A Spectro-polarimetric Imager for Intelligent Transportation Systems

Daniel F. Huber[a], Louis Denes[b], Martial Hebert[a], Milton Gottlieb[b], Boris Kaminsky[b], Peter Metes[b]

[a]Robotics Institute, Carnegie Mellon University,
[b]Carnegie Mellon Research Institute, Carnegie Mellon University,
5000 Forbes Avenue, Pittsburgh, PA 15213

ABSTRACT

We have built a portable spectro-polarimetric machine vision system that operates at video frame rates. Our system contains only electronically controllable components, including an imaging acousto-optic tunable filter (AOTF), a phase retarder, acceptance and imaging optics, and a standard CCD-based camera. The device operates like an ordinary camera, except that a computer controls the spectral and polarization content of light to be viewed. For example, by sweeping the wavelength over the AOTF's range (visible and near infrared), one can obtain a spectral signature for each pixel in an image. Alternately, the camera can switch between two wavelengths, allowing for high-speed discrimination of closely matched colors in a scene. In addition to digitally controlling the wavelength, our imager uses a liquid crystal retarder to filter images based on polarization signatures of objects. We have implemented a number of algorithms to take advantage of the unique capabilities of our sensor, some of which can be applied to problems specific to transportation systems. We present two image processing applications that highlight the different methods we use to analyze scenes with our system. One application uses spectral processing to locate vegetation in a scene; the second uses polarization signatures to detect glare from hazardous road conditions such as water and ice.

Keywords: AOTF, multi-spectral, polarization

1. INTRODUCTION

Multi-spectral imaging techniques have been used extensively in remote sensing systems. Recent advances in imaging acousto-optical tunable filter (AOTF) technology have enabled us to build a practical, portable, land-based, multi-spectral imaging system[1] (figure 1). This system is entirely electronically controlled and works in real-time, which makes it a useful sensor for intelligent transportation systems as well as more for general computer vision applications.

In multi-spectral imaging, narrowband filters are used to sample a scene at each of several bands across the wavelength range of interest. The scene can then be analyzed in terms of its spectrum in addition to the normal spatial coordinates. A color camera gives some indication of the spectral content of a scene; but, because sensitivity curves for each color cover a wide range of wavelengths, the red, green, and blue (RGB) intensities offer only a summary of the original spectra in the scene. Due to this loss of information, two different spectra, called metamers, can produce identical RGB values. By using

Figure 1: The spectro-polarimetric camera

narrowband filters, multi-spectral imaging systems obtain enough detail about the spectra to discriminate between such objects which would be indistinguishable when viewed with a color camera.

Although multi-spectral processing is not new to vision research, traditional approaches are somewhat cumbersome and slow, typically limited to a small number of bands obtained by mechanically inserting filters into the optical path. The limitation in the number of bands means that the spectra will be necessarily coarse. In systems that support larger numbers of bands, the switching time between bands is usually too large for practical use in situations where the camera or the scene is in motion. Since these are precisely the situations in which sensors on vehicles must operate, it would seem that multi-spectral imaging is inappropriate for transportation systems. However, the system that we have developed is based on an AOTF, a device capable of switching between a large number of independent bands at high speeds (41 bands with a 30 microsecond switching time). Our system also includes an optional liquid crystal phase retarder which allows independent filtering based on the polarization signatures in a scene. Because our system is targeted for operation in demanding outdoor environments, it is necessarily rugged, compact, and modular.

In the next section, we discuss the details of the design of our system, focusing primarily on the hardware aspects of the device. First, we describe each of the three main modular components: the camera, control electronics, and host computer. Then we explain the different modes in which the system operates. In section three, we demonstrate two image-processing applications which capitalize on the multi-spectral or polarimetric capabilities of our imaging system. The first application, vegetation detection, uses spectral bands to discriminate two classes of terrain, while the second application, wet road detection, uses polarimetric signatures to identify hazardous road conditions.

