

The following paper posted here is not the official IEEE published version. The final published version of this paper can be found in the Proceedings of the IEEE Conference on Intelligent Transportation Systems (5th : 2002 : Singapore):pp.100-105

Copyright © 2002 IEEE.

Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

A Comparative Study of Different Color Spaces for Foreground and Shadow Detection for Traffic Monitoring System

Pankaj Kumar
engp8919@nus.edu.sg

Kuntal Sengupta *
ksengupta@advancedinterfaces.com

Adrian Lee
eng91009@nus.edu.sg

Surendra Ranganath
elesr@nus.edu.sg

Department of Electrical and Computer Engineering
National University of Singapore
4 Engineering Drive 3
Singapore 117576

Abstract

Segmenting foreground objects of interest in real time is an important step in many applications of video surveillance, vehicle tracking and traffic monitoring. Background subtraction is a method very often used to segment moving objects in image sequences. In this paper we present a fundamental unbiased study of different color spaces for detecting foreground objects and their shadows in image sequences. Different color spaces show different efficiency in detection of foreground objects and their shadows. This empirical study was done with the motivation of determining, which color space is best for foreground segmentation and shadow detection. This study of the quality of foreground and shadow detection as a function of color space is unique kind and especially relevant to color image sequences. Our study includes five color spaces “RGB”, “XYZ”, “ $YCrCb$ ”, “HSV” and the normalized “rgb”. We use an empirically substantiated model of shadows formulate the detection scheme for each color space. We use a statistical technique to model the background pixels. The results are compared in terms of true detection, misses and false detection of pixels and also detection of the moving foreground objects as blobs. The results show that “ $YCrCb$ ” is the best color space for optimal foreground and shadow detection.

1 Introduction

Design of a computer vision based system for traffic monitoring and behavior analysis has gained high

importance for Intelligent Traffic and Vehicle System (ITVS) applications [5, 6, 12, 10]. Visual sensors are easy to install and maintain. They provide wide coverage and more information of the traffic compared to other sensors. In many vision systems, the capability of extracting moving objects from video is a fundamental and crucial problem. The drawback is, most systems are either too slow to be practical, or work only in very controlled situations. Recently, faster computers have allowed researchers to consider more complex and robust models for real-time streaming and analysis of data. These new methods have allowed researchers to begin modeling real-time processes under varying conditions [1, 4].

The foreground detection is usually achieved by motion detection and/or background suppression. However neither motion detection nor background suppression methods can distinguish between moving objects and moving shadows. Motion detection by frame differencing and foreground detection by background subtraction both detect shadows as foreground. Sometimes by using a higher threshold the false detection due to shadow can be suppressed but higher thresholds lead to other problems like misses and splits. Split is a case where single object gets divided into two or more objects. Good moving object segmentation [7, 2] and its shadow removal [11, 6] are important for good detection and tracking results. In this work we show that foreground object segmentation and shadow detection is a function of the color space we choose. So from the very start it is important that one chooses a proper color space to apply sophisticated background adaptation [14, 13] and maintenance [15] for better segmentation.

This paper presents an empirical and comparative study of the different color spaces for model based background removal and moving object and shadow detec-

* Advanced Interfaces Technologies
403 South Allen St.
State College, PA 16801

tion. In Section 2 we present a brief review of different color spaces and transformations amongst themselves. Section 3 describes the shadow model and how it matches experimental observation. In Section 4 we explain the background model and the segmentation criterion for detection of foreground and shadow. Three types of results are presented in Section 5; two are based on quantitative measures while third one is qualitative. Finally in Section 6 we discuss the conclusions and some directions of future work.

2 Different Color Spaces and Relationships Between Them

Color spaces are different bases for representing intensity and color information in color images. Usually color spaces have three components or channels for representing all possible color and intensity information. For example RGB is a very commonly used three dimensional color space with color components or channels Red, Green, and Blue. Different combination of these three channels are capable representing all possible shades. In our experiments we consider the following five color spaces “*RGB*”, “*XYZ*”, “*YCrCb*”, “*HSV*” and the normalized “*rgb*”. We give the mathematical relation of each of these color space to “*RGB*”. Using the relationship of different color spaces to “*RGB*” we can compute their relations amongst each other also.

