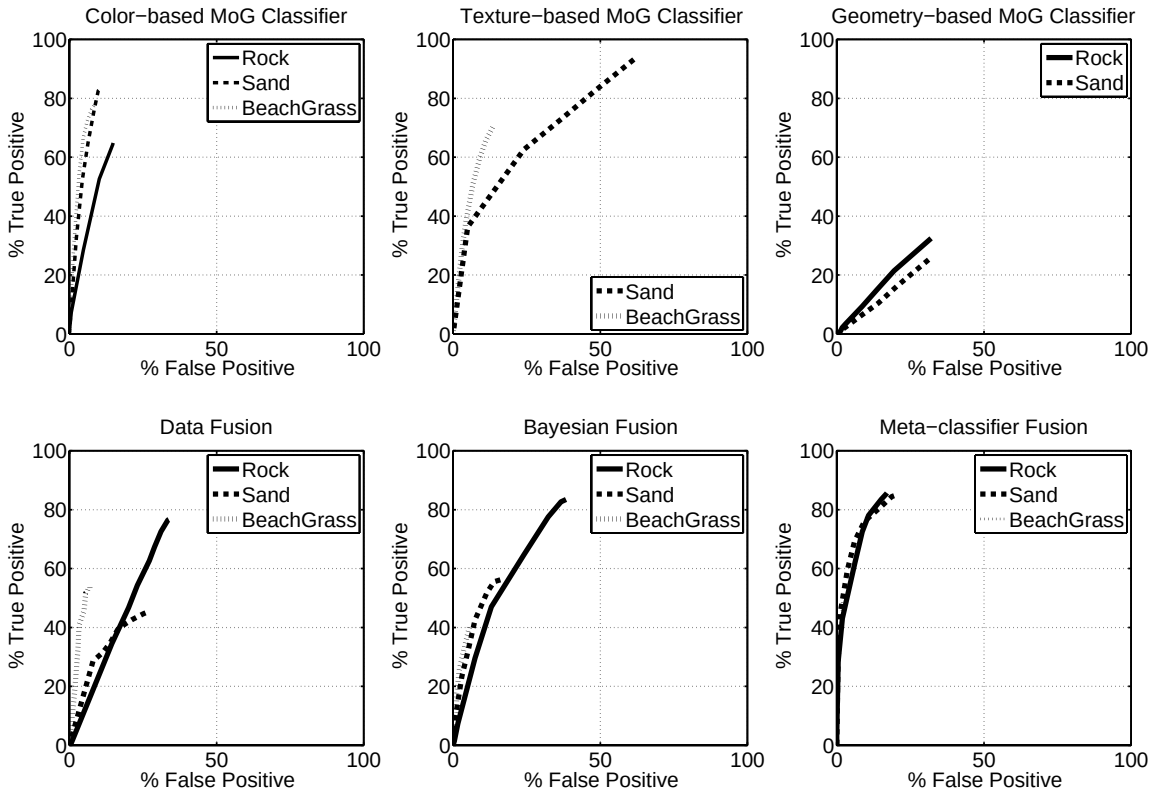

Figure 8: ROC curves: Low-level classifiers (top row), high-level classifiers (bottom row)

Low-level Classifier Results—Six days of experiments were conducted with a total of approximately 50 traverses and a total distance traveled of 500 meters. Every traverse included approximately 250 images. Every 20th image was included in the test set to minimize overlap. Data from the first traverse of the final day was used for training data. Classifier accuracy was assessed using images from the remaining traverses on the final day. The performance of the low-level classifiers is shown in Figure 8 as series of ROC curves.

It was observed that the performance of the color-based classifier was improved over that observed in experiments on MER imagery. This was likely due to the greater color variation present in an average beach scene. Relatively poor results were observed from the range-based classifier. The reason for this decrease in performance may be related to the poor accuracy and resolution of stereo-based range data for these experiments relative to MER imagery data, which used state-of-the-art JPL stereo processing software operating on high-quality images. This performance decline illustrates the sensitivity of range-based classification to data quality, and strengthens the motivation for classifier fusion.

High-level Classifier Results—High-level classifier performance is shown in Figure 8. In keeping with the MER results, the classifier fusion methods perform significantly better than the data fusion approach. Data fusion exhibits a bias towards the "rock" class yielding high false positives and degrading the detection rate for other classes. In this experiment setting, use of high-level classifiers does not increase the number of observable terrain classes since the color-based classifier is able to distinguish all terrain classes present in the setting. However, the ROC curves show a performance increase as a result of merging texture- and range- based classifiers with color-based results. In the meta-classifier fusion results, it is clear that although individual performances of other low-level classifiers are below color-based results, they contribute to the training of meta-classifier yielding improved results.