2. SPECTRO-POLARIMETRIC IMAGING SYSTEM

The system described in this paper was targeted for operation in outdoor, real-world environments. As such, the primary design considerations were portability, ruggedness, and turnkey operation. The system is composed of three parts: the camera enclosure, the control electronics, and the host computer (figure 2). The system operates much like an ordinary CCD camera, except that the spectral and polarimetric content of light to be viewed is electronically controlled via computer. During operation, the host computer sends commands to the controller to select the desired spectral and polarization parameters. The controller interprets the commands, generates the required analog signals, and passes them to the camera enclosure. Finally, the incoming image is optically filtered and sampled by the camera, and the resulting video signal feeds back to the host computer where it is digitized by a frame grabber. Subsequent software processing of the image can then be used to determine the next set of control parameters.

2.1. AOTF camera

The multi-spectral capabilities of our imaging system are due to the AOTF located in the camera enclosure. The AOTF itself is a tellurium dioxide (TeO_2) crystal with acoustic transducers bonded to the surface on one side. An RF signal applied to the

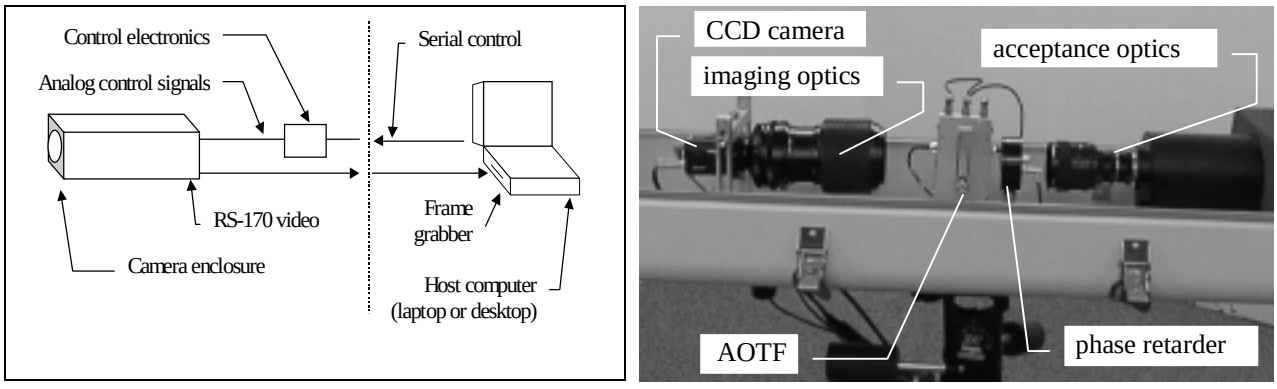

Figure 2: Left: Block diagram of the imaging system hardware. Right: Components of the camera enclosure. Light enters from the right, passing through the acceptance optics, retarder, AOTF, and imaging optics before being recorded by the CCD.

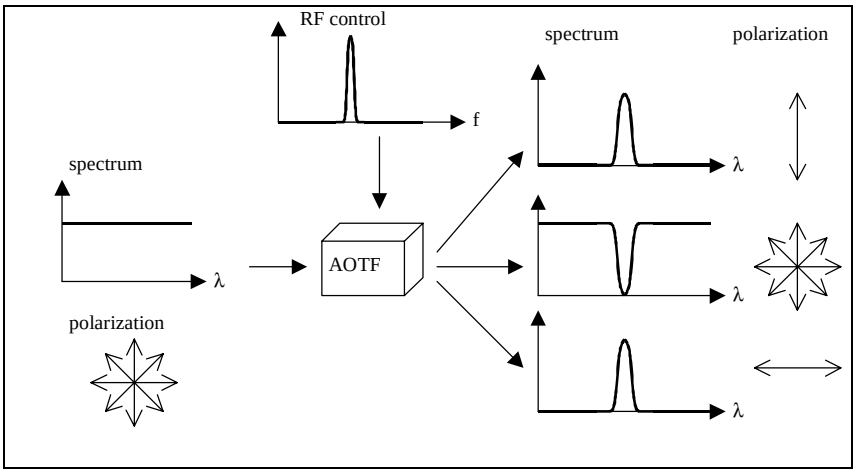

Figure 3: An example of the filtering action of the AOTF. Unpolarized white light enters the AOTF, which is controlled with a narrow band RF signal. The light is decomposed into three components: vertically polarized bandpass filtered, unpolarized notch filtered, and horizontally polarized bandpass filtered.

