Vehicle Count Estimation on Traffic Images

Jan Andre Alcantara[1], Niño Byron Aguirre[2],John Ferdinand Trinidad[3], Joel Ilao Ph.D.[4]

[1,2,3,4,] *Computer Technology Department, College of Computer Studies*
De La Salle University – Manila
jan_alcantara@dlsu.ph, bbbyron25@yahoo.com, john_ferdinand_trinindad@dlsu.ph, joel.ilao@dlsu.edu.ph

Abstract: Computer Vision has been used as an effective approach to developing cost-efficient traffic surveillance systems. Traffic surveillance systems are now used for vehicle detection, tracking, recognition and counting. However, performing these tasks are not easy especially when challenges in traffic arise such as varying lighting conditions, effect of shadows and occluded vehicles that highly affect the performance of almost all traffic surveillance systems. In this paper, we approach the challenges of varying lighting conditions through a background modelling scheme that updates the background image, a shadow removal scheme to remove shadows among the vehicles in the images, and a segmentation scheme to segment occlusive vehicles. In this paper, background modelling is performed by removing the moving foreground objects and leaving an empty background image. Image noise is removed through a 3x3 median filter. Shadow removal is done by converting the RGB image to its LAB equivalent image – having three channels: L for the lightness and the remaining A and B channels are arbitrary color channels. For shadow detection, if the sum of the means of A and B is less than a threshold (256 in this paper), then the pixels in L that are less than or equal to the difference of its mean and one third of its standard deviation can be classified as shadow pixels. The system was able to count vehicles in a traffic video recorded from DLSU's Archer's Eye having 704x480 resolution. Furthermore, system performance with regards to the detection and counting of vehicles was affected by the presence of non-vehicle entities in the recorded traffic video.

Key Words: Vehicle Counting; Vehicle Detection; Background Modelling;, Shadow Removal; Noise Removal;

1. INTRODUCTION

Computer Vision is commonly used in various applications and one of the many is in traffic surveillance. Today, computer vision has been the focus of traffic related researches such as vehicle detection, recognition, tracking and counting as it provides its users efficient ways to perform various tasks when hardware availability is limited.

Additionally, the use of computer vision for these applications implies a cost-efficient way for monitoring traffic instead of investing to different traffic sensors – that are costly and more susceptible to physical interruptions from the environment.

Today, there are numerous researches on traffic surveillance and traffic surveillance systems. (Ilao, Dang, Daquioag, Ramos, & See, 2007) proposed methods in tracking vehicles with the use of hypothesis generation and hypothesis verification modules on Philippine roadways, arriving at an overall performance of 74.58% for correctly identified vehicles. (Lai, Leong, Ortouste, Yu, & Ong, 2008) proposed a vision based system for traffic analysis by obtaining entities that contribute to traffic which can be used as traffic parameters for traffic management and informatio systems. They also targeted Philippine roadways and resulted to 68.75% accuracy in their proposed system.

Being able to count vehicles in traffic is a challenging aspect in traffic surveillance. The scenarios wherein vehicles are occluded and presence of shadows, and varying lighting conditions affects system performance, making the traffic surveillance less effective.

Shadows within images contribute to the possible errors during detection of vehicles, which later affects other tasks such as vehicle tracking, recognition and counting. Thus, removing these shadows greatly helps improve the accuracy and performance of the system. According to (Suny & Mithila, 2013), shadows among objects can be classified into two classes: cast– and self– shadows. Cast shadows are the ones that we see (usually in the surfaces like the ground), projected by the object and opposite from the direction of the direct light source. Self shadows, however, are the parts of an object that are not illuminated by direct light.

In this paper, we aim to address the challenges brought by occlusion, shadows and lighting changes.

2. PREPROCESSING

The pre-processing stage in this paper involves the acquisition of the traffic images by converting a traffic video into frames. Modelling of the static background (i.e. the empty highway) is then performed, in preparation to removal of the moving foreground objects. Removing noise from each frame is done before removing the shadows via a 3x3 median filter. After the noise and shadows were removed, background subtraction is performed using the obtained background image.