Data Fusion for Local Terrain—Local classification of terrain based on fusion of vibration and color features was tested using data captured by the vibration sensor and belly-mounted camera. These data were collected while the rover traversed sand, beach grass, and rock. A total of 21 minutes of vibration data were collected (1260 one-second segments), with over 2500 associated local images. Half of the data was used for establishing the meta-parameters and training each SVM classifier. The other half was used to test the classifiers.

The results for local terrain classification are shown in Figure 9. The left plot shows results for pure vibration-based classification. It can be seen that all terrains are moderately well distinguished, with an average accuracy of 65% at full classification. The center plot shows results for

pure color-based classification. Here "beach grass" is nearly all detected, with very few false positives. "Rock" and "sand" are also well distinguished. The average accuracy is 77% at full classification. Finally, the right plot shows the results for data fusion of color and vibration. An improvement over vibration-only and color-only classifiers was exhibited, with an average accuracy of 84%. This result suggests improved classification performance can be derived from fusion of visual and tactile information. This is likely due to the insensitivity of tactile features to variations in illumination.

Computation Times

All algorithms in this work except SVM classification were implemented in Matlab. On a Pentium 1.8 GHz desktop computer, pixel-wise MoG classification of a 512 x 512 image took an average of 5.2 seconds. Patch-wise MoG classification (for range-based, data fusion and meta-classifier fusion) required an average of 2.4 seconds. Bayesian fusion took 1.2 seconds to form classifier decisions. The most computationally expensive element of the algorithms is texture feature extraction, requiring approximately 14.8 seconds of computation time for three levels of Haar wavelet transforms and computing the pixel-wise texture signature of 512 x 512 grayscale image. In total, classifying a 512 x 512 frame takes approximately 29.0 sec/frame. These times could be significantly reduced in a C-code implementation.

SVM classification was implemented with C++, using the LIBSVM library, with additional optimization for linear kernels (Chih-Chung & Chih-Jen, 2001). Classification of a 512x512 color image took an average of 0.61 seconds using a linear kernel. Classification using a Gaussian kernel took an average of 77.5 seconds for a 512x512 color image.

After feature extraction, texture classification times were identical to those for color classification. Patch-wise classification (for range and data fusion) averaged less than 0.01 seconds per patch for the linear SVM, and less than 0.04 seconds per patch for the Gaussian SVM. The number of patches in each image varied from 10 to 400.

5. CONCLUSION

Knowledge of the physical properties of terrain surrounding a planetary exploration rover can be used to allow a rover system to fully exploit its mobility capabilities. The ability to detect or estimate terrain physical properties would allow a rover to predict its mobility performance and knowledge of terrain properties could allow a system to adapt its control and planning strategies to enhance performance.

This paper has compared the performance of various methods for terrain classification based on the fusion of visual and tactile features. It was shown that classifier fusion methods can improve overall classification performance in two ways compared to low-level methods. First, classifier fusion yielded a more descriptive class set than any of the low-level classifiers can attain individually. Second, the rate of false positives decreased significantly while the rate of true positives increased. This shows that in challenging planetary surfaces, stand alone visual features are may not be sufficiently robust for mobile robot sensing; however, classifier fusion techniques improve sensing performance significantly.

Future research will focus on integrating additional tactile sensing modes such as wheel sinkage and torque with visual classifier fusion algorithms.

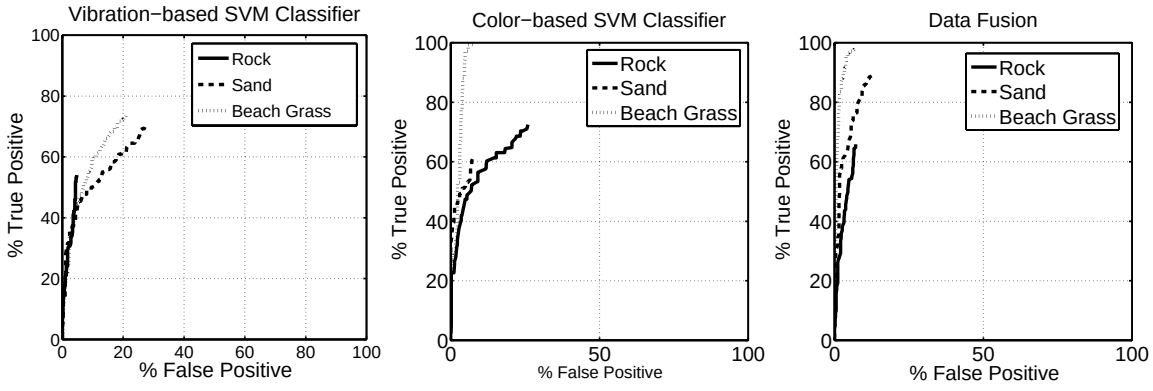

Figure 9: Classifier results for local vibration-based classification (left), color-based classification (middle), and data fusion of color and vibration (right)

ACKNOWLEDGMENT

This work was supported by the NASA Jet Propulsion Laboratory (JPL) through the Mars Technology Program.

REFERENCES

[1] Urquhart, M. and Gulick, V (2003). "Lander detection and identification of hydrothermal deposits," abstract presented at *First Landing Site Workshop for MER*.

[2] Iagnemma, K. and Dubowsky, S. (2002, March). "Terrain estimation for high speed rough terrain autonomous vehicle navigation," *Proceedings of the SPIE Conference on Unmanned Ground Vehicle Technology IV*.

[3] Kelly, A., et al. (2006, June). "Toward Reliable Off Road Autonomous Vehicles Operating in Challenging Environments," *The International Journal of Robotics Research*. 25(5/6).

[4] Dima, C.S., Vandapel, N., and Hebert, M. (2004). "Classifier fusion for outdoor obstacle detection," *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 1, 665-671, doi: 10.1109/ROBOT.2004.1307225.

[5] Manduchi, R., Castano, A., Thalukder, A., and Matthies, L. (2005, May). "Obstacle detection and terrain classification for autonomous off-road navigation," *Autonomous Robots*, 18, 81-102.

[6] Squyres, S. W., et al., (2003). "Athena Mars rover science investigation," J. Geophys. Res., 108(E12), 8062, doi:10.1029/2003JE002121.

[7] Rasmussen, C., (2001, December). "Laser Range-, Color-, and Texture-based Classifiers for Segmenting Marginal Roads," in *Proceedings of. Conference on Computer Vision & Pattern Recognition Technical Sketches*, Kauai, HI.

[8] Angelova, A., Matthies, L., Helmick, D., Sibley, G., Perona, P. (2006). "Learning to predict slip for ground robots," *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, Orlando, Florida. May, 2006.

[9] Bellutta, P., Manduchi, R., Matthies, L., Owens, K. and Rankin,K. (2000, October). "Terrain perception for Demo III," *Proceedings of the Intelligent Vehicles Symposium*, 326-331, doi: 10.1109/IVS.2000.898363

[10] Vandapel, N., Huber, D.F., Kapuria, A., Hebert, M. (2004). Natural Terrain Classification using 3-D Ladar Data. *Proceedings of the International Conference on Robotics and Automation (ICRA)*, 5, 5117- 5122.

[11] Mandelbaum, R., McDowell, L., Bogoni, L., Reich, B., and Hansen M. (1998). "Real-Time Stereo Processing, Obstacle Detection And Terrain Estimation Form Vehicle-Mounted Stereo Cameras," *Proceedings of the 4th IEEE Workshop on Applications of Computer Vision,* 288, Princeton, New Jersey.

[12] Brooks, C. and Iagnemma, K. (2005). "Vibration-based Terrain Classification for Planetary Rovers," *IEEE Transactions on Robotics*, 21, 6, 1185-1191.

[13] Sadhukhan, D., Moore, C., and Collins, E. (2004). "Terrain Estimation Using Internal Sensors," in *Proceedings of International Conference on Robotics and Applications* (*IASTED)*, 84(11), 1684-1704, doi: 10.1109/5.542415.

[14] Ojeda, L., Borenstein, J., Witus, G., and Karlsen, R. (2006). "Terrain characterization and classification with a mobile robot," *Journal of Field Robotics*, 23(2), 103-122, doi: 10.1002/rob.20113.

[15] Castano, R., et al. (2005). "Current Results from a Rover Science Data Analysis System," *Proceedings of 2005 IEEE Aerospace Conference,* Big Sky. 356-365, doi: 10.1109/AERO.2005.1559328.

[16] Gor,V., Castaño, R., Manduchi, R., Anderson, R., and E. Mjolsness (2001). "Autonomous Rock Detection for Mars Terrain," *Space 2001, AIAA*.

[17] Thompson, D. R., Niekum, S., Smith, T. and Wettergreen, D. (2005). "Automatic Detection and Classification of Features of Geologic Interest," *Proceedings. of IEEE Aerospace Conference,* 366-377, doi: 10.1109/AERO.2005.1559329.

[18] McGuire, P. C., et al., (2005). "The Cyborg Astrobiologist: scouting red beds for uncommon features with geological significance," International Journal of Astrobiology, 4, 101-113.

[19] Dima, C.S., Vandapel, N., and Hebert, M. (2003). "Sensor and classifier fusion for outdoor obstacle detection: an application of data fusion to autonomous road detection," *Applied Imagery Pattern Recognition Workshop*, 255- 262, doi: 10.1109/AIPR.2003.1284281.

[20] Bishop, C.M., (1995). *Neural networks for pattern recognition.* New York: Oxford University Press.

[21] Bilmes, J. (1997). "A Gentle Tutorial on the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models," Technical Report, University of Berkeley.

[22] Vapnik, V.N. (1995). *The Nature of Statistical Learning Theory*. New York: Springer.

[23] Chih-Chung C. and Chih-Jen L, (2001). LIBSVM: a library for support vector machines. Software retrieved January, 2006 available at http://www.csie.ntu.edu.tw/~cjlin/libsvm

[24] Goldberg, S., Maimone, M., and Matthies, L. (2002). "Stereo vision and rover navigation software for planetary exploration," *IEEE Aerospace Conference,* Big Sky, 5, 2025-2036, doi: 10.1109/AERO.2002.1035370.

[25] Espinal, F., Huntsberger, T.L., Jawerth, B., and Kubota T. (1998). "Wavelet-based fractal signature analysis for automatic target recognition," *Optical Engineering, Special Section on Advances in Pattern Recognition*, 37(1), 166-174.

[26] Manduchi, R. (1999). "Bayesian Fusion of Color and Texture Segmentations," *In Proceedings. of International Conference on Computer Vision (ICCV)*, 2, 956-962, doi: 10.1109/ICCV.1999.790351.

[27] Wolpert, D. H. (1990).Stacked generalization,. Los Alamos, NM, Tech. Rep. LA-UR-90-3460, 1990.

[28] Mars Analyst's Notebook (2006). Retrieved May 24, 2006, from http://anserver1.eprsl.wustl.edu/.

[29] Ansar, A., Castano, A., and Matthies, L. (2004, September). "Enhanced real-time stereo using bilateral filtering**,"** *2nd International Symposium on 3D Data Processing, Visualization, and Transmission,* 455-462., doi: 10.1109/TDPVT.2004.1335273

BIOGRAPHY

Ibrahim Halatci *is a technical support engineer in the Engineering Development Group at the Mathworks Inc. He has recently received his MS degree from the Mechanical Engineering department of the Massachusetts Institute of Technology. He received his B.S. degree with honor in Mechatronics Engineering from Sabanci University in 2004. His research interests include control systems, its application to robotics and learning for mobile robots*

Christopher Brooks *is a graduate student in the Mechanical Engineering department of the Massachusetts Institute of Technology. He received his B.S. degree with honor in engineering and applied science from the California Institute of Technology in 2000, and his M.S. degree from the Massachusetts Institute of Technology in 2004. Hs research interests include mobile robot control, terrain sensing, and their application to improving autonomous robot mobility. He is a member of Tau Beta Pi.*

Karl Iagnemma is a principal research scientist in the Mechanical Engineering department of the Massachusetts Institute of Technology. He received his B.S. degree summa cum laude in mechanical engineering from the University of Michigan in 1994, and his M.S. and Ph.D. from the Massachusetts Institute of Technology, where he was a National Science Foundation graduate fellow, in 1997 and 2001, respectively. He has been a visiting researcher at the Jet Propulsion Laboratory. His research interests include rough-terrain mobile robot control and motion planning, robot-terrain interaction, and robotic mobility analysis. He is author of the monograph Mobile Robots in Rough Terrain: Estimation, Motion Planning, and Control with Application to Planetary Rovers (Springer, 2004). He is a member of IEEE and Sigma Xi.