transducers creates a traveling acoustic wave inside the crystal. As light passes through the crystal, it is optically filtered and spatially separated into three components: two are bandpass filtered according to the frequency spectrum of the applied RF signal, and one is notch filtered using the complement of the RF spectrum. The bandpass filtered components are plane polarized orthogonally to one another. As an example, if an RF signal with the frequency-domain spectrum shown at the top of figure 3 is applied to the AOTF, unpolarized white light will be optically separated into the components shown on the right. More complex RF spectra can be applied in order to create different optical filtering functions.

Following the optical path through the camera enclosure, incoming light first passes through the acceptance optics, which provide active zoom control and collimate the light before it passes through the AOTF. After passing through the phase retarder (discussed in the next section), light enters the AOTF, where it is decomposed into the three components described previously. We align the imaging optics to focus the vertically polarized bandpass filtered image onto the CCD camera image plane while excluding the other two images.

Apart from the AOTF, all components in the camera enclosure are readily commercially available. The enclosure itself is weatherproof and ruggedized, suitable for mounting on the exterior of a vehicle. Using the active zoom on the acceptance optics, the camera's field of view is adjustable from 1.6 to 16 degrees. The AOTF can switch between two bands in 30 microseconds and operates with a 10-nanometer spectral resolution over a range from 450 to 850 nanometers, corresponding to the visible spectrum and near infrared. The operating range of the AOTF is chiefly determined by the choice of crystal material. We are currently investigating other materials that will extend the AOTF's range further into the infrared.

2.2. Phase retarder

The liquid crystal variable phase retarder provides optional polarimetric control for the imaging system. Conceptually, a retarder operates by resolving incident light into two orthogonal components with electric vectors oriented along what are called the fast and slow axes of the device. By varying the voltage applied to the retarder, we control the amount of phase delay between the components, achieving the overall effect of changing the polarization state of the incoming light. The AOTF acts as an analyzer by transmitting only vertically polarized light, and by varying the retardation, we change the intensity of light reaching the CCD camera.

For natural scenes in the visible spectrum, polarized light is predominantly linearly polarized, and under direct illumination from the Sun, Moon, or under overcast skies, the polarization of light is primarily horizontal. To observe the amount of horizontally linearly polarized light in a scene, we mount the retarder with the fast axis oriented 45 degrees from horizontal. If the input voltage to the retarder is set to achieve $\lambda/2$ phase delay, horizontally polarized light is rotated 90 degrees and becomes vertically polarized. In this case, the light would be transmitted by the AOTF. With no phase delay, the light is blocked by the AOTF. By switching between these two extremes, objects can be discriminated based on their polarization. By sweeping the retardation from 0 to $\lambda/2$, we can obtain polarization signatures of points in the scene.

With our retarder, we can obtain general characteristics of an object's polarization signature but not the exact polarization angle and magnitude obtained by polarization cameras such as that designed by Wolff.[2,3] However, we are currently designing variable phase retarders that would operate over a wider range than Wolff's system. Due to the limitations of the liquid crystal technology on which the retarder is based, the device's switching time is somewhat larger than that of the AOTF. In the worst case, the latency is 21 ms, which is slightly longer than the 17 ms required to obtain one field with the CCD camera. Consequently, the system must operate at only 15 frames per second when the retarder is used.

