Vegetation Detection for Mobile Robot Navigation

David M. Bradley, Scott M. Thayer, Anthony Stentz, and Peter Rander

CMU-RI-TR-04-12

February, 2004

Robotics Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213

Abstract

Described is the development and testing of a robust vegetation detector for mobile robot navigation. A multispectral sensor was created out of a near-infrared and a visible light video camera. Vegetation was then detected by subtracting each pixel in the red channel of the visible-light image from the corresponding pixel in the near-infrared image and thresholding the result. This computationally-efficient technique has been verified to be a robust chlorophyll detector in natural environments.

This work was sponsored by DARPA, under contract ”Perception for Off-Road Mobility (PerceptOR)” (contract number MDA972-01-9-0016). The views and conclusions contained in this document are those of the authors and should not be interpreted as representing official policies or endorsements, expressed or implied, of the U.S. Government.

Contents

1 Introduction

Current autonomous navigation techniques work well for environments such as hallways and on roads, where obstacles are static and usually rigid. In these cases, size and shape are sufficient for determining which obstacles can be driven over and which need to be avoided. In off-road driving, however, the assumption that every obstacle is rigid and lethal quickly creates problems. In situations such as a cornfield or a field of thick, tall grass, there may be dense geometric obstacles on all sides of the robot. This can lead to frustrating behaviors where the robot attempts an unnecessary detour because, for example, it happens to see a single blade of grass as a lethal object. Although a large vehicle would easily be able to drive through tall grass or a cornfield, it must first be able to reliably discriminate between a stand of grass and a roll of barbed wire, or between cornstalks and thin trees. As a general rule, the more chlorophyll a material has, the easier it is to drive through. Grass is easy to drive through while bramble is not. Leaves offer far less resistance than tree trunks. Rapid and accurate detection of chlorophyll-rich vegetation can help find the path of least resistance through a cluttered environment.

Methods have been developed to detect vegetation from 3-D point clouds [5, 11], but current 3-D laser scanners limit the usefulness of this approach at longer ranges. Accurate classification requires fairly dense point clouds which are hard to acquire at medium to long ranges with LADAR because of beam divergence, and scanning speed limitations. Point cloud acquisition problems become worse as the average speed of the vehicle increases. Higher speed operation requires classifying regions farther from the vehicle, so that the vehicle has time to stop if it detects an obstacle in its path. Vehicle stopping distance increases as the square of speed, and in order to maintain a particular angular field of view the robot must classify an area of stopping distance times speed every second. Covering this area with a uniform density of points means that the 3D point acquisition speed must be $O(s^3)$, where s is the vehicle speed. Additionally, scanning sensors such as LADAR that do not sense their entire field of view simultaneously need an accurate estimate of their current pose in order to register data accurately as the robot moves. Not only do higher speeds make jarring (and hard to measure) bumps more likely, but they also require measuring points farther ahead, which amplifies the effects of any errors in pose. New technologies such as flash LADAR may alleviate registration problems, but passive sensors may still be preferable to enhance stealth. 2-D cameras provide a passive means of rapidly collecting data from large regions of terrain.

Fortunately there are well-established techniques in the field of satellite photogrammetry for measuring chlorophyll content using a multi-spectral camera [2, 4, 6, 7, 8, 9, 16, 13]. A simple pixel-by-pixel comparison between red and Near-InfraRed (NIR) reflectance provides a powerful and robust way to detect vegetation. Although the viewpoint of a satellite is drastically different from that of a mobile robot, the technique is still effective despite additional complications such as shadows and views of the sky.

This paper describes how to build a multi-spectral camera out of two video cameras, and shows that it can robustly detect vegetation over a wide range of vegetation health and weather conditions. Sections 2.1, 2.2 and 2.3 give an overview of spectral

vegetation detection and its application to navigation. Simulation results validating the approach are presented in Section 2.4. The constructed system is described in Section 3.1, and field results are shown in section 3.2. Section 3.3 describes an experiment with an alternative sensor setup consisting of a single CCD camera combined with the reflectance channel of a laser scanner, and discusses the relative merits and disadvantages of this approach. Finally, section 4 provides an analysis of the results.

2 Spectral Vegetation Detection

2.1 Physical Principles

The spectral properties of chlorophyll-rich vegetation are primarily determined by the absorption spectra of water and chlorophyll, and the refraction of light at cell walls [16]. The water present in cells absorbs light with wavelengths longer than 1400 nm. Chlorophyll strongly absorbs visible light, especially red and blue wavelengths [2]. The remaining light is efficiently scattered by the critical internal reflection caused by the change in refractive index from water to air at the cell wall. As a result, those wavelengths between 800 nm and 1400 nm that escape both water and chlorophyll are strongly reflected in all directions.

The sharp difference between the reflectance of vegetation at 670nm (red) and at 1000nm (NIR) has long been exploited in the field of satellite remote sensing. Kauth and Thomas [8] noticed that plotting NIR reflectance against Red reflectance for satellite images produced a scatter diagram with a line of points formed by pixels containing bare soil, and a cluster of points from pixels completely covered with vegetation. Points with a mixture of vegetation and soil appear between the soil line and the vegetation point. They called this the ”tasseled cap-agram” because of it’s resemblance to a tasseled cap. Figure 1 shows this scatter plot created from one of our images. Because our camera also includes a view of the sky, our scatter plot contains a blue sky line (marked in blue) as well as the soil line (marked in red) and the vegetation point (marked in green). Clouds blend into the soil line, but are still very distinct from vegetation. Pixels containing vegetation and blue sky are remarkably well separated from everything else.

