Automatic Road Traffic Density Estimation using Image Processing Algorithms

Daniel Arani Osuto, Heywood Ouma Absaloms and Edward Ng'ang'a Ndungu

Abstract: In the modern world, urban centres are growing at a very high rate. Growing with them is road traffic congestion. Traffic jams, especially at peak hours, have become routine. As a result, traffic management is one of the most pressing issues in today's towns. Several alternatives are being sought to deal with the problem. These include: expansion of road networks, regulating the number of vehicles on the roads, and deployment of Intelligent Transportation Systems (ITSs). Other than the ITSs, the other alternatives (however effective) have many practical challenges in their implementation. ITSs are based on a wide range of technologies such as loop sensors and video surveillance systems. Vision based ITSs have proved advantageous over the traditional methods based on loop sensors. In these modern systems, video surveillance cameras are installed along the roads and road intersections where they are used to collect traffic data. The data is then analysed to obtain traffic parameters such as the road traffic density.

In this paper, a simple and elegant approach for estimating the road traffic density during daytime using image processing and computer vision algorithms is presented. The video data collected is first broken down into frames which are then pre-processed in a series of steps. Finally, the vehicles are detected and extracted from the images and counted. Then the traffic density is obtained as the number of vehicles per unit area of the road section. The proposed approach was implemented in MATLAB R2015a and average vehicle detection accuracies of 96.0% and 82.1% were achieved for fast moving and slow moving traffic scenes respectively.

Keywords: *Image processing, intelligent transportation systems, Road traffic density*

I. INTRODUCTION

Vision based ITSs have attracted much attention among researchers in the last two decades. Important road traffic parameters such as the road traffic density can be obtained by these systems and over the years, many approaches have been proposed [1]. The overwhelming majority of these approaches are dependent on motion detection or background modelling and subtraction to detect vehicles. This limits their application only to free flowing traffic scenes or scenes with static backgrounds. In cases where the traditional static background subtraction method is used, changing lighting conditions are not factored in [2], and segmentation results may thus not always be reliable. Dynamic background modelling is excellent to a large extent in handling changing scene conditions [2], [3]. Unfortunately, this method cannot be used for stationary traffic, a common problem in the developing world. In addition to this, many of the proposed approaches do not consider the performance in different illumination conditions. In this paper, the background modelling and subtraction methods are avoided. Instead, a combination of simple but robust image processing and computer vision algorithms are used to extract the vehicles from the traffic video data for traffic density estimation. This enables the algorithm to effectively handle stationary traffic scenes.

II. RELATED WORK

Vision based vehicle detection has long been explored. Although initial efforts were not so successful, lately much improvement has been achieved. Commercial software for this purpose exist but are dogged with problems such as inability to handle vehicle occlusions from the camera's view [4], [5], limited functionality in severe weather conditions [6], [7], shadows and night detection [8], [9]. Inter-system compatibility is one other drawback associated with today's video analytic algorithms thus severely limiting their deployment, as they do not generally work with already installed hardware unless the hardware is from the same vendor as the algorithms. Open platforms such as Open Network Video Interface Forum (ONVIF) have been formulated [10], but so far remain at the specifications stage and are not yet standardized. Each vendor understands these specifications differently and as a result, the integration of their products remains at the very basic level.

Ambardekar, et al. [11] and, Sivaraman and Trivedi [1] give good general overviews of the state of the art in vehicle detection and classification algorithms. Avery, et al. [9] used images from uncalibrated video cameras to count and classify trucks and heavy vehicles on the basis of length. The vehicles were extracted using a dynamic background subtraction method and then their lengths were extracted only when they reached a particular point in the scene while traveling in a straight line. The limitation of this approach is that it cannot work in heterogeneous traffic scenes where the vehicles are not moving in a straight line. Similarly, Pancharatnam and Sonnadara [12] used adaptive

D. A. Osuto, Department of Telecommunication and Information Engineering, JKUAT (phone: +254710385733; e-mail: danielosuto@gmail.com).

H. A. Ouma, Department of Electrical and Information Engineering, University of Nairobi (e-mail: habsaloms@yahoo.com).

E. N. Ndungu, Department of Telecommunication and Information Engineering, JKUAT (e-mail: ndunguen@gmail.com).

background subtraction to detect moving vehicles and tracked them using their bottom coordinates before counting and classifying them on the basis of size into large, medium and small classes. Although good results were reported, similar to [8], the system relied on the assumption that a vehicle will only occupy one lane at a time for its effective performance. It was shown that great errors occur when this is not true.