2.3. Controller and host computer

The controller acts as an intermediary between the host computer and the camera and provides a uniform interface across multiple platforms. Any computer with a serial port and a frame grabber can control the camera with no hardware redesign. Inside the controller, a Motorola 6816 microcontroller interprets the serial commands from the host computer and generates the necessary control signals for the retarder and the AOTF. The retarder voltage is controlled digitally using the manufacturer-supplied controller. However, generation of the RF signal for the AOTF is more complicated. The microcontroller programs up to four voltage-controlled oscillators (VCO's), the outputs of which are combined and amplified to produce the RF signal. We control the frequency of each VCO independently, and amplitude control is achieved with a digital step attenuator which allows up to 19 dB of attenuation.

Our simple camera interface gives us great flexibility in choosing the host platform. Our goal is to use a laptop PC, making the system truly portable, but current bus bandwidth limitations on PCMCIA frame grabbers prevent video frame rate digitization. At this time, we are using a 200 MHz Pentium-Pro PC with an ImageNation PX500 frame grabber capable of capturing 30 frames per second.

2.4. Operational modes

From an operational standpoint, determining the control strategy for the system is an important yet challenging issue. The problem is difficult because a finite frame rate prevents acquisition of all data simultaneously, and even if instant acquisition were possible, additional time is needed to process the massive amount of data produced by the camera. When choosing a control strategy, one must decide where to draw the line in the tradeoff of time and information. Two control strategies lie at the extreme ends of this range of choices: sweep mode and switching mode. In sweep mode, we use a single VCO to perform a full spectral sweep over the functional range of the camera (450 to 850 nm), collecting 41 narrowband images spaced at 10 nm intervals. This method gathers the maximum amount of information in a scene, but at the expense of a significant time delay since more than one second elapses between the capture of the first and last images in the sequence. Obviously, this method presents problems when the scene or the camera are in motion.

In switching mode, the camera alternates between a small number of parameter settings. The parameters for each image can use one VCO to acquire a single band in the scene, or they can be chosen to create more complex filters by combining multiple VCO's and the retarder. In the extreme, an arbitrary waveform generator could be used to generate arbitrarily complex filters. We are currently working to add this capability to our system. The filters used in switching mode can be viewed as basis functions for the spectra in the scene. For example, three filters with the shape of the photoreceptor response of the eye could be used to simulate human visual perception. Maloney argued that many spectra can be accurately approximated by a linear combinations of a small number (5-7) of basis functions.[4] Using such basis functions in our system is one way to balance the tradeoff of time and information gain. For a given application, optimal basis functions could be obtained by analyzing sweep mode spectra using Maloney's methods.

3. IMAGE PROCESSING

Multispectral analysis can be applied to a variety of computer vision problems, including target discrimination, image segmentation, and object recognition. We have implemented a number of algorithms to take advantage of the unique capabilities of our sensor, some of which can be applied to problems specific to transportation systems. For example, by choosing the appropriate models and training data, our image segmentation algorithm can be used to detect road signs. The same methods could be used to locate road boundaries or other distinguishing environmental features.

In this section, we present two image processing applications that highlight the different methods we use to analyze scenes with our system. The first application is a reliable vegetation detection algorithm that is useful for off-road navigation systems. The second application uses polarization signatures to highlight wet road surfaces and is more applicable to highway-bound vehicles.

3.1. Terrain classification: vegetation detection

We have implemented a real-time vegetation detection algorithm as part of a larger effort to develop a terrain classification module for the Unmanned Ground Vehicle (UGV) project. We have already demonstrated autonomous planning, mapping, and off-road navigation skills using the NavLab II, a modified Army HMMWV, but despite its impressive capabilities, NavLab II is unable to distinguish between rocks and tall grass, trees and hillsides, or even mud and hard ground.[5] As a consequence, the vehicle plans and navigates conservatively, avoiding all objects that may be potential hazards. By classifying the terrain in a scene, we can reduce the number of false positive obstacles, such as tall grass, as well as false negatives, such as water and mud.