2.2 Vegetation Indices

Numerous schemes have been proposed to segment the tasseled cap plot. Jordan [7] assumed that lines of equal vegetation all intersect at the origin and developed the Ratio Vegetation Index (RVI):

$$\mathrm{RVI} = \frac{\rho_{\mathrm{NIR}}}{\rho_{\mathrm{RED}}} \tag{1}$$

The RVI measures the slope of the line between the origin of red-NIR space and the red-NIR value of the pixel. The infinite range of the RVI is cumbersome, however, and it has been replaced by the Normalized Difference Vegetation Index (NDVI) which is functionally equivalent and only varies from -1 to 1 [9].

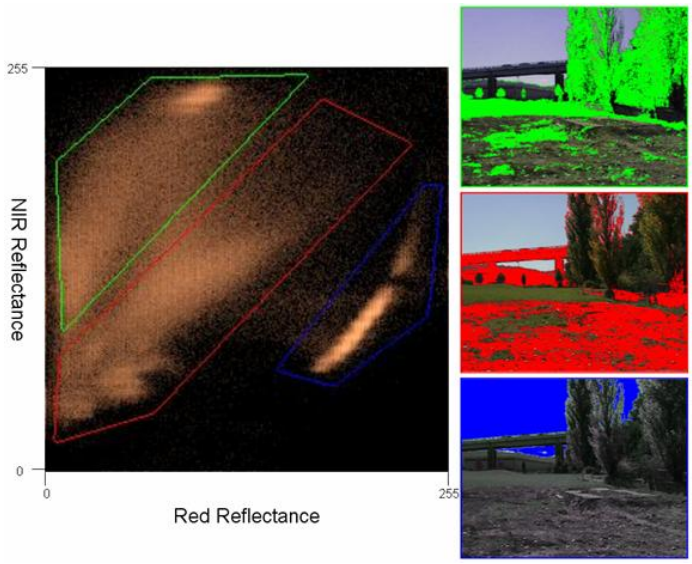

Figure 1: Scatter plot of NIR reflectance vs. red reflectance for all pixels in a typical image. Different regions in the scatterplot clearly correspond to different types of pixels in the image. Pixels in the green region correspond to vegetation, and pixels in the blue region correspond to sky.

$$\text{NDVI} = \frac{\rho_{\text{NIR}} - \rho_{\text{RED}}}{\rho_{\text{NIR}} + \rho_{\text{RED}}} \tag{2}$$

Crippen [4] proposed a more computationally efficient version of the NDVI known as the Infrared Percentage Vegetation Index (IPVI), which eliminated the unnecessary red radiance subtraction:

$$\text{IPVI} = \frac{\rho_{\text{NIR}}}{\rho_{\text{NIR}} + \rho_{\text{RED}}} = \frac{\text{NDVI} + 1}{2} \tag{3}$$

The other main approach to vegetation detection is to measure the distance in the scatter plot from the soil line. Richardson and Wiegand [13] pursued this approach with their Perpendicular Vegetation Index (PVI):

$$\text{PVI} = sin(\alpha)\rho_{\text{NIR}} - cos(\alpha)\rho_{\text{RED}} \tag{4}$$

Where α is the angle between the soil line and the NIR axis. A common special case of the PVI is when α equals $45°$. Here the PVI simplifies to what has been called the Vegetation Index [10] or the Difference Vegetation Index (5).

$$\text{VI} = \text{DVI} = \rho_{\text{NIR}} - \rho_{\text{RED}} \tag{5}$$

Attempts have also been made to compensate for variations in the soil background, most notably the Soil Adjusted Vegetation Index (SAVI) proposed in [6] (6):

$$\text{SAVI} = \frac{(\rho_{\text{NIR}} - \rho_{\text{RED}})(1+L)}{\rho_{\text{NIR}} + \rho_{\text{RED}} + L} \tag{6}$$

Where L is selected based on where isovegetation lines intersect the soil line. Qui et al. provide a formula for automatically determining L from the current image data in their Modified Soil Adjusted Vegetation Index (MSAVI) [12]. The closed form of this formula is known as MSAVI2 (7):

$$\text{MSAVI2} = \frac{2(\rho_{\text{NIR}}+1) - \sqrt{(2\rho_{\text{NIR}}+1)^2 - 8(\rho_{\text{NIR}} - \rho_{\text{RED}})}}{2} \tag{7}$$

2.3 Applying Chlorophyll Detection to Navigation

Once the chlorophyll content of the scene is determined, it has to be incorporated somehow into the robot's cost map. There is a continuum of possible approaches for utilizing the Chlorophyll content information between these two extremes:

1. Use the red and NIR responses directly to estimate the cost of each obstacle in the scene. Essentially this involves estimating the joint probability distribution of the cost C of the obstacle given the sensor readings, $f_C(C = c|\rho_{\text{NIR}}, \rho_{\text{RED}}, \Theta)$, where Θ is a vector containing all of the other relevant sensor readings. This approach has the disadvantage of potentially requiring vast amounts of training data.

2. Perform a binary classification of obstacles as either vegetation or non-vegetation. This approach conveys less information, but is more robust to noise and requires less training data.

Because of the variability of outdoor lighting and the limited dynamic range of the cameras used, this work has focused on binary classification. When used for binary classification all of vegetation indices discussed in section 2.2 produce linear decision boundaries in the red-NIR plane (table 1 and figure 2). DVI corresponds to a line with a constant slope of one and a variable intercept. NDVI is the reverse, with a constant intercept of zero and a variable slope. MSAVI2 varies both slope and intercept but maintains a fixed relationship between them. Finally, PVI allows arbitrary slopes and intercepts. Nonlinear classifiers were not investigated as figure 3 suggests that they would provide little added benefit.