In [13], a video-based vehicle detection and classification system for real time traffic data collection using uncalibrated video cameras was proposed. The system eliminated the need for the complicated camera calibrations. It struck a good balance between algorithm complexity and effectiveness in real time applications. The paper also noted one other critical limitation of all background modelling and subtraction based algorithms (both static and dynamic) for foreground extraction: they do not account for transient lighting changes in the scene. This is confirmed by the results in [14]. Ince [15] used invariant moments and shadow aware foreground masks to count vehicles and classified them using a perspective projection of the scene geometry. The algorithm was tested on real world data and showed to be computationally efficient.

III. THE PROPOSED ALGORITHM

A multi-stage vehicle detection and counting algorithm for handling both free flowing and slow moving traffic or stationary traffic is proposed. Fig.1 illustrates the concept of the algorithm.

First, the collected colour video data is broken down into constituent frames which are then converted into greyscale so as to simplify and facilitate subsequent processing. The vehicles are then extracted from the video frames and their negatives using a Laplacian of Gaussian (LoG) edge detection method and mathematical morphology. The vehicles so obtained are counted and their number used to calculate the traffic density as the number of vehicles per unit area of the road section at any given time.

A. Negative transformation

For each grey frame extracted from the video, its negative is computed. This is to ensure that as much relevant edge detail as possible is extracted in the segmentation stage, thus minimizing spurious edge discontinuities. Fig. 2 shows an image from a typical traffic scene and its negative.

(a) (b)

Fig. 2. (a) Original Frame (b) Negative Frame

B. ROI mask modelling

One of the extracted frames is used to model a Region of Interest (ROI) polygon. This polygon is ultimately used to mask the processed binary frames so as to limit the counting and classification of vehicles to those found only within the region of interest. The size of this polygon is chosen empirically such that the vehicle intra-class variations are minimized. Fig. 3 shows a traffic scene and the modelled ROI mask for the free flowing traffic.

(a) (b)