Intensity-based terrain classification with a monochrome camera is difficult because different terrain types may produce the same image intensity and relative intensities vary considerably with lighting conditions. Color cameras alleviate this problem somewhat, but the terrain in which we are interested frequently contains only muted shades of greens and browns, which may still be impossible to distinguish.[6] However, with the ability to analyze the complete spectrum of a point in the scene, the terrain classification problem becomes easier. For example, it is known that the chlorophyll in vegetation leads to a unique spectral signature exhibiting a sharp rise in reflectance in the 690 to 730 range (figure 4). This effect can be seen even in winter, when brown grass looks much like dirt when viewed with a color camera. Other researchers have shown that the spectra of different types of trees are visually distinct.[7]

Our algorithm takes advantage of the characteristic shape of vegetation spectra to discriminate between vegetation and non-vegetation. For this application we operate in switching mode, using band ratios – ratios of intensities between pairs of bands. Band ratios have been shown to be generally useful in multi-spectral imaging, as have spectral derivatives.[8] Unfortunately, spectral derivatives require sampling of the entire spectrum, which is incompatible with real-time performance. Another advantage of the band ratio method is its robustness under varying illuminants – it performs well under typical outdoor lighting conditions, including full sun, partly cloudy, and completely overcast. For vegetation detection, we use the ratio of the band centered at 730 nm, corresponding to the peak in the infrared, and the band at 650 nm, corresponding to a point in the red region where vegetation's spectrum tends to reach a local minimum. One can apply

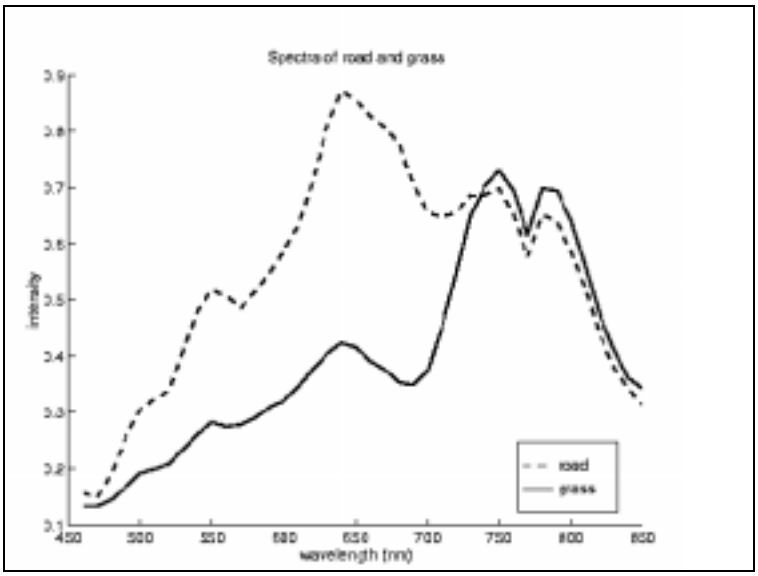

Figure 4: Example of a typical spectrum of grass (solid line) with a spectrum of a grey road surface (dashed line) for comparison. The grass spectrum shows a marked peak between 700 and 850 nm even though the image was acquired during winter when the grass was brown.