2.4 Validation of Approach

To ensure robust performance across a variety of different types of vegetation and non-vegetation, the vegetation detection system was simulated and tested on the USGS digital spectral library [3]. This library contains the spectral signatures of over 800 different materials. Many of these spectra are from rare minerals that are not relevant to our application. The relevant subset of the library used for testing contained 105 common types of vegetation and 169 common soil mixtures, artificial materials, and

Index	Equation of Decision Boundary
DVI	$\rho_{NIR} = \rho_{RED} + T$
NDVI	$\rho_{NIR} = \left(\frac{T+1}{1-T}\right)\rho_{RED} + T$
PVI	$\rho_{NIR} = cotan(\alpha)\rho_{RED} + cosec(\alpha)T$
MSAVI2	$\rho_{NIR} = \left(\frac{-2}{3-2T}\right)\rho_{RED} - \left(4T^2 - 8T + 3\right)$

Table 1: Equations for the linear decision boundaries of selected vegetation indices, where T is the threshold used and α is the angle with the NIR axis.

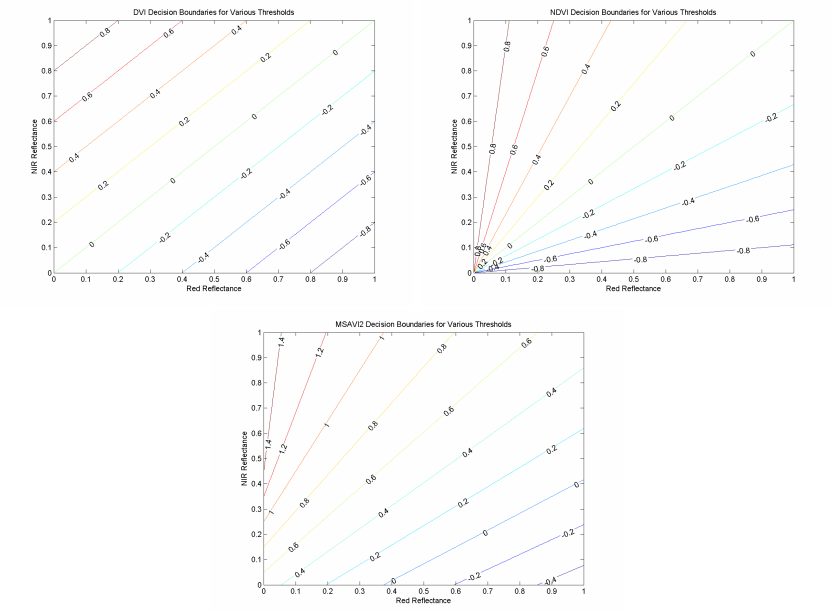

Figure 2: Binary classification decision boundaries produced by DVI (top left), NDVI (top right) and MSAVI2 (bottom center). Pixels above and to the left of a boundary are classified as vegetation. Each boundary is labeled with the corresponding threshold value.

coatings. Figure 3 shows the distribution of these materials over the red (600-670nm) and NIR (900-1000nm) spectral bands measured by the fielded system. The distribution of NDVI and DVI values of these materials shows that both indices are quite useful for detecting vegetation (figure 4) but neither is optimal. The PVI, with its ability to select both slope and intercept, classifies the most materials correctly (258 of 274 materials). A yellow aspen leaf and dry grass are misclassified as non-vegetation (figure 5). Various man-made materials and natural mixtures involving hematite are misclassified as vegetation (figure 6).

Unfortunately our camera system suffers from two sources of noise not present in the USGS spectral signatures. First, the spectral signatures were collected under ideal

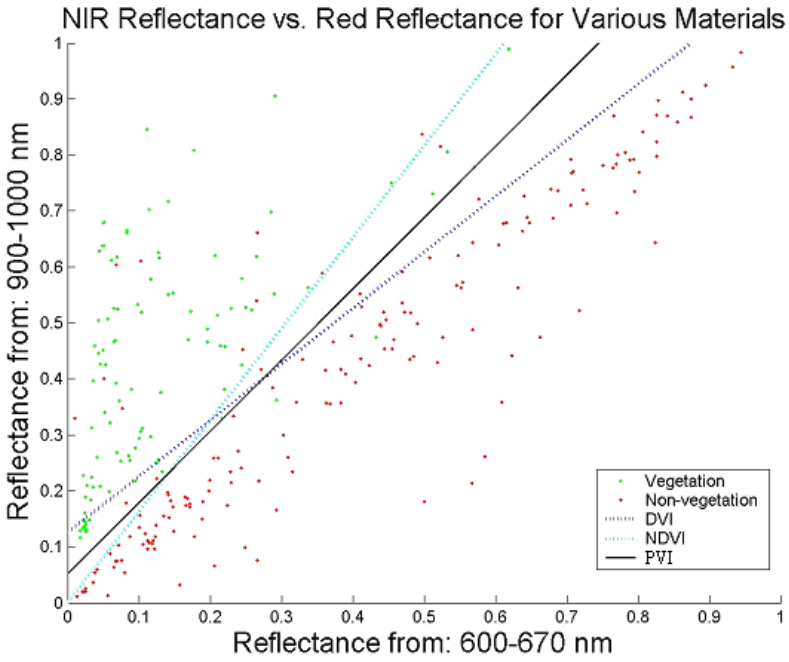

Figure 3: Scatter plot of near infrared reflectance vs. red reflectance for 274 materials commonly found outdoors. Also shown are the decision boundaries of the best DVI, NDVI, and PVI-based classifiers. Data generated from the USGS digital spectral library [3].

