Estimated Vehicle Density Based on Video Processing Using the Gaussian Mixture Model Method

A A Putri\textsuperscript{1}, A Achmad\textsuperscript{2} and Suwadi\textsuperscript{2}

\textsuperscript{1} Institute of Technology Sepuluh Nopember, Telematic Engineering, Surabaya, Indonesia
\textsuperscript{2} Institute of Technology Sepuluh Nopember, Telecommunication Engineering, Surabaya, Indonesia

Abstract. Congestion is a problem that cannot be decomposed properly, especially at big cities in Indonesia, one of them is in Surabaya. To solve this problem is needed to observe characteristics of the road such as the number of vehicles, vehicle speed and vehicle density. This research took the data in several areas in Surabaya by put the camera in the middle of a crossing bridge. This research aims to determine the density value by comparing traffic volume with vehicle speed. The value of speed and number of vehicles can be known using Gaussian Mixture Model method, where object detection is detected by background substraction process, which is video divided into several frames and in each frame images is divided into two part there are foreground and background. The object will be count 1 when blob analysis detect the foreground. In the theory to know the density value of a traffic flow, with compare the number of vehicles and the speed of vehicle. From the data has been taken, it is known that the number of vehicles is opposite with average speed the vehicle that passes. If average velocity of vehicle greater than the number of vehicle so the value density vehicle getting smaller. so that information can be found that the road is experiencing congestion conditions.

1. Introduction

Big Cities in Indonesia especially Surabaya has a critical cases in the daily life and one of them is congestion. Congestion that happen can be caused by many factors including the traffic volume and the lack awareness of society obeying the traffic rules, even alot of public transportation stops out of place, pedestrians do not cross in place or even motorists do not follow traffic lights. However, this condition in Surabaya will be worse during rush hour.

The main roads including Basuki Rahmar, A Yani, and Bungurasih have a great intensity movement so that the information can be known from the vehicles volume, vehicles speed, and the vehicles time that have reached, those are when vehicles volume increased causing the vehicle speed decreased, then the time to travel the road is getting bigger. One solution to this problem is to expand the capacity of the road, but if this is done it will require considerable costs so that what can be done is by examining the parameters, this can be obtained before the parameters include vehicle speed, traffic volume and road capacity.

The data retrieval in the form of traffics video will be processed using the Gaussian mixture model, where the video will be divided into several frames, in each frame will produce an image and then the process of background substraction will divide into two parts, namely background and foreground, through blob analysis and tracking object, value 1 will be counted when the foreground is detected. From these results can be seen the value of vehicle speed and number of vehicles, where the indicator of traffic performance is vehicle density (k) is the ratio between vehicle volume (q) and vehicle speed (v), theoretically the magnitude should not be more than 1, which means if the value approaches 1, the road condition is near saturation, in another word in the traffic field it is in a solid state. If the condition occurs repeatedly while the capacity is inadequate then a condition will occur called congestion. This condition can be seen by looking at the average speed, the lower the average speed, the greater the level of congestion in the road. In this case according to the Indonesian Road Capacity Manual (1997) the maximum limit of the ratio between volume and capacity is 0.75.)
2. Previous Work On Vehicle Density Assessment

In the previous research, there were several results published in an effort to develop vehicle density calculations. The vehicle detection process on the road is used for tracing vehicles, to calculate the average speed of the process through the image processing video is divided into perframe and knows the frame rate of the video. Background subtraction meaning is taking the moving video background. The system is by recording the video on the highway, after taking perframe, then the background will be taken. Vehicle detection from video was taken from analyzing the video background by dividing image into background and foreground from BLOB analysis. In a way, if the bbox shows the foreground detecting the vehicles process. Vehicle tracking system, the extracted feature has been tracked over sequential frame retrieval from input traffic video data. System represents an between scheme feature matching scheme in the previous frame, which has been stored in the track graph metric and instaneous frame. The speed estimation scheme formula is Speed = (Distance) / (Total frame * frame rate). [1] The relations of the density and the speed of vehicles can be measured by using several moving observation methods including Greenshields model, Greenbergs model and underwoods model. The value was taken when the vehicle moves on the highway within the opposite directions, Greenshield’s result believes that the flow is not disturbed, the speed and density are linearly related. Greenberg thinks that the relation between speed and density is not predictable when the density is low, because the density is close to zero. However, Speed tends to rise infinitely. Underwood thinks that the relation of density and speed is exponent. The main disadvantage is the speed to zero only because the density value is infinite. After knowing each of the value, then look for the regression values using statistic calculations, so the result of the coefficients from the density of zero point five to zero point six can be obtained from the three models [2]. The density calculation is not only for counting the number of vehicles but calculating the number of road map complexity and maximum capacity so that it proposes to calculate SJRatio and total distance from the map to predict the relation of them [3]. Along with vehicle speed and high mobility constrained by road conditions from uneven density distribution, in the analysis of vehicle correlation and density, variations were found in the case of high vehicle density intersections with worse congestion [4].