2.1 Background Modelling

When using computer vision for traffic surveillance or even in simple object and motion detection systems, background subtraction has been tested as an effective approach in separating the non-moving objects (background) from the moving objects (foreground). In background subtraction, it is often necessary to have a background model that will not contain any foreground objects so that separating the two will be easier afterwards. (Lai, Leong, Ortouste, Yu, & Ong, 2008) made use of Codebook Model Algorithm to identify their background image through a sequence of training images and obtain a static background image. (Ilao, Dang, Daquioag, Ramos, & See, 2007), however, suggested a background modelling approach that updates the background in a traffic video. Please refer to Eq. 1, for equation used for background modelling.

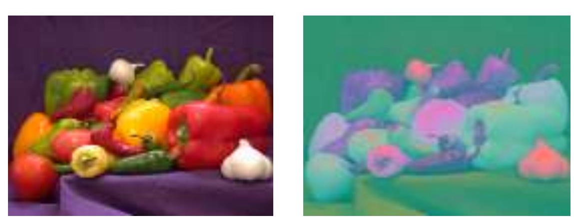
$$\beta_{t+1} = \beta_t(1-\alpha) + \alpha(F_t(1-D_t) + T(D_t)) \qquad \text{(Eq. 1)}$$

where:

β_{t+1} = next background estimate
β_t = current background estimate
αe = learning rate constant set to 0.05
F_t = current frame processed
D_t = binary moving mask

2.2 Shadow Detection and Removal

To detect shadows, the RGB image is converted into LAB equivalent image (Figure 1).



(a) *(b)*

Figure 1 (a)RGB image, (b)LAB equivalent image. Both used 'peppers.png'

The LAB color space has three channels where L is the lightness and A and B are two arbitrary color channels. According to (CIELAB - Color Models - Technical Guides, 2000), the L channel of the LAB color space ranges from 0 to 100, corresponding to different shades from black to white (Figure 2). The A and B channel are both ranging from -128 to + 127, indicating the green to red ratio and blue to yellow ratio, respectively.

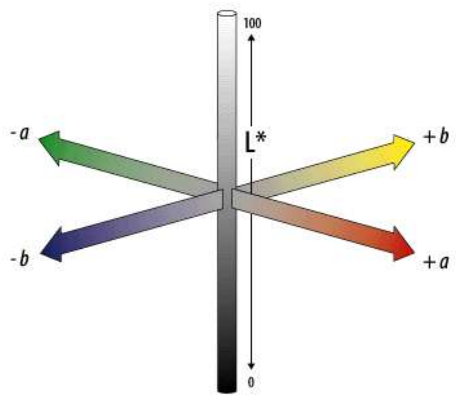


Figure 2 LAB Color Space

In this paper, we follow the suggested approach by (Suny & Mithila, 2013) in detecting and removing shadows from an RGB image.

Shadow detection starts by first converting the RGB image to a LAB image. The next step is to compute for the mean values of the pixels in L, A and B planes of the image separately. Their algorithm in classifying shadow pixel from non-shadow is as follows:

If $mean(A) + mean(B) \leq 256$, then pixels $P(x,y)$ in $L \leq (mean(L) - d(L)/3$ can be classified as shadow pixel and others as non-shadow pixel.

where:
A,B,L = three channels of the LAB image
P(x,y) = pixel with a value in the L channel
d = standard deviation

(Suny & Mithila, 2013) suggested in their paper to correct the chromaticity values of the image in removing shadows. Using a mean shift algorithm, they segmented the image according to its color values. Since the segments of the shadow regions are adjacent to the non-shadow regions, (Suny & Mithila, 2013) suggested to choose the closest in chromaticity to the segment of interest. They scaled the shadowed segment's chromaticity values so that the average of the chromaticity in that region matches the average of the chromaticity of the chosen non-shadow region.

3. VEHICLE COUNTING: PROPOSED APPROACH

In this paper, we propose an approach in counting vehicles by combining the methods proposed by (Ilao, Dang, Daquioag, Ramos, & See, 2007) and (Suny & Mithila, 2013). Please see (Figure 3) for the System block diagram.