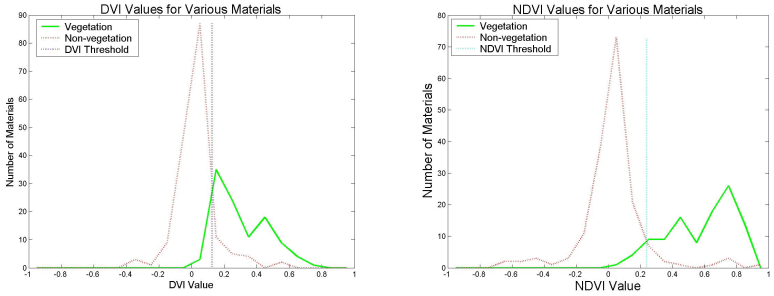

Figure 4: Distribution of the NDVI and DVI values for 274 materials commonly found outdoors.

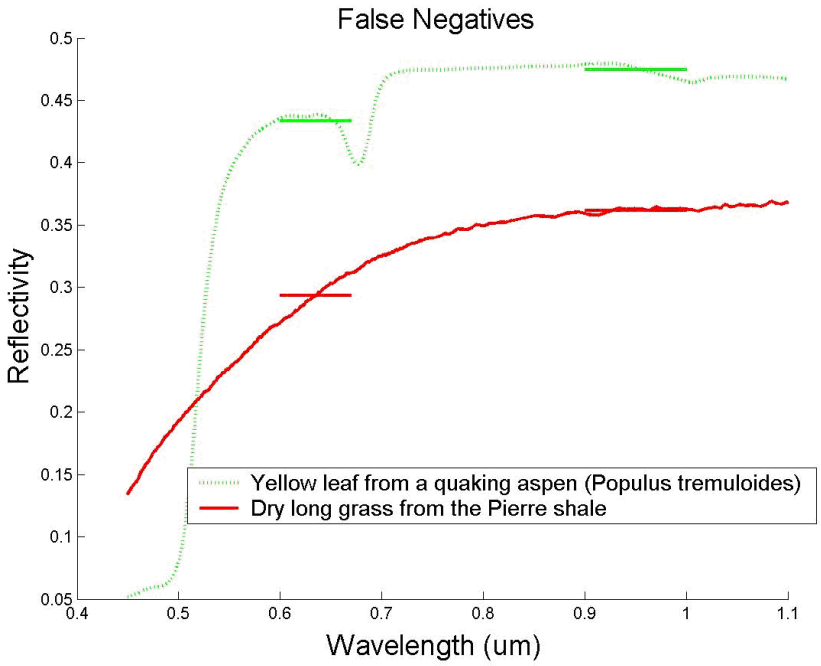

Figure 5: Vegetation spectra misclassified as non-vegetation by an optimal linear discriminator using the red and NIR bands (600-670nm and 900-1000nm). Horizontal line segments indicate the simulated camera responses shown in figure 3.

illumination conditions, which are rare outdoors. Sunny days in particular present a challenge as our cameras do not have the dynamic range to make accurate measurements of both shadowed and brightly lit areas simultaneously. Second, because of gamma correction, variable gain offset between the visible and NIR cameras, and various other factors our camera setup does not respond linearly. Despite these problems, the USGS dataset provides a valuable tool for identifying materials on the decision boundary between the vegetation and non-vegetation classes. These materials, or artificial calibration targets with similar responses, can then be used to tune the vegetation detector to the actual camera and weather conditions.

3 Field Results

3.1 Sensor Design

To create our multi-spectral sensor we used two Sony FCB-EX480 color cameras. One camera is used for visible light and one for NIR by removing the NIR-cut filter. The optical axes of both cameras are aligned using a cold mirror as a beamsplitter (see figure 7). The cold mirror reflects visible light while transmitting IR radiation (see