Fig. 3 (a) Frame (b) ROI Mask

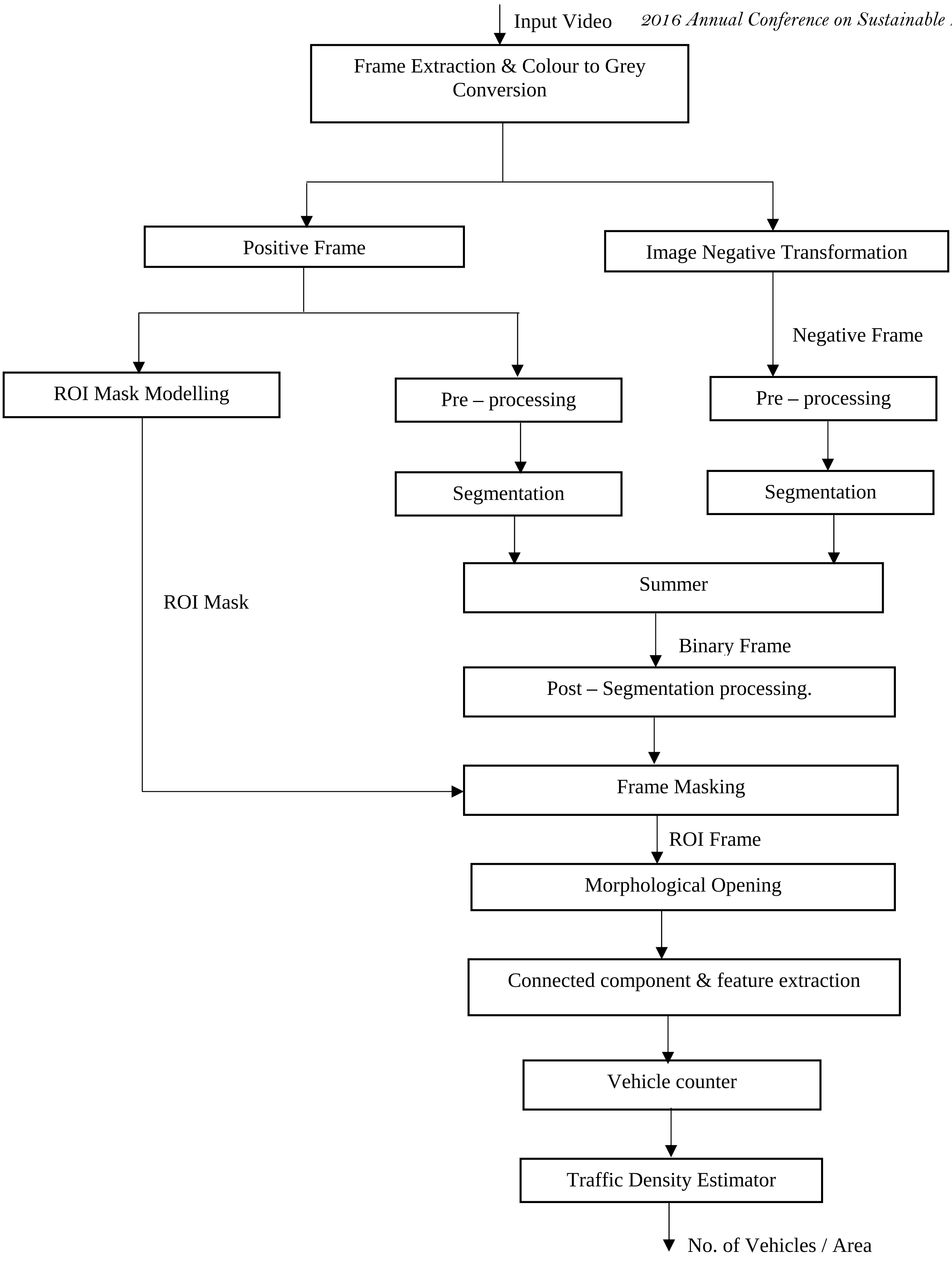

Fig. 1. The proposed algorithm

C. Top-hat transformation

The segmentation performance is improved by compensating for non-uniform illumination of the scene using the morphological top-hat transformation prior to the segmentation stage. This is computed as

$$g(x,y) = f(x,y) - (f(x,y) \circ b(x,y)) \qquad (1)$$

Where *g(x, y)* is the uniform background frame, *f(x, y)* is the input frame and $f(x,y) \circ b(x,y)$ is the morphological opening of *f(x, y)* using a structuring element (SE), *b(x, y).* The size of this structuring element is chosen such that it is larger than any object of interest in the scene so as to avoid deletion of any vehicle in the subtraction process.

D. Image smoothing and blurring

The uniform background frame is smoothed using a median filter to remove random noise and then aggressively blurred using a Gaussian filter so as render 'noise' edges into the background and therefore reduce the chances of their detection. Finally, the blurred frame is contrast enhanced linearly so as to emphasize the edges while preserving the mean intensity values of the frames using a contrast stretching algorithm prior to segmentation.

E. Image segmentation

To extract objects in both the frame and its negative, the Laplacian of a Gaussian (LoG) edge detection method is used due to its excellent edge detection properties and relative simplicity [16]. This preserves generality unlike the trial and error thresholds normally used in many of the reported approaches. The LoG of a two dimensional image is computed as

$$\nabla^2 G(x,y) = \left[\frac{x^2 + y^2 - 2\sigma^2}{\sigma^4}\right] e^{-\frac{x^2+y^2}{2\sigma^2}} \qquad (2)$$

Where ∇^2 is a Laplacian operating on the Gaussian smoothed image, *G(x, y)* and σ is the standard deviation of the image pixel intensities. Fig. 4 and Fig. 5 show video frames so segmented.

(a) Original frame

(b) Segmented frame

Fig.4. Free flowing traffic scene

(a) Original frame

(b) Segmented frame

Fig. 5. Slow moving traffic scene

F. Summation

After segmentation, the two branches are added to eliminate double counts and to ensure that as many objects are detected as possible. This addition is possible since the frame and its negative are spatially registered and therefore the objects which occur simultaneously in both the frame and its negative reinforce each other. The output of the summer give the complete edge map, and therefore the binary image of the frame.

G.Post-Processing And Feature Extraction

The obtained binary frame is then subjected to morphological filtering. First, the segmented binary frame is closed so as to eliminate any spurious disjoints between connected components. Then the holes in the connected components are filled to ensure that true sizes of objects are used in subsequent stages. Next, a skeletonizing algorithm is run once before pruning the image to get rid of spurious branches of objects that result after segmentation. Then, the processed frame is masked using the modelled ROI mask so as to limit the counting and classification to the objects found in the region of interest only. Finally, the irrelevant small objects within the ROI such as pedestrians are deleted using a morphological opening operation. The morphological opening is done after masking the frame so as to ensure that unwanted objects on the border of the region of interest are deleted as well.

H. Vehicle counting and traffic density estimation

The resulting connected components in the fully processed frame represent vehicles on the road at that time. These components are counted to give the total number of vehicles on the given section of the road at the given time. Fig. 6 shows the result of the count for the frame shown in Fig. 4, and Fig. 7 (b) shows the result of the count for the frame shown in Fig. 5. Fig. 7 (a) shows the ROI polygon used for the slow moving traffic scene. With this value, the road traffic density can be calculated as

$$\text{Traffic Density} = \frac{\text{Number of vehicles}}{\text{Area of traffic scene}} \quad (3)$$

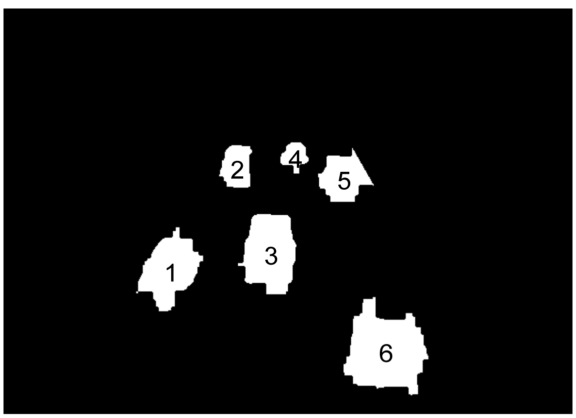

Fig. 6. Post-processed frame

Fig. 7(a). ROI polygon

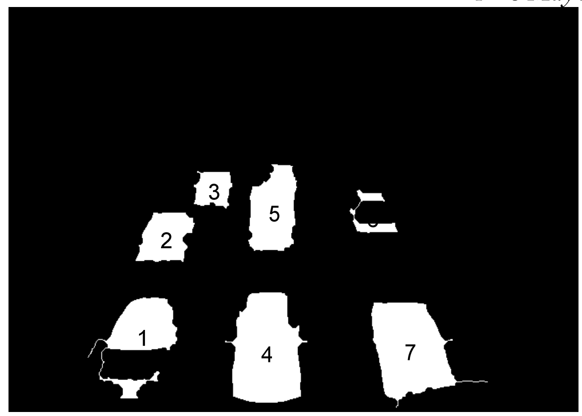

Fig. 7(b). Post-processed frame

Note from Fig. 7(b) that object 5 consists of two vehicles (Fig. 7(a)) but are represented as one vehicle. This is due to the fact that the car is occluded from the camera's view by the larger vehicle in front of it. This scenario highlights the importance of proper installation of cameras meant for traffic management systems.

IV. Experimental Results

Video data from a road section was collected using a 5MP camera mounted above the road under which the subject vehicles passed. In order to assess the performance of the system under various illumination levels across the day, the data was collected at 0630hrs before the sun is up; 1230hrs when the sun is overhead and the shadows are negligible, and 1630hrs when both reflections from the road surface and shadows are strongest. Data was also collected from a traffic scene that involved very slow moving traffic so as to assess the performance of the proposed system on such traffic scenes or on stationary ones. Each collection period lasted 20 minutes, resulting in at least 36000 frames each time. The vehicle detection results for both free flowing traffic and slow moving traffic datasets are given in Table 1.

Table 1. A summary of the detection results

Dataset	Total no. of vehicles in the video data	Total no. of correct detections	Total no. of wrong detections	Detection accuracy
Morning	220	209	11	95.0%
Midday	306	291	15	95.1%
Evening	438	425	13	97.0%
Overall (Free flowing traffic)	964	925	39	96.0%
Slow moving traffic	246	202	42	82.1%

The vehicle detection accuracy for the slow moving traffic dataset was generally poorer than those for the free flowing traffic datasets as seen Table 1. The reason for this is that occlusions were more severe in the slow moving traffic scene than for the free flowing traffic scenes. Consequently, at times, two or even more vehicles could be detected as one vehicle rather than as separate vehicles as illustrated in Fig. 7.
The relatively low camera position was the main cause of detection errors since it was difficult to 'see' the spaces between the vehicles on the same lane when the vehicles involved were very close together as seen in Fig. 7. Attempts to raise the camera position were not successful due to practical limitations.

V.CONCLUSION

This paper attempted to solve the problem of vehicle detection and counting in natural traffic scenes using video surveillance systems for both free flowing and slow moving or stationary traffic scenes. Stationary or slow moving traffic scenes have little reported about them and the majority of the proposed systems make use of motion detection based approaches and are therefore inappropriate for these scenes. This is despite the fact that slow moving or stationary traffic is the main problem facing traffic management authorities in most towns around world.
The proposed algorithm detected vehicles with a good degree of accuracy under different illumination conditions during the day for both free flowing and stationary traffic scenes. The shadows were also well handled with a good degree of success as shown in Fig. 4 and Fig. 6. The vehicle detection algorithm used a novel approach in which the vehicles were simultaneously extracted from the traffic video data frames and their negatives using the Laplacian of a Gaussian edge detector. Edge linking was achieved through mathematical morphology and summation of the positive and negative edge maps. However, despite the success of the algorithm, it was noted that over-segmentation occurred for very large trucks: with cabins and their trailers being detected as separate vehicles. The algorithm also had problems with occluded vehicles. To minimize these problems, it is suggested to raise the position of the camera to be high enough with respect to the ROI.