3. Method

Input data is needed from offline digital video which record a moving vehicle taken from a crossing bridge. The input parameter is a parameter of several methods from a process that is entered into the system by the user. The input parameters is needed in this system are the parameters of Region of Interest (ROI), filter median parameters, object detection parameters.

After input data and all parameters are obtained, the system will extract video into a some of frames. Each frame that has been extracted will find the background using the Gaussian Mixture Model method. Background is an image of immovable objects, such as highways, sidewalks, and houses. The background was obtained will be compared to frame so that it will produce a new image, there is a foreground. Foreground is an image of a moving object. Foreground is a grayscale image with 3 gray levels, each of which indicates a stationary object, a moving object, and a moving object's shadow.

The results of the foreground image usually have some noise. Noise usually occurs due to light reflection, falling leaves, moving object shadows, or other undesirable disturbances found in the image. To reduce or even eliminate noise, a smoothing process is needed, which is a process to refine the image. The smoothing method used in this study is the median filter. The way the median filter works is to replace the value of each pixel with the median value of its neighboring pixels. After the image becomes smooth or noise is reduced, a shadow removal process is needed to remove shadows, which is stating the shadow of a moving object as an immovable object so that the result is a binary image or has two values, each of which states a moving and immovable object.

After getting a binary image, the system will detect the object. Object detection is done by checking pixels one by one in the Region of Interest. If the pixel that is examined is a pixel from moving object,
so other neighboring pixels that are same object will be searched. Each object that is detected will be counted the number of pixels. After the object is detected, the blob analyst does it, where each object said to be moving will be analyzed as a vehicle so that the automatic detection process of the number of vehicles can be carried out based on the number of objects detected in the blob analysis.

After the number of vehicles is known the next step is to calculate the speed of the vehicle in the road segment. Vehicle speed calculation starts from installing the region of interest manually, where the function itself is as the selection of vehicles to be detected. The next step is to enter some parameters including the horizontal distance between the cameras, the height of the camera installation, the viewing angle covered by the camera, the angle of view that is not caught on camera, and the video frame rate.

![Flowchart](image)

**Figure 1.** Flowchart program

### 3.1 Input Video

Data enter this system in the form of offline video, recording of vehicles on the road using the iPhone 6s camera. Videos are taken with the position of the camera lens reaching all areas of the one-lane road. In video recording, the parameters to measured including:
the angle of installation camera, the height of the camera and the horizontal distance between the camera to starting point covered.

Where:

- $H$: Camera installation angle
- $d_1$: Area angle that is not visible
- $d_2$: Angle covered by camera
- $D$: Camera height from road surface
- $d_1$: Horizontal distance between camera and vehicle
- $d_2$: Areas not visible to the camera
- $d_3$: Area visible to the camera

Area can get:

$$d_1 : H \tan \left( \theta \right) = H \tan \left( \frac{\theta}{2} \right)$$  \hspace{1cm} (1)

$$d_2 : D - d_1$$  \hspace{1cm} (2)

**Figure 3.** Angle camera to know vehicle in street

in taking video on the highway it takes a wide range of power to get the optimal angle value

**Figure 4.** Camera tendencies and grid view results for vertical and oblique aerial image

### 3.2 Selection of ROI Area

At this step the ROI (Region of Interest) will be used as an area for detecting mobile vehicles.
3.3 Background Subtraction With Gaussian Mixture Model

Gaussian Mixture Model (GMM) is a method requires a clustering algorithm to group pixels which divided into two types, foreground and background. Then the system has been implemented divides the frame into a background image and foreground image using a background algorithm an image that becomes the background of a moving object. While foreground is a picture of a detected moving object. Foreground image obtained is filtered to leave some noise in the image. So that the image will be easier to process and the object of the vehicle will appear more clearly. This noise is a captured image more clearly.

3.4 Object Detection Process with Blob Analysis And Blob Tracking

Detect objects based on blob analysis using the connected component method after filtering process, where the analysis is carried out when each set of pixels with gray level is one, categorized as one object. Detected objects will be labeled to simplify the process. From the blob analysis, will be obtained information about centroid, area, height, and width of an object obtained from a rectangle. After that the object tracking process is carried out with blob tracking. Blob tracking aims to determine that objects detected at frame e-I with i-l frame are the same or different objects. This can be known by measuring the distance from the centroid of each blob detected. If the distance is less than the specified threshold, the object detected in these frames is the same object.