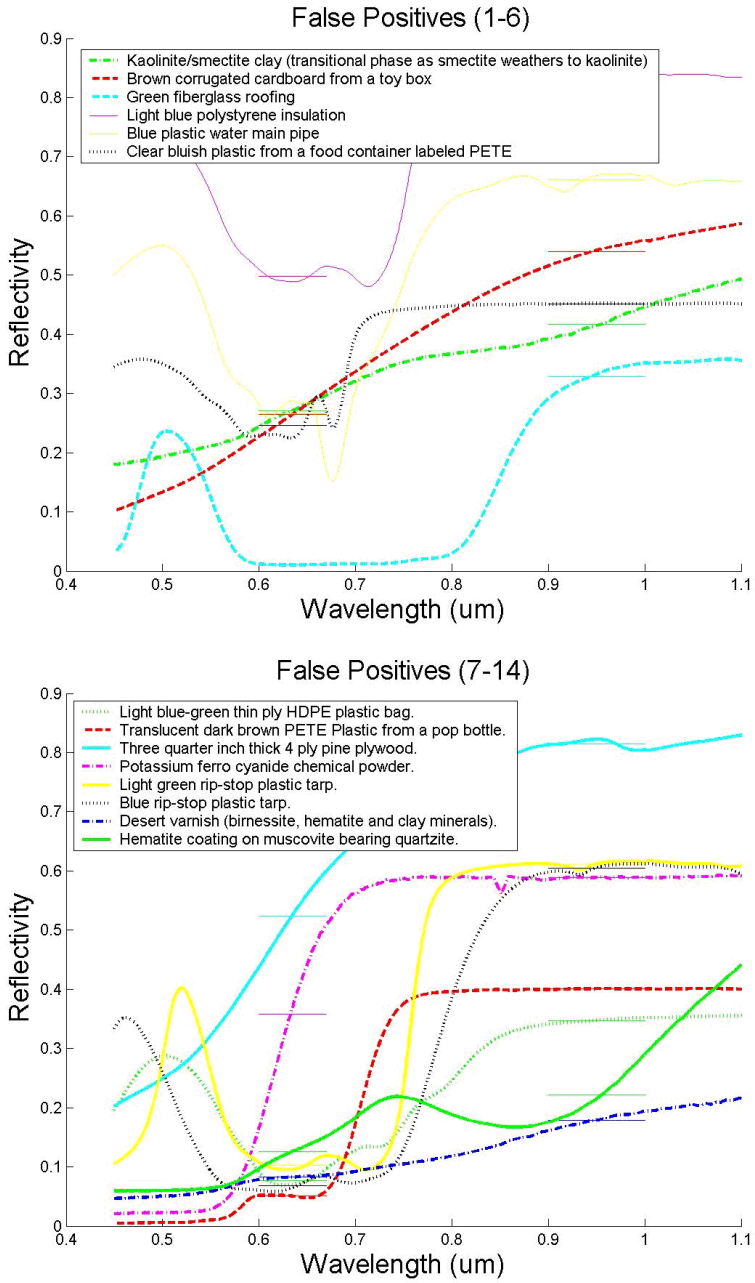

Figure 6: Non-vegetation spectra misclassified as vegetation by an optimal linear discriminator using the red and NIR bands (600-670nm and 900-1000nm). Horizontal line segments indicate the simulated camera readings.

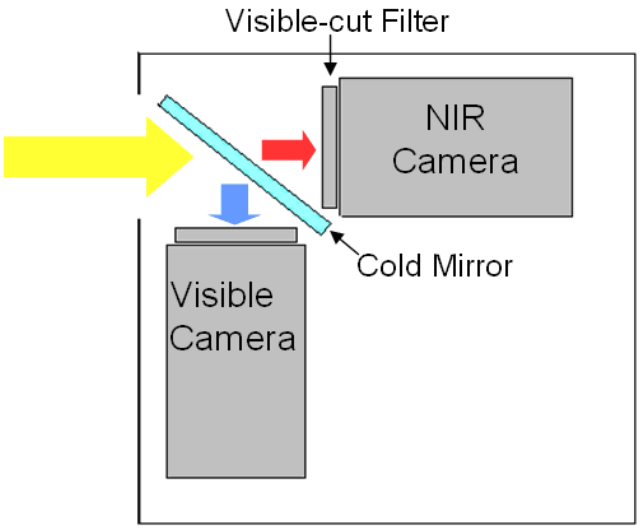

Figure 7: Diagram of multispectral camera setup

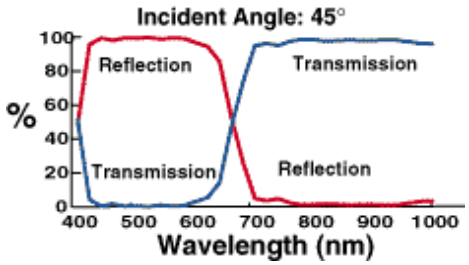

Figure 8: Reflection and transmission properties of the cold mirror (figure courtesy of Edmund Optics).

figure 8). This prevents the signal power loss traditionally associated with using a beamsplitter. Because the NIR camera is simply a visible-light camera with the IR-cut filter removed, it is fitted with a Hoya RM-90 visible-cut filter to ensure that it only measures NIR radiation. To prevent saturation as the camera changes direction relative to the sun, the NIR camera is set to auto gain and the gain of the visible light camera is slaved to match the gain of the NIR camera plus a constant offset to account for the current sky conditions, as clouds change the ratio of NIR illumination to visible illumination.

The video feed from the cameras is digitized with two imagenation PXC200A framegrabbers. The PVI is then calculated for each pixel and thresholded to produce real-time vegetation detection results. A simple threshold proved adequate because the output of the vegetation indices can be reasonably approximated with three distinct

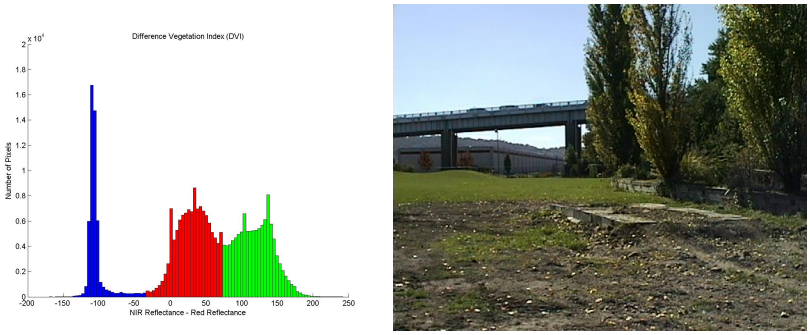

Figure 9: Distribution of DVI values (left) of the pixels in the typical image shown (right).

Gaussian distributions, one each for vegetation pixels, sky pixels, and everything else (see figure 9).