ACKNOWLEDGEMENTS

This paper borrows its material entirely from a thesis named 'Vision Based Automatic Road Traffic Density Estimation and Vehicle Classification for both Stationary and free flowing Traffic Scenes Using an Ensemble Pattern Classifier' and a Journal paper named 'Vision based road traffic density estimation and vehicle classification for stationary and moving traffic scenes during daytime'. Both the thesis and the Journal paper were authored by the corresponding author* of this paper under the supervision of the other two authors. The Journal paper was published in the Journal of Sustainable Research in Engineering (JSRE).

REFERENCES

[1] S. Sivaraman and M. M. Trivedi, "A Review of recent developments in vision-based vehicle detection," in 2013 IEEE Intelligent Vehicles Symposium (IV), Gold Coast, Australia, 2013.

[2] K. Kim, T. H. Chalidabhongse, D. Harwood and L. Davis, "Real-time foreground–background segmentation using codebook model," Elsevier, 2005.

[3] C. Stauffer and W. E. Grimson, "Adaptive background mixture models for real-time tracking," in IEEE International Conference on Computer Vision and Pattern Recognition, 1999.

[4] V. Heidari and M. R. Ahmadzadeh, "A method for vehicle classification and resolving vehicle occlusion in traffic images," Pattern Recognition and Image Analysis (PRIA), 2013 First Iranian Conference on, pp. 1-6, 6-8 March 2013.

[5] C. C. C. Pang, W. W. L. Lam and N. H. C. Yung, "A Novel method for resolving vehicle occlusion in a monocular traffic-image sequence," IEEE Transactions on Intelligent Transportation Systems, vol. 5, no. 3, pp. 129-141, 2004.

[6] A. Sharma, D. Bullock, S. Peeta and J. Krogmeier, "Detection of inclement weather conditions at a signalized intersection using a video image processing algorithm," The Civil Engineering at DigitalCommons@University of Nebraska -

Lincoln, vol. 32, pp. 150-155, 24-27 September 2006.

[7] Minnesota Department of Transportation, Office of Traffic, Safety & Technology and SRF Consulting Group, Inc., "Evaluation of non-intrusive technologies for traffic detection," United States Department of Transportation, Federal Highway Administration, Minneapolis, 2010.

[8] S. H. Yu, J. W. Hsieh, Y. S. Chen and W. F. Hu, "An automatic traffic surveillance system for vehicle tracking and classification," in Scandinavian Conference on Image Analysis (SCIA2003), Göteborg, Sweden, 2003.

[9] R. P. Avery, Y. Wang and G. S. Rutherford, "Length-based vehicle classification using images from uncalibrated video cameras," University of Washington, 22 July 2004. [Online]. Available: http: //faculty.washington.edu/yinhai/wangpublication_files/ITSC_05_VC.pdf. [Accessed 21 March 2015].

[10] Open Network Video Interface Forum Inc., "ONVIF™ core specification," August 2013. [Online]. Available: http://www.onvif.org/specs/core/ONVIF-Core-Specification-v240.pdf. [Accessed 20 June 2015].

[11] A. Ambardekar, M. Nicolescu, G. Bebis and M. Nicolescu, "Vehicle classification framework: a comparative study," EURASIP Journal on Image and Video Processing, 2014.

[12] M. Pancharatnam and D. U. J. Sonnadara, "Vehicle counting and classification from a traffic scene," in Proceedings of the 26th National IT Conference, Colombo, Sri Lanka.

[13] G. Zhang, R. P. Avery and Y. Wang, "A Video-based vehicle detection and classification system for real-time traffic data collection using uncalibrated video cameras," Department of Civil and Environmental Engineering, University of Washington, Washington, 2006.

[14] N. Buch, J. Orwell and S. A. Velastin, "Detection and classification of vehicles for urban traffic scenes," in Visual Information Engineering, 2008. VIE 2008. 5th International Conference, Xian China, 2008.

[15] E. Ince, "Measuring traffic flow and classifying vehicle types: a surveillance video based approach," Turkish Journal of Electrical Engineering and Computer Science, vol. 19, no. 4, 2011.

[16] R. C. Gonzalez and R. E. Woods, Digital image processing, Third Edition, Upper Saddle River, New Jersey: Pearson Education Inc., 2008.