3.5 Moving Speed Detection Process

Speed will be detected after the tracking object is known. Each object from the tracking blob will calculate the centroid value on the first frame and the i-l frame. the first is to do the process of finding the distance between pixels is done by finding the coordinates of the point and calculating the distance between the center points. This can be obtained from blob analysis.

Simply to find the center point is

1. Calculate the x coordinates, with:

   \[ \text{Posisi } x = \text{bbox}(1) + \left( \frac{\text{bbox}(3)}{2} \right) \]  

2. Calculate the y coordinates, with:

   \[ \text{Posisi } y = \text{bbox}(2) + \left( \frac{\text{bbox}(4)}{2} \right) \]  

Where:
4. RESULT

A. Detect objects using the Gaussian Mixture Model method

In the first step, input is matched with all distributions until the most suitable distribution is found. A pixel can be said to enter in a distribution if the pixel value is within a distance of 2.5 standard deviations from a distribution. For matching inputs used inequality.

\[ \mu_k - 2.5 \times \sigma_k \leq X_{t} \leq \mu_k + 2.5 \times \sigma_k \]

where \( X_{t} \) is the value of the intensity pixel \((i, j)\) on the \(t\)-frame, \( \mu_k \) is mean on pixel \((i, j)\) from \(k\)-Gaussian \( \sigma_k \) as a standard deviation in pixels \((i, j)\) from \(k\)-Gaussian.

In the second step is parameter update. At this step an update the values of GMM parameters is will be used to process the next input. The updated value consists of weight, means, and standard deviation. The weight value is updated using equation 10. The means value is updated using equation 11. The standard deviation value is updated using equation 11. Each of these equations applies to each pixel \((i, j)\).

\[ \omega_{k,t} = (1 - \alpha) \omega_{k,t} - \alpha (M_{k,t}) \]

with \( \omega_{k,t} \) is weight from Gaussian \(k\)th from time \(t\), \( \alpha \) is learning rate and the value is one. for suitable models and 0 for other models. After the weight date value is normalized the total weight of the total distribution is not more than 1. The standard deviation value of the Distribution signal is updated every time there is a pixel value that matches the distribution.

\[ \sigma^2 = (1 - \rho) \sigma_{k,t}^2 - \rho (X_{t} - \mu_{k,t}) (X_{t} - \mu_{k,t}) \]

where \( \omega_{k,t} \) is the value of Gaussian-\(k\) at frame \(t\), \( \mu_{k,t} \) is the mean from Gaussian at \(k\)-th at frame-\(t\), \( \sigma_{k,t} \) is standard deviation from Gaussian-\(k\) at frame \(t\), \( \alpha \) is learning rate, and the value of \(M_{k,t} \) is 1 for the suitable model and 0 for the model that doesn’t match. After the weight value is updated, do normalization so that the total weight of all distributions is correct.

In the second step is Selection of Background Distribution. At this step, models are chosen that reflect background. First of all models are sorted by \( \sigma^2 \) so the distribution that most reflects the background remains above and that does not reflect background below which will be replaced by another distribution. To select \(B\) the first distribution used as a background distribution is used equation 12, where \(T\) is a predetermined threshold value.

\[ B = \arg\min_k \sum_{t=1}^{b} W_k \]

In the third step is filtering by removing the shadow and reducing noise, how it works with replace intensity value of each pixel with value of kmedian from neighboring pixels. The range pattern of neighboring pixels is called a predetermined window. The median value is calculated by sorting in the pixel window section itself. The output of smoothing is a frame with the intensity obtained from the median value. The process is illustrated in Figure 6.

![Figure 6. Smoothing process with median filter.](image-url)
In the fourth step, shadow removal is a process where the shadow detected in foreground picture that will be removed. In shadow removal the output of GMM is grayscale image which only has 3 types of intensity there are 0 for fix objects, 127 for shadow objects, and 255 for moving objects, after that will be changed to 2 types of intensity or commonly called binary, 0 for fix objects and shadow, and 1 for moving objects.

Figure 7 Result of detection object

A. Count Number of Vehicle

In calculating the number of vehicles using the Gaussian mixture model. The first step in the GMM method is the initialization of the normal distribution from the background. This step aims to produce an initial background model. At this step the first frame is considered as the background. The intensity of each pixel in the first frame is used as the initial average value of the background model. While the threshold is constant, which is 30 and for the initial variant is 1 for each pixel.

Comparison of two background frames The most underlying background modeling is to ensure that minimal background pixel changes are still detected as background not as objects. Reducing and matching the background has an important role, to know the background part that exists in an image. The basic idea of background subtraction is | framet-framet1 | <Threshold * Variