HYPERSPECTRAL VIDEO FOR ILLUMINATION-INVARIANT TRACKING

Amit Banerjee, Philippe Burlina, and Joshua Broadwater

Johns Hopkins University Applied Physics Laboratory, Laurel, MD 20723, USA

ABSTRACT
Recent advances in electronics and sensor design have enabled the development of a hyperspectral video camera, which can capture hyperspectral datacubes at near video rates. In this work, we show how high-speed hyperspectral imaging can be used to address several challenging problems in video surveillance. In particular, we combine traditional methods for hyperspectral image analysis and computer vision to achieve illumination-invariant motion detection and object tracking. Experiments using real hyperspectral video images are provided.

Index Terms— Hyperspectral video, illumination invariant analysis, deshadowing, tracking, motion detection

1. INTRODUCTION
Most modern video cameras provide imagery with high spatial and temporal resolution that is ideal for detecting and tracking moving objects. However, their low spectral resolution limits their ability to classify or identify objects based on color alone [1],[2]. Conversely, traditional hyperspectral sensors offer high-resolution spectral and spatial imagery at low temporal resolutions with modest frame rates (up to 0.5 Hz). Hence, they have been utilized extensively for object detection and classification, but usually in static, non-dynamic environments.

Recent groundbreaking work has led to the development of novel hyperspectral video (HSV) cameras that are able to capture hyperspectral datacubes at near video rates. This technological breakthrough was made possible by recent innovations in fast electronics and sensor design. While standard video cameras capture only 3 wide-bandwidth color images, the HSV camera collects many narrow-bandwidth images of the scene. Since the hyperspectral video camera is able to simultaneously capture images with high temporal, spatial, and spectral resolution, it combines the advantages of both video and hyperspectral imagery in a single sensor.

In this work, we demonstrate the utility of the HSV camera for motion detection and object tracking. These are critical parts of many video surveillance systems that have been the focus of extensive research in the computer vision community. By capturing hyperspectral imagery at near video rates, we can now incorporate radiative transfer theory methods into standard vision algorithms and achieve illumination-invariant video surveillance. Specifically, the proposed approaches for motion detection and object tracking rely on estimating the reflectance spectra of the objects in the scene. By comparing the reflectance spectra, it is possible to mitigate the effects of slow and fast-changing illumination conditions on motion detection and object tracking. Furthermore, processing images in the reflectance domain allows for the use of simple algorithms for motion detection and object tracking.

This paper is organized as follows. Section 2 is an overview of the HSV camera to describe how it captures hyperspectral images at high speeds. Section 3 details the approach to estimate the illumination-invariant reflectance from the observed radiance measurements, and Section 4 explains how the reflectance spectra are used to detect motion and track moving objects. Section 5 presents experiments on real HSV data to demonstrate spectral-based change detection and tracking of objects through occlusions and illumination changes. Section 6 concludes the paper.

2. THE HYPERSPECTRAL VIDEO CAMERA
The HSV camera is a passive sensor that measures the optical spectra of every pixel from 400 - 1000nm. It acquires datacubes using a line scanning technique. An oscillating mirror scans the scene up to 10 times a second. For each mirror position one horizontal scan line is acquired and its pixels are decomposed by a spectrometer into a full spectral plane. The spectral plane is captured by a CCD, and is built into a datacube as the mirror completes a vertical scan of the scene. Some of the important specifications of the camera are included below. The Surface Optics Corporation developed the HSV camera used in this study. Some of the important specifications of the camera are included in Table 1.