- *RGB* - has primaries, Red, Green, Blue and it is the most commonly used color space.
- *HSV* - describes any color in terms of the three quantities - Hue, Saturation, and value (HSV). This is the color space used by [11] for shadow detection. The mathematical relation between HSV and RGB is given by

$$H = \begin{cases} \theta & \text{if } B \leq G \\ 360^\circ - \theta & \text{if } B > G \end{cases} \quad (1)$$

$$\text{where } \theta = \cos^{-1} \frac{\frac{1}{2}[(R-G)+(R-B)]}{[(R-G)^2+(R-B)(G-B)]^{\frac{1}{2}}}$$

$$S = 1 - \frac{3}{R+G+B}[\min(R, G, B)] \quad (2)$$

$$V = \frac{1}{3}(R+G+B) \quad (3)$$

This is a non-linear relationship taken from [3].

- *XYZ* - the fundamental color space of the Commission Internationale d'Eclairage (CIE) and is the standard based on color matching experiments on human observers. It has a linear conversion from the *RGB*

color space given as

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7151 & 0.0721 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (4)$$

- *YCrCb* - used by most image compression standards like JPEG, H.261, and MPEG has the following relation with the *RGB* space

$$\begin{bmatrix} Y \\ C_r \\ C_b \end{bmatrix} = \begin{bmatrix} 0.257 & 0.504 & 0.098 \\ 0.439 & -0.368 & -0.071 \\ -0.148 & -0.291 & 0.439 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} \quad (5)$$

- *rgb* - normalized RGB used by [9, 8] to detect foreground objects and remove shadows is obtained as follows.

$$\begin{aligned} r &= \frac{R}{R+G+B} \\ g &= \frac{G}{R+G+B} \\ b &= \frac{B}{R+G+B} \end{aligned} \quad (6)$$

3 Shadow Model

Shadow occur when the object totally or partially occludes the light coming from the light source and reaching the background. The shadow has two parts to it, called umbra and penumbra. The umbra corresponds to the area where the direct light is almost totally blocked, whereas the area where light is partially blocked is called penumbra. Reflectance of surfaces is dependent on the wavelength of the light falling on it, but is independent of the intensity of light impinging on it. If the light intensity falling on the object decreases or increases by a constant then the light reflected from it will also change by a constant additive term. Using this hypothesis we form the equations relating the different color channels of a background pixel when it is in shadow region and when it is not.

$$\begin{aligned} R_{shadow}(i, j) &= R_{bg}(i, j) + n + e \\ G_{shadow}(i, j) &= G_{bg}(i, j) + n + e \\ B_{shadow}(i, j) &= B_{bg}(i, j) + n + e \end{aligned} \quad (7)$$

where n is a constant and e is error due to signal acquisition and minor variation in reflectance property of the material due to light intensity variation. R_{bg} , G_{bg} , and B_{bg} are the original background values of the pixel and R_{shadow} , G_{shadow} , and B_{shadow} are the pixel values when it is in shadow. This model is verified by empirical observations. Figure 1 shows *RGB* plots of different

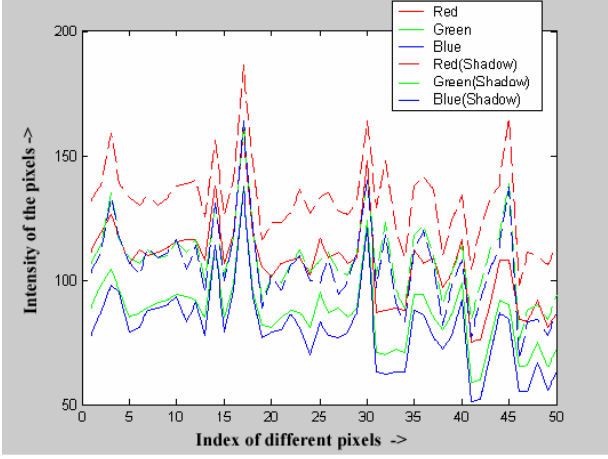


Figure 1. The plot shows *RGB* intensity values of pixels when they are in shadow region and intensity values of the same pixels when they are not in shadow region. The *RGB* intensity values of the pixels in the shadow region are a constant additive term less than when the corresponding pixels is not in shadow.

pixels when they are in shadow region and when they are not for an image sequence. The plots clearly shows that the *RGB* values of the pixels in shadow region follows the *RGB* values, by almost a constant additive term less than when the same pixels are not in shadow. But if the light impinging on a surface is less than a threshold then usually its color information lost. Its not picked up by the camera or even the human eye, which is the best camera in existence. So its impossible to model umbra regions of a shadow.