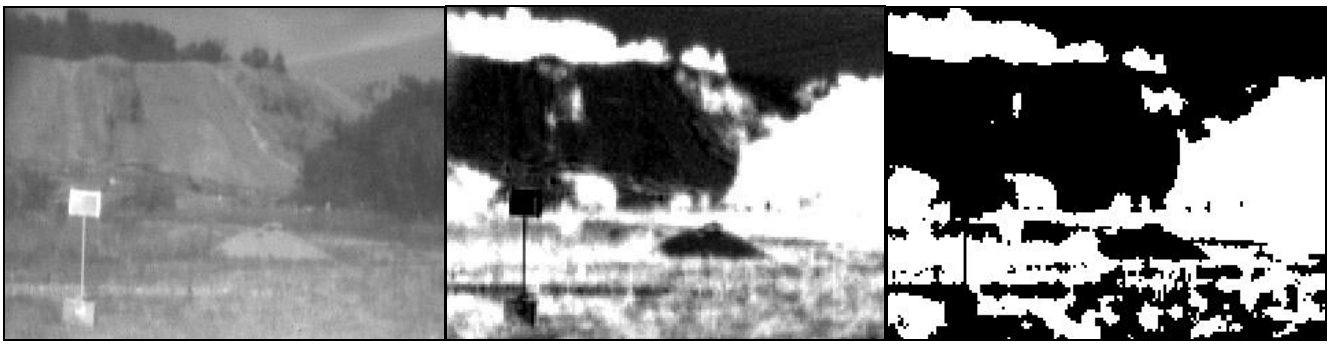

Figure 5: Left: A scene with a large dirt mound in the center, some tall grass to the left, and two smaller dirt mounds hidden behind the tall grass. Center: The scene has been processed to highlight areas of vegetation in brighter shades and non-vegetation in darker shades. Right: After thresholding, the three dirt mounds are clearly differentiated from the tall grass.

thresholding to the resulting image to obtain a boolean classification result, or the ratio itself can be used as a confidence measure. Matthies reported plans to develop a similar vegetation detector using a filter-wheel camera to sample scenes at a small number of wavelengths.[9]

The scene shown in figure 5 was taken at the slag heap, an area of rough terrain near Carnegie Mellon University (CMU) where we test our autonomous off-road vehicles. The scene shows a large hill of dirt in the distance, a flat area of mixed grass and dirt in the foreground, and a reference card, which was not used in this experiment, on the left. In the center of the image lies an area with several mounds of dirt interspersed with foot-high grass. Our current obstacle detection on NavLab would classify the entire region as an obstacle even though a human driver would be able to drive through the tall grass and avoid the dirt piles. The detector's normalized band ratio and thresholded output are shown in the center and right panels of the figure. Notice that the dirt mound is clearly distinguishable from the tall grass and that two partially hidden mounds are now visible behind the tall grass. By integrating the vegetation detector's output with NavLab's current obstacle detection sensors, the vehicle could choose to drive through the relatively safe tall grass rather than avoiding the area completely.

3.2. Polarimetry: road hazard detection

Wet or icy roads present hazardous situations for human and autonomous drivers alike. Aside from the increased danger of losing control of the vehicle, glare from wet pavement makes it difficult to see lane markings and other road features. The problem is more pronounced in autonomous vehicles since they typically don't have the ability to fall back on other visual cues and background knowledge. In high-glare situations, a polarizing filter can significantly reduce reflections due to the fact that water, and therefore wet roads, tends to horizontally polarize reflected light. We foresee two ways in which our imaging system could be applied to wet or icy road conditions: glare reduction and glare detection.

Glare reduction works similarly to a polarizing filter inserted in the optical path. However, the phase retarder accomplishes this electronically and can be programmatically enabled and disabled, which is advantageous since an additional filter reduces the overall intensity of the image. For glare detection, we identify horizontally polarized reflections, thereby detecting possible hazardous road conditions. In a computer-assisted driving scenario, this additional information could be used to alert the driver of the danger, or in an autonomous vehicle, the controller could modify its driving habits accordingly – perhaps by slowing down and limiting steering adjustments to reduce the chance of skidding. Initial experiments indicate that wet road detection is possible with our system. For this application, we use switching mode to tune the retarder to the two values that maximize and minimize the intensity of reflected light in the two separate frames. We then subtract the two images, scale them, and optionally threshold the result to highlight the polarized regions (figure 6). This simple algorithm can differentiate a high-glare region from an area with no glare but equivalent brightness in the image.