<table>
<thead>
<tr>
<th>Spectral Bandwidth</th>
<th>400-1000nm (visible &amp; NIR)</th>
</tr>
</thead>
<tbody>
<tr>
<td># of Bands</td>
<td>22, 44, 88 (evenly spaced)</td>
</tr>
<tr>
<td>Maximum Speed</td>
<td>10, 5, 2.5 cubes/second</td>
</tr>
<tr>
<td>CCD size</td>
<td>512x512, 500fps</td>
</tr>
<tr>
<td>CCD pixel size</td>
<td>18µm x 18µm</td>
</tr>
<tr>
<td>Dynamic range</td>
<td>16 bits/pixel</td>
</tr>
<tr>
<td>Dimensions</td>
<td>7&quot; x 9.5&quot; x 26&quot;, 27 lbs.</td>
</tr>
</tbody>
</table>

Table 1. Specs for the SOC700-10Hz HSV camera.
3. ESTIMATING OBJECT REFLECTANCE

This section presents our approach to estimate the reflectance spectra of an object from its radiance spectra. Based on a simple model of radiative transfer theory [3], it automates the work of Piech and Walker [4] to estimate the functional mapping from radiance to reflectance spectra. The algorithm described below applies principally for objects with Lambertian or diffuse surfaces. Highly specular objects will require additional processing and are not considered here. While other algorithms may provide better reflectance estimates, the approach developed here is fast and allows for simple tracking of objects in variable illumination conditions.

First, the radiometric transfer function can be simplified to

\[ L(x, y, \lambda) = R(x, y, \lambda)[A(\lambda) + F(x, y)B(\lambda)] \]

where \( A(\lambda) \) represents the radiance due to sunlight, \( F(x, y) \) represents the amount of sky light at pixel \((x, y)\) (i.e., in shadow zones the amount of sky not blocked by the object creating the shadow), and \( B(\lambda) \) represents the radiance due to sky light. The terms \( A(\lambda) \) and \( B(\lambda) \) are considered to be independent of pixel location when small areas are imaged (i.e., the sunlight and skylight terms do not vary over the small area being imaged). Also note that unlike the empirical line method (ELM) [5], we must treat the skylight and sunlight terms separately due to the ground-based orientation of the camera.

For in-scene estimation of the parameters in Eq. (3), typically the reflectance of one of the objects in the scene is known. One approach is to identify objects in the scene that have nearly flat reflectance signatures (i.e. constant and independent of wavelength) in full sunlight conditions. For example, asphalt roofing tar has a relatively flat reflectance [6]. Eq. (3) becomes

\[ L_{flat}(\lambda) = k[A(\lambda) + FB(\lambda)] \]

where \( k \) represents an unknown flat reflectance value independent of wavelength. If the entire image is in sunlight, then the reflectance can be simply calculated as

\[ R(x, y, \lambda) = k \frac{L(x, y, \lambda)}{L_{flat}(\lambda)} \]

(3)

to within some unknown offset \( k \). To remove the effects of \( k \), each pixel is normalized to have the same energy. The result is an image with minimal illumination differences making tracking and identification of objects much simpler.

For images with shadow zones, the process is slightly more complicated. First, a shadow mask must be estimated. The energy of each pixel, computed using either the \( L_1 \) or \( L_2 \) norm of its spectra, is thresholded to produce the shadow mask. Given the shadow mask, a shadow line must be found that crosses across the same material. For the pixels in the full sunlight condition, Eq. (5) is applied to estimate the reflectance. Using the estimated reflectance, the skylight effects can be estimated such that

\[ kF(x, y)B(\lambda) = \frac{L(x, y, \lambda)}{R(\lambda)} \]

(4)

for pixels of the same material just inside the shadow zone. Thus estimates for both full sun and full shade conditions are available.

Using these estimates and the shadow mask, pixels in full sun can be converted to reflectance using (5). For pixels in shade, their reflectance can be calculated using

\[ R(x, y, \lambda) = \frac{L(x, y, \lambda)}{kF(x, y)B(\lambda)} \]

(5)

Again, we do not know the offsets due to \( k \) or \( F(x, y) \), but this can be handled by normalizing the resulting reflectance as was done with (5).