4 Background Model and Segmentation Criterion

The intensity and color of each background pixel are monitored over time for a few seconds with no foreground object. Each channel of a pixel in different color spaces is modeled using a normal distribution $N(\mu, \sigma^2)$. also the image noise over time can be modeled by a zero mean normal distribution $N(0, \sigma^2)$. This normal distribution model for the pixel is used underlying model for many background subtraction techniques [8, 9, 14, 13], [14, 13] use multiple normal distribution to model a pixel.

4.1 Algorithm for Segmentation

The procedure for segmentation involves the following steps.

1. Obtain the mean and standard deviation of each pixel in the background images.
2. In the current frame apply the following check to each pixel. If the pixel is in a region defined by its background model mean and variance then declare it to be a background pixel. For example for *RGB* space, if

$$|R(i, j) - \mu_R(i, j)| < c \times \sigma_R(i, j) \quad (8)$$

$$|G(i, j) - \mu_G(i, j)| < c \times \sigma_G(i, j)$$

$$|B(i, j) - \mu_B(i, j)| < c \times \sigma_B(i, j) \quad (9)$$

then the pixel is declared to be background. Where c is a constant, (i, j) is the pixel index, μ is the mean, and σ the standard deviation. In our implementations $c = 5$.

3. Those pixels which are not classified as background are tested for being shadow or foreground and then classified accordingly. Different color spaces have different segmentation criterion. For example in the *RGB* color space the following criterion is used:

$$\text{if } \begin{cases} 0.9 < \frac{R(i, j) - \mu_R(i, j)}{G(i, j) - \mu_G(i, j)} < 1.1 \\ \text{and} \\ 0.9 < \frac{G(i, j) - \mu_G(i, j)}{B(i, j) - \mu_B(i, j)} < 1.1 \end{cases} \quad (10)$$

then pixel (i, j) is shadow pixel otherwise it is foreground pixel.

4. Finally, the rest of the pixels are classified as foreground.

Similar procedure holds for the rest of the color spaces, except that the segmentation criterion for shadows are quite different.

5 Results and Discussions

We compare the results of our experiments by two quantitative criteria and one qualitative evaluation based on human observation. Section 5.1 evaluates the goodness of the segmentation based on true, false and misses in classification of pixels as foreground. Section 5.2 evaluates the results in terms of true detection, false detection, misses and splits of the foreground object. In Section 5.3 we evaluate based on visual observation of the results when compared with the real scene.

5.1 Accuracy Measurement Results

An artificial foreground object with multiple random colors is overlayed on to an image sequence as shown in

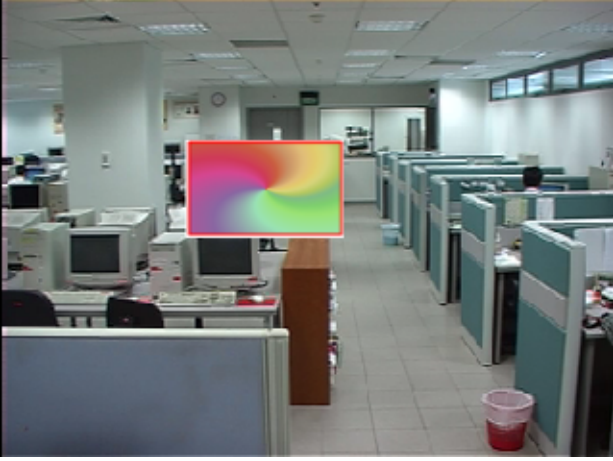


Figure 2. A frame from the test sequence. The size of random colored foreground is known. This test sequence is used to obtain the results shown in Table 1.

Figure 2. We apply the foreground detection scheme to it in all the color spaces. The results are follows:

The results shown in Table 1 is analyzed as follows.