Figure 6: Results of the wet road detection experiment. Top row: The two original images of a partially wet road with the retarder tuned to achieve minimum and maximum reflection. Portions of the scene – the top center and left and bottom right – remain dry for comparison. Lower left: The scene has been processed to emphasize the contrast between dry and wet pavement. Bottom right: The enhanced image has been thresholded to identify high-glare regions. Notice that areas with high surface reflectance and no glare, such as the stem of the letter P, are correctly identified as having no glare.

4. FUTURE WORK

We are currently working to advance the imaging system hardware, to develop, improve, and extend our classification applications, and to begin on-vehicle testing. Our primary hardware focus is the development of a new AOTF that will extend the upper range of the system further into the infrared region. Secondly, we are working to integrate an arbitrary waveform generator into the imaging system, allowing us to generate complex filtering functions. We are investigating methods for choosing the number and shapes of these filters.

On the software side, we are continuing to develop our terrain classification system and extend it to the general problem of classification and object recognition. Multi-spectral classification can be applied in a variety of applications. For example, we have implemented a minimum Mahalanobis distance classifier that could be used to detect and recognize road signs.[10] Many signs (e.g. stop signs) are brightly colored and therefore exhibit distinctive spectral signatures which can be discriminated using multi-spectral imagery. Of course, color cameras could also detect stop signs but would have trouble differentiating a green interstate highway sign from a background of trees, whereas our system would not. Geometric cues, such as shape, could be used in conjunction with the road sign detector to improve robustness, or the detector's output could act as a filter for input to a sign-reading module. Preliminary results of our classifier are encouraging. Figure 7 shows a typical scene containing a stop sign. We applied our classification algorithm using a database of sign and vegetation models. The result, shown in the right image, shows stop sign's octagonal shape in stark contrast to the background of vegetation and unknown surfaces. Additional work is required to reduce the system's sensitivity to changing illumination conditions.

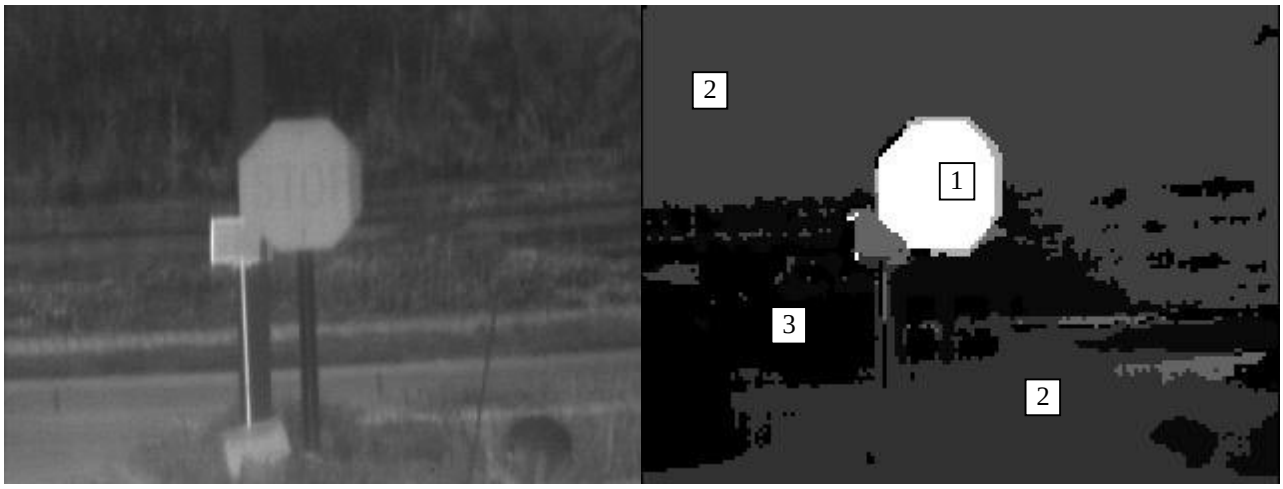

Figure 7: An example of classification using a database of traffic signs and vegetation. Left: The original scene with stop sign and reference card. Right: The classified scene showing three main classes 1) the sign 2) vegetation and 3) unknown.