3.2 Experimental Results

With our camera system we collected tens of thousands of images from three different sites. A qualitative analysis has shown the system to be surprisingly robust in a variety of natural environments and weather conditions. Here we have included sample images from video sequences we captured. Video clips of the sequences are available at [1]. Figure 10 shows a sample result from a sunny day with grass, deciduous trees and a variety of man-made objects. The same scene on a rainy day (figure 11) shows even better performance because of reduced shadows and specular reflections as well as less frequent changes in the gain of the camera as its orientation varies relative to the sun. As chlorophyll breaks down in dead vegetation it becomes harder to detect, but the system was still able to easily discriminate between tree trunks and dead grass in mid-November (figure 12) and to detect pine needles and dead grass in early December (figure 13, 14, and 15).

The system would occasionally fail on some man made objects such as red vehicles (figure 16) and clothing (figure 17) that had extremely high NIR reflectance, but this problem was not observed with any natural objects. The video clips show some misclassification when the camera is turning sharply, but this is due to the fact that the cameras had not been synchronized and were capturing frames at different times. Earlier videos also show problems caused by a delay between when the NIR camera changed its gain settings and when the visible camera's settings were updated. Both problems require only minor technical fixes and will be fixed in future versions.

3.3 Results With Alternate Sensor Configurations

For our implementation we chose to use two cameras, however there are some single camera solutions available (primarily for airborne remote sensing). Unfortunately these cameras start around $10,000, roughly five times the price of our system. We have also

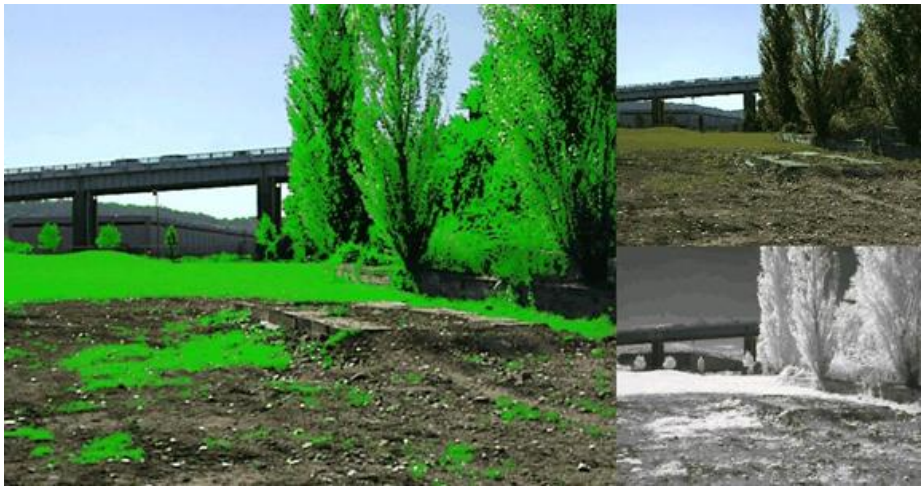

Figure 10: Sunny day with healthy vegetation and man-made objects. On the left, pixels detected as vegetation are marked in pure green with the level of confidence indicated by the pixels brightness. On the right are the raw visible and NIR images.

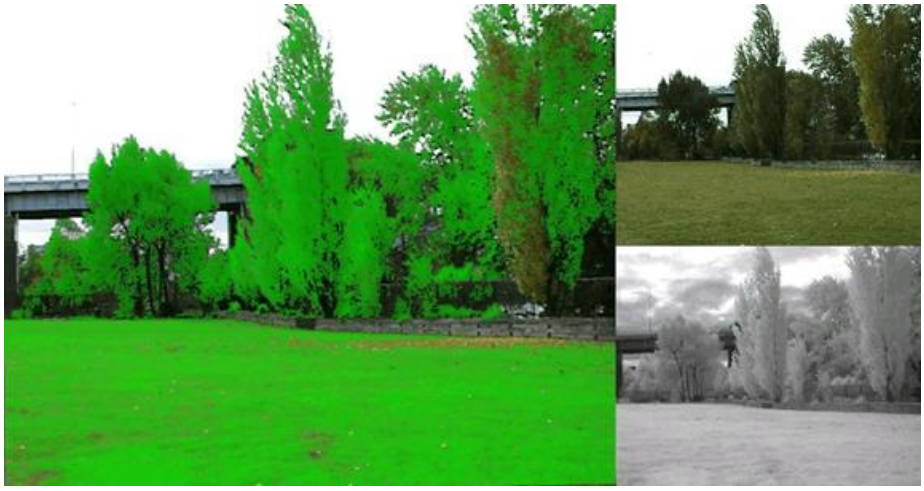

Figure 11: Similar view on a cloudy day with some dead vegetation. On the left, pixels detected as vegetation are marked in pure green with the level of confidence indicated by the pixel's brightness. On the right are the raw visible and NIR images.

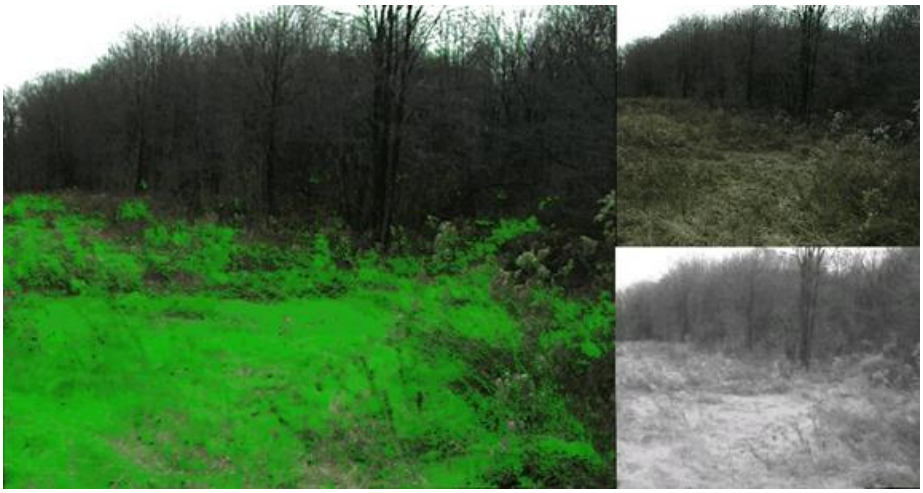

Figure 12: Field of tall dead grass in mid-November. On the left, pixels detected as vegetation are marked in pure green with the level of confidence indicated by the pixels brightness. On the right are the raw visible and NIR images.