- *RGB* color space detects bright light area as a foreground object when it is actually a background and therefore this adds on to the false detection in *RGB* space. True detection and misses in *RGB* space is bad, due to the problem of clipping.
- *HSV* color space is also sensitive to bright light areas and classifies those pixels as shadow or foreground. This leads to high false detection and *HSV* also has low true detection and a high miss rate.
- *YC_rC_b* or *YUV* color space is insensitive to bright light areas. Gives high true detection very low miss rate and also low false detection rate.
- *XYZ* color space is also insensitive to bright light areas but the results of the true detection percentage and miss rate is not as good as *YC_rC_b* color space. *XYZ* has the lowest false detection rate among all the color spaces examined. If there is a need to have a least false detection application, this is a correct choice.
- *rgb* color space is insensitive to bright light area but when the foreground passes over the bright light background then there are lot of errors in classification. Therefore this leads to a slightly lower true detection and a slightly higher misses on average. The false detection rate is better than *RGB* and *HSV* color spaces.

Color Space	TD	Misses	FD
<i>RGB</i>	97.32%	2.68%	0.73%
<i>HSV</i>	88.12%	11.88%	5.76%
<i>YC_rC_b/YUV</i>	97.72%	2.28%	0.35%
<i>XYZ</i>	96.7%	3.3%	0.255%
<i>rgb</i>	91.46%	8.54%	0.365%

Table 1. Accuracy Measurement Results. This table shows the result of True Detection (TD) Misses and False Detection (FD) of the pixels averaged over many frames. TD is the ratio of the total number of true detection to the total number of foreground pixels. Misses is the ratio of the total number of misses to the total number of foreground pixels. FD is the ratio of the total number of false detection to the total number of background pixels. The *YC_rC_b* color space gives optimal result.

The *YC_rC_b* is therefore a good choice for further work as it has the highest true detection rate and lowest miss rate, as shown in Table 1 on pixel level classification. This is usually the desired output from a segmentation algorithm.

5.2 Connected Component Results

The output image has been investigated for its number of connected components. The segmentation results of different color spaces are processed with a median filter and then labeled into different blobs by 8 connectivity and then the blobs of size below a threshold have been deleted. The results of this processing for *YC_rC_b* color space is shown in Figure 3. The Figure 4 shows the result of the blob detection by *XYZ* space to show the phenomena of splits.

The results shown in Table 2 is analyzed as follows:

- *RGB* color space gives a high true detection rate, a low false detection and low number of split. This color space is considered one of the better color spaces when compared to the others.
- *HSV* color space gives a high true detection and a low number of split but it has the highest false detection. Therefore this color space is not satisfactory.
- *XYZ* color space gives the least number of true detection and the highest number of split which is bad as can be seen in Figure 4(b). It has the lowest number of false detection which is good. Since it is unable to detect many foreground objects, this color space is not satisfactory.

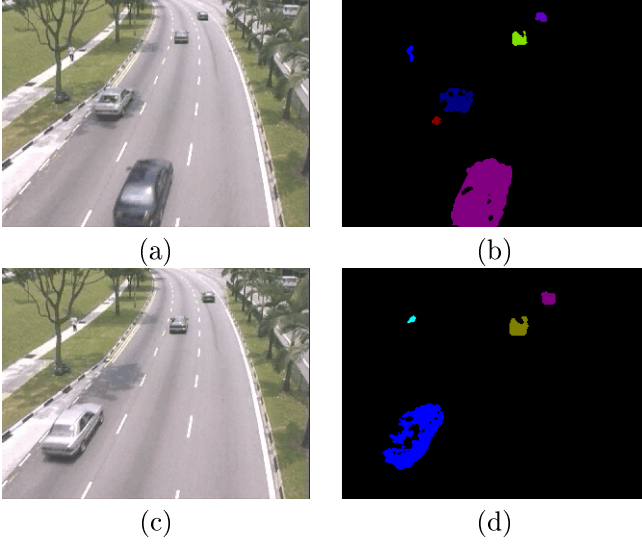


Figure 3. Figure 3(a) and (c) shows the original image before processing. Figure 3(b) and (d) shows the result of $YCrCb$ color space for the connected components.