Variation in illumination is a major problem in outdoor computer vision applications. Depending on the time of day, weather conditions (sunny versus cloudy), or even scene geometry, the measured spectral signature for a given object can change significantly. The ability to compensate for varying illuminants is known as color constancy, and in order to perform classification accurately in an outdoor environment, one must either solve the color constancy problem or use methods that do not rely on strict color constancy. No color constancy algorithm can achieve constancy in all situations, and indeed most existing algorithms make substantial assumptions about scene geometry and the form of the illuminant.[11] However, by assuming that the illumination source is sunlight, reasonable constancy may be obtainable.

In our current experiments, we compensate for lighting changes using a reference card of known spectral signature placed in the scene. Using this card, we can normalize a target scene based on the relative intensity of the reference card spectra in the target and model scenes. Since it is not always practical to place a reference card in the scene in advance, we will look for alternate methods to achieve color constancy in the future. The state-of-the-art color constancy algorithms are generally tested in laboratory conditions using planar surfaces of color patches, so it will be interesting to see how these methods operate in outdoor environments.

5. ACKNOWLEDGEMENTS

This research has been sponsored in part by the Office of Naval research (ONR) under Contract N00014-95-1-059 and by DARPA under contract DAAE07-96-C-X075, entitled "Technology Enhancements for Unmanned Ground Vehicles," monitored by U.S. TACOM.

6. REFERENCES

1. L.J. Denes, et. al., "Multi-Spectral Imaging Filters," *Proceedings of the 1997 Image Understanding Workshop (IUW)*, New Orleans, LA., May, 1997, pp. 1483-1489.
2. L. B. Wolff, T. A. Mancini, P. Pouliquen, and A. G. Andreou, "Liquid Crystal Polarization Camera," *IEEE Transactions on Robotics and Automation*, vol. 13, no. 2, pp. 195-203, 1997.
3. L. B. Wolff, "Applications of Polarization Camera Technology," *IEEE Expert*, vol. 10, no. 5, pp. 30-8, Oct. 1995.
4. L. T. Maloney, "Evaluation of linear models of surface spectral reflectance with small numbers of parameters", J. Opt. Soc. Am. A, vol. 3, no. 10, pp. 1673-83, 1986.
5. M. H. Hebert, C. Thorpe, and A. Stentz, eds., *Intelligent Unmanned Ground Vehicles – Autonomous Navigation Research at Carnegie Mellon*, Kluwer, Boston, 1997.

6. I. L. Davis, A. Kelly, A. Stentz, and L. Matthies, "Terrain Typing for Real Robots," *Proceedings of the Intelligent Vehicles '95 Symposium*, Detroit, MI, Sept. 1995, pp. 400-5.
7. L. Cheng, M. Hamilton, C. Mahoney, and G. Reyes, "Analysis of AOTF hyperspectral imagery," *SPIE Conference on Algorithms for Multispectral and Hyperspectral Imagery*, Orlando, FL, April 1994, SPIE vol. 2231, pp. 158-66.
8. M. A. Sturgeon, et. al., "Spectral and polarimetric analysis of hyperspectral data collected by an acousto-optic tunable filter system," *SPIE Conference on Algorithms for Multispectral and Hyperspectral Imagery*, Orlando, FL, April 1994, SPIE vol. 2231, pp. 167-76.
9. L. Matthies, A. Kelly, T. Litwin, G. Tharp, "Obstacle Detection for Unmanned Ground Vehicles: A Progress Report," *Proceedings of the Intelligent Vehicles '95 Symposium*, Detroit, MI, Sept., 1995, pp. 66-71.
10. R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, ch. 2, Wiley-Interscience, New York, 1973.
11. D. A. Forsyth, "A Novel Algorithm for Color Constancy," in *Physics-Based Vision Principles and Practice – Color*, G. E. Healey, S. A. Shafer, and L. B. Wolff, eds., pp. 241-71, Jones and Bartlett, London, 1992.