Figure 13: Pine Trees, road, and dead grass in early December. On the left, pixels detected as vegetation are marked in pure green with the level of confidence indicated by the pixels brightness. On the right are the raw visible and NIR images.

Figure 14: Pine forest scene, early December. On the left, pixels detected as vegetation are marked in pure green with the level of confidence indicated by the pixel's brightness. On the right are the raw visible and NIR images.

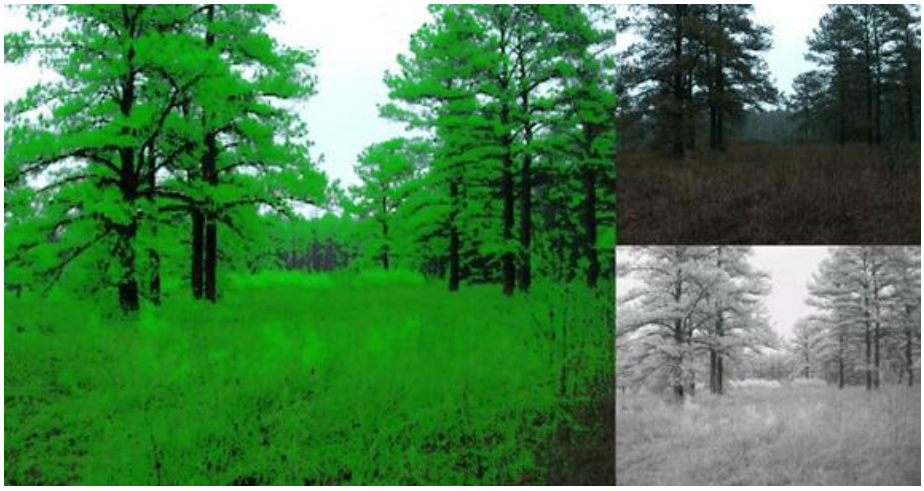

Figure 15: Tall dead grass and pine trees, early December. On the left, pixels detected as vegetation are marked in pure green with the level of confidence indicated by the pixel's brightness. On the right are the raw visible and NIR images.

Figure 16: Surprisingly, misdetection occurs much more often with red vehicle paint than green vehicle paint. Notice the strong NIR reflectance of the paint.

Figure 17: Synthetic clothing also causes misdetection. Notice how bright the jacket appears in NIR.

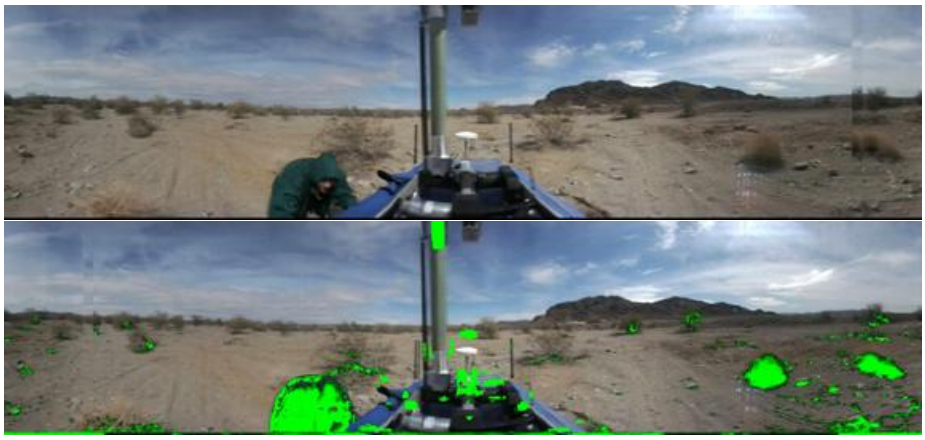

Figure 18: Panoramic image obtained from a camera mounted on top of a rotating SICK laser scanner (top). Computing the DVI using the reflectance channel of the laser and the red channel of the camera detects both chlorophyll and shadows (bottom).

experimented with combining a laser range finder and a camera for spectral vegetation detection. The Sick laser range finders mounted on the PerceptOR vehicles [14, 15] use a 900nm laser and can output a laser reflectance channel as well as range measurement. Combining the laser reflectance channel with the red channel of a color camera mounted on top of the laser allows us to use the same vegetation indices described previously. Figure 18 shows an image acquired using this setup in Yuma, Arizona. The strong response of the green plastic windbreaker shows that the vegetation index is working, since that particular green plastic is one of the known false positive materials. It is the detections on the bushes, however, that illustrate the drawbacks of this sensor arrangement. Because the reflectance image is lit from the laser whereas the color image is lit from above, shadows in the color image are classified as vegetation. It should be noted, though, that objects that self-shadowed were predominantly vegetation.