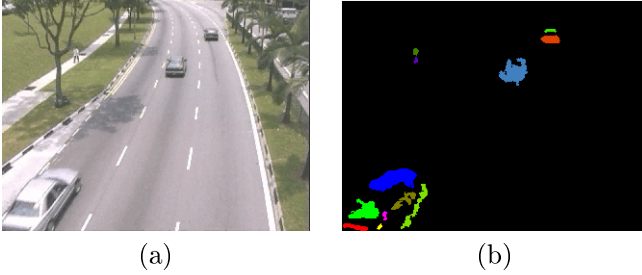


Figure 4. Figure 4(a) shows the original image before processing. Figure 4(b) shows the XYZ color space result where splits are observed on the bottom left vehicle.

- $YCrCb$ color space gives a high true detection, a reasonably low false detection and the lowest number of split. Therefore this color space is satisfactory.
- rgb color space gives an almost averaged result compared to the rest of the color spaces. Since the result is not the best, it is rejected.

The $YCrCb$ is therefore a good choice for our application as it has reasonably high true detection and low misses and splits which is what we wanted as it is able to detect foreground objects reasonably well.

5.3 Visual Comparison of Results

Various traffic data have been used for experimentation. We have found that the $YCrCb$ color space gives

Color Space	TD	Misses	FD	Splits
RGB	72.5%	0%	25%	41.25%
HSV	88.75%	0 %	151.25%	22.5%
$YCrCb$	85%	0%	36.25%	35%
XYZ	43.75%	0%	5%	196.25%
rgb	67.5%	0%	41.25%	96.25%

Table 2. Connected Component Results. These are the average blob detection rates of the foreground objects. True Detection(TD) is the ratio of the correctly detected blobs with the total number of blobs detected. False detection (FD) is the ratio of the falsely detected blobs with the total number of blobs detected. Splits is the number of extra objects detected due to splits to the number of original objects which underwent split. Combining all results $YCrCb$ color space seems to give the best results.

fairly good results as compared to the rest of the color spaces. The result of segmentation and shadow classification is shown for $YCrCb$ space is shown in Figure 5

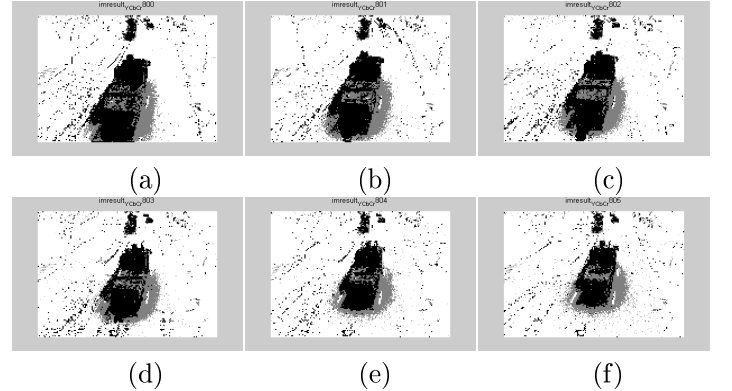


Figure 5. a, b, c, d, e and f are six continuous frames for which segmentation in $YCrCb$ color space is shown. A minority of the shadow pixels have been classified as background. A minority of the foreground image have been classified as shadows. The segmentation result is better than other color spaces.

We observed from our experiments that in RGB color space, the algorithm yielded false classification in bright light backgrounds. When lighting condition is not very bright, foreground pixels were misclassified as shadows. For HSV color space, the algorithm yielded similar misclassification of the brightly light background as foreground and was not able to detect shadows properly. For XYZ color space, the algorithm gives false classification on brightly light backgrounds and shadows were

not well detected. The YC_rC_b color space algorithm is robust to brightly light backgrounds and was able to detect shadows quite precisely. Since YC_rC_b and YUV have the same equations when dealing with shadows, both of them gave good detection of shadows. For rgb color space, the algorithm by itself is not able to detect shadows and but it is able to detect foreground objects well.

We conclude that YC_rC_b is the most suitable color space for the detection of foreground and shadow in traffic image sequences and is robust to camera artifacts.