4. ILLUMINATION-INVARİANT ALGORITHMS

Given the ability to estimate the reflectance spectra of an object observed under varying illumination conditions, we are able to develop simple algorithms for illumination-invariant motion detection and object tracking. The algorithms outlined below assume a stationary HSV camera. However they can be generalized for moving cameras given accurate motion compensation. For illumination-invariant motion detection, a reference hyperspectral image frame is required that has been converted to reflectance. For the subsequent images in the HSV sequence, the processing steps are:

- Estimate the reflectance spectra for every pixel.
- Use the spectral angle mapper (SAM) to compare the reflectance spectra of the current (C) and reference (R) hyperspectral images.
  - If \( \text{SAM}(C(x, y), R(x, y)) < \tau \), then pixel \((x, y)\) is part of a moving object.
  - Else, pixel \((x, y)\) is a background pixel.

Since reflectance spectra are independent of the illumination conditions, this algorithm will mitigate false alarms due to sudden or slowly-changing illumination effects.

For illumination-invariant tracking, we exploit frame-to-frame matching of the reflectance spectra of the tracked object. Since the reflectance spectra are invariant to the imaging conditions, the objects can be tracked by simply comparing the spectra of every pixel with the spectra of the tracked object using the Mahalanobis detector:

\[ m_M(x) = (x - m)^T \Sigma^{-1} (x - m) \]

(6)

where \( x \) is the pixel under test and \( m \) and \( \Sigma \) are the mean and covariance of the object’s spectra. Pixels whose Mahalanobis distance is below \( T \) are considered to be the tracked object.
5. EXPERIMENTAL RESULTS

Figure 1 displays the two color images from an 88-band hyperspectral video sequence used for the change or motion detection experiment. The top image is the reference image taken in the morning, and the second image is the test image taken in the late afternoon. Note the differences in illumination and the varying shadow regions. There images are well-registered since there is negligible camera motion. The differences in the images are the three cars driving along the road in the background distance (highlighted by the red boxes) and the car parked in the foreground.

Using the approach outlined in Section 3, the reflectance spectra for every pixel in both images are estimated. The corresponding pixels in both images are compared using the spectral angle mapper (SAM) criterion. This approach was applied for the 88-band hyperspectral image and the 3-band color image derived from the hyperspectral datacube. The detection results are shown in Figure 2. Using color, the algorithm generates a high false alarm rate (Figure 2 top). Using spectra, the cars in the foreground and background are detected (green boxes) and some false alarms due to the blowing leaves are also detected (Figure 2 bottom).

The ability of HSV camera track objects is demonstrated using a real hyperspectral video sequence with 44 bands at 5 cubes/second. Two people wearing red shirts that are nearly indistinguishable for the human eye or an RGB camera walk in and out of shadows. The objective is to use the spectra of the person’s clothing to a) track him through different illumination conditions, b) distinguish him from the other person wearing a red shirt, and c) maintain track while he walks in and out of the camera’s field of view.

The plots of the observed radiance spectra of the red shirt are given in Figure 3(a). Note the spectral differences between the two shirts and how the spectra change when observed in sun and shadow. Using the algorithm described in Section 3, the reflectance spectra of the shirts are estimated from the observed radiance spectra. As can be seen in Figure 4(b), the estimated reflectance for each shirt is consistent, whether they are observed in sun or shadow. Further, the spectral differences in the two shirts are also maintained in the reflectance spectra. This indicates that the reflectance spectra can be used to track objects through varying illumination conditions and discriminate between the multiple objects with similar color.

The sequence of HSV frames in Figure 4 show the result of reflectance-based tracking. The pixels in green denote the tracked person’s red shirt. Note how the track is maintained as the person walks in and out of shadows, and in and out of the camera’s FOV. The spectral differences in the two shirts are also exploited to distinguish between the two people in red shirts.
Fig. 3. Plots of the radiance and reflectance spectra of the red shirts in sun and shadow.

6. CONCLUSIONS

We introduce a new framework for object tracking that utilizes a novel hyperspectral video camera. The HSV camera extends the capabilities of modern video cameras to provide imagery with high spatial, temporal, and spectral resolution. It has the potential to augment and in some cases revolutionize the methods applied for solving various computer vision tasks by exploiting tens of high-resolution spectral bands, instead of 3 low-resolution color bands.

7. REFERENCES