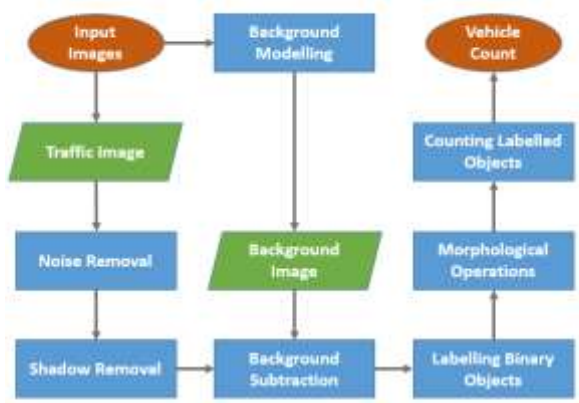


Figure 3 Diagram for Vehicle Counting

By making use of the background modelling approach suggested by (Ilao, Dang, Daquioag, Ramos, & See, 2007), we get the empty background image β_t (Eq. 1), with the moving foreground objects removed from the scene. Having obtained the background image, we then make use of a 3x3 median filter to remove the noise from the RGB images and the shadow removal approach suggested by (Suny & Mithila, 2013) to remove the shadows in those images before performing background subtraction. After separating the foreground objects from the background, we convert the resulting image into its binary equivalent image through a given threshold. After having the binary image, we apply some morphological operations such as erosion and dilation to remove excess pixels and fill in some gaps in the image. Labelling the binary image through the number of connected regions is then performed before getting the count.

Furthermore, since vehicles are not the only objects that contribute to traffic, smaller sized moving objects such as humans and animals may also be detected. With this, we set a minimum area of 1200 pixels for the labelled regions so that objects in the image with smaller area would not be counted.

4. EXPERIMENETAL RESULTS

The traffic video used has a resolution of 704x480 pixels. Figure 4 shows the vehicle counting of the proposed system.

Figure 4 Top to bottom: image from traffic video, background image, subtracted image, subtracted image after morphological operations, image with counted vehicles.

The video images in the figure above were screenshots of the first (left column) and five hundredth (right column) frames from the traffic video, respectively. The background images were constructed via 300 frames in the video. The background image on the left column was constructed from frames 1 to 300, while the background image on the right column of the figure was from frames 500 to 800 of the input video.

5. CONCLUSION

From the last figure (please see Figure 4), the proposed approach was able to count vehicles from traffic images via the bounded boxes around them. However, since the location of the traffic camera includes the surveillance of pedestrians on roadside, the performance of the system is affected. Hence, future works may include the exclusion of non-vehicles in traffic during the vehicle detection process.

6. ACKNOWLEDGMENTS

The authors would like to thank the Information Technology Services Office of DLSU for providing them recorded videos from the Archer's Eye, for this study.

7. REFERENCES

Chiu, C.-C., Ku, M.-Y., & Wang, C.-Y. (2010). Automatic Traffic Surveillance System for Vision-Based Vehicle Recognition and Tracking. *Journal of Information Science and Engineering*, 611-629.

CIELAB - Color Models - Technical Guides. (2000). Retrieved from dba.med.sc.edu: http://dba.med.sc.edu/price/irf/Adobe_tg/models/cielab.html

Ilao, J., Dang, M., Daquioag, A., Ramos, T., & See, F. (2007). Hypothesis Generation of Vision-Based Vehicle Tracking System. *7th Philippine Computing Science Congress (PCSC 2007).*

Jacques, J., & Jung, C. (2005). Background Subtraction and Shadow Detectio in Grayscale Video Sequences. *18th Brazilian Symposium on Computer Graphics and Image Processing*, (pp. 189-196).

Lai, F., Leong, A., Ortouste, R., Yu, P., & Ong, C. (2008). Vision-Based Intelligent System for Traffic Analysis. *Thesis Document, College of Computer Studies*.

Suny, A., & Mithila, N. (2013, July). A Shadow Detection and Removal from a Single Image Using LAB Color Space. *International Journal of Computer Science Issues, 10*(4), 270-273.