6 Conclusion and Future Works

This work has empirically compared the suitability of different color spaces for segmentation of foreground and detection of shadow for shadow removal. We have investigated five different color spaces for their suitability in detecting shadows and found YC_rC_b space to be a suitable solution for foreground segmentation and shadow detection. The false classification due to shadow and highlights were greatly solved. False classification due to shadows and highlights takes place due to intensity change of the background pixels. Shadows are due to the obstruction of the light source by the foreground. Highlights are camera artifact due to the camera design of maintaining constant brightness in image sequence. The model can handle situations where the background of the scene is not completely static but contains motions such as movement of tree leaves and branches. There are due to specular reflections in the background and fast illumination changes. The next step in this study will be to study how different background modeling techniques work with different color spaces. There is literature which has compared different background modeling techniques and proposed more sophistication to existing techniques but there has not been much work to study how different color spaces behave with different modeling techniques. Using the results of this work have successfully designed systems which can track vehicles in real time. We are looking into behavior analysis from the deciphered track of the vehicles.

References

- [1] A. Elgammal, D. Harwood, and L. Davis. Non-parametric model for background subtraction. *IEEE ICCV'99 FRAME-RATE WORKSHOP*, page <http://www.eecs.lehigh.edu/FRAME/Elgammal/bgmodel.html>, 1999.
- [2] N. Friedman and S. Russell. Image segmentation in video sequences: A probabilistic approach. *Thirteenth Conf. on Uncertainty in Artificial Intelligence (UAI)*, August 1-3, 1997.
- [3] Gonzalez and R. E. Woods. *Digital Image Processing*. Prentice Hall, 2002.
- [4] T. Horprasert, D. Harwood, and L. Davis. Statistical approach for real-time robust background subtraction and shadow detection. *IEEE ICCV'99 FRAME-RATE WORKSHOP*, page <http://www.eecs.lehigh.edu/FRAME/Horprasert/index.html>, 1999.
- [5] R. M. Inigo. Application of machine vision to traffic monitoring and control. *IEEE Transactions on Vehicular Technology*, 38(3):112–122, August 1989.
- [6] M. Kilger. Shadow handler in a video-based real-time traffic monitoring system. In *IEEE Workshop on Application of Computer Vision*, pages 11–18, 1992.
- [7] D. Koller, J. Weber, T. Huang, J. Malik, G. Ogasawara, B. Rao, and S. Russell. Towards robust automatic traffic scene analysis in real-time. In *Proceedings of International Conference on Pattern Recognition*, Israel, November 1994.
- [8] P. Kumar, K. Sengupta, and S. Ranganath. Detecting humans: analysis and synthesis as a pattern recognition problem. In K. N. Ngan, T. Sikora, and M.-T. Sun, editors, *Visual Communications and Image Processing 2000*, volume 4067, pages 720–730, Bellingham, Washington 98227-0010 USA, June 2000. SPIE-The International Society for Optical Engineering.
- [9] S. J. McKenna, S. Jabri, Z. Duric, A. Rosenfeld, and H. Wechsler. Tracking groups of people. *Computer Vision and Image Understanding*, 80:42–56, 2000.
- [10] G. Medioni, I. Cohen, F. Bremond, S. Hongeng, and R. Nevatia. Event detection and analysis from video streams. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(8):873–889, August 2001.
- [11] A. Prati, I. Mikic, C. Grana, and M. M. Trivedi. Shadow detection analysis for traffic flow analysis: A comparative study. In *Submitted to IEEE Int'l Conf. on Intelligent Transportation Systems*, Massachusetts Institute of Technology, August 2001.
- [12] C. E. Smith, C. A. Rishards, S. A. Brandt, and N. P. Papanikolopoulos. Visual tracking for intelligent vehicle-highway systems. *IEEE Transactions on Vehicular Technology*, 45(4):744–758, November 1996.
- [13] C. Stauffer and W. Grimson. Adaptive background mixture models for real time tracking. In *Proceedings of Computer Vision and Pattern Recognition*, pages 246–252, June 1999.
- [14] C. Stauffer, W. Grimson, and R. Romano. Using adaptive tracking to classify and monitor activities in a site. In *Proceedings of Computer Vision and Pattern Recognition*, pages 22–29, June 1998.
- [15] K. Toyama, J. Krumm, B. Brumitt, and B. Meyers. Wallflower: Principles and practice of background maintenance. In *Proceedings of International Conference on Computer Vision, Greece*, pages 255–261, September 1999.