4 Analysis of Field Results

For natural environments the combination of low red reflectance and high near infrared reflectance seems to be remarkably specific to vegetation. We have observed consistently good results at three different sites (two around Pittsburgh and one in Virginia) under a variety of weather conditions and vegetation health. Additionally, the vegetation indices have been tested on satellite measurements taken all around the world and the USGS spectral database [2]. Man-made objects such as vehicles painted with pigments that are reflective in NIR can be misclassified as vegetation. In practice, though, vegetation detection will be used with other navigational sensors that can detect these failures. For instance, the truck that was misclassified in Figure 16 provides very consistent depth returns to a laser range scanner, indicating a solid object. A stand of tall grass, on the other hand, returns a range of depth readings to a laser scanner because

Vegetation Index	DVI	PVI	NDVI	MSAVI2
Time per image	37.5	44.7	46.7	198.2

Table 2: Execution time of various vegetation indices implemented in Matlab and run on 470x595 pixel images. Results were obtained by averaging 190 runs on a 2.6 GHz Pentium 4.

some pulses bounce of the first row of plants while others penetrate farther into the stand [11]. As described in section 2.3, when used for binary classification the different vegetation indices are merely linear classifiers operating on a 2-D feature space. With two free parameters, PVI is the most general of the indices and can be tuned to the best performance. The other indices provide suggestions for how to pick the parameters of PVI, but need not be implemented since PVI is relatively efficient (see Table 2). Sample processed results video clips are available online at [1].

Although the PVI-based system has proven to be very robust, the parameters still need to be adjusted between significantly different lighting situations (i.e. cloudy days vs sunny days). A human operator can quickly and easily tune the threshold, so for situations when a human operator is remotely supervising several vehicles (as in the PerceptOR program [14, 15] this presents no difficulty. Alternatively, automatic threshold-tuning could be accomplished by placing an test pattern of known red and NIR reflectance in the camera's field of view and adjusting the parameters of the vegetation index and the gain on both cameras to match a predetermined response.

5 Conclusions

A robust and computationally efficient vegetation detector is a useful tool for telling pliable vegetation apart from true mobility obstacles. We have demonstrated a simple camera system that captures red-channel and NIR images of a scene and applies the perpendicular vegetation index to reliably detect vegetation. Simulation results and field experiments have shown the system to be robust across a variety of weather and vegetation-health conditions.

References

[1] D. Bradley. Example vegetation detection results videos, available at **http://www.davidbradley.info/vegetationDetection/index.html**, 2004.

[2] R. N. Clark, G. A. Swayze, K. E. Livo, R. F. Kokaly, S. J. Sutley, J. B. Dalton, R. R. McDougal, and C. Gent. Imaging spectroscopy: Earth and planetary remote sensing with the usgs tetracorder and expert systems. *J. Geophys. Res.*, December 2003.

[3] R. N. Clark, G. A. Swayze, R. Wise, K. E. Livo, T. M. Hoefen, R. F. Kokaly, and S. J. Sutley. Usgs digital spectral library splib05a, usgs open file report 03-395, 2003.

[4] R. E. Crippen. Calculating the vegetation index faster. *Remote Sensing of Environment*, 34:71–73, 1990.

[5] M. Hebert and N. Vandapel. Terrain classification techniques from ladar data for autonomous navigation. In *Proc. Collaborative Technology Alliances Conference*, May 2003.

[6] A. R. Huete. A soil adjusted vegetation index (savi). *Remote Sensing of Environment*, 25:295–309, 1988.

[7] C. F. Jordan. Derivation of leaf area index from quality measurements of light on the forest floor. *Ecology*, 50:663–666, 1969.

[8] R. J. Kauth and G. S. Thomas. The tasseled cap - a graphic description of the spectral-temporal development of agricultural crops as seen by landsat. In *LARS: Proceedings of the Symposium on Machine Processing of Remotely Sensed Data*, pages 4B–14–4B–51, West Lafayette, IN, 1976.

[9] F. J. Kreigler, W. A. Malila, R. Nalepka, and W. Richardson. Preprocessing transformations and their effects on multispectral recognition. In *Proc. of the Sixth International Symposium on Remote Sensing of Environment*, pages 97–131, Ann Arbor, MI, 1969.

[10] T. M. Lillesand and R. W. Kiefer. *Remote sensing and image interpretation*, chapter 1. John Wiley, New York, 1987.

[11] J. Macedo, R. Manduchi, and L. Matthies. Ladar-based discrimination of grass from obstacles for autonomous navigation. In *Proc. of the International Symposium on Experimental Robotics*, Honolulu, HA, December 2000.

[12] J. Qi, A. Chehbouni, A. R. Huete, Y. H. Kerr, and S. Sorooshian. A modified soil adjusted vegetation index: Msavi. *Remote Sensing of Environment*, 48:119–126, 1994.

[13] A. J. Richardson and C. L. Wiegand. Distinguishing vegetation from soil background information. *Photogrammetric Engineering and Remote Sensing*, 43:1541–1552, 1977.

[14] A. Stentz, A. Kelly, H. Herman, P. Rander, and R. Mandelbaum. Integrated air/ground vehicle system for semi-autonomous off-road navigation. In *Proceedings of AUVSI-2002*, July 2002.

[15] A. Stentz, A. Kelly, P. Rander, H. Herman, O. Amidi, R. Mandelbaum, G. Salgian, and J. Pedersen. Real-time, multi-perspective perception for unmanned ground vehicles. In *Proceedings of AUVSI-2003*, July 2003.

[16] R. Willstatter and A. Stoll. *Utersuchungenuber Chlorophyll*. Springer, Berlin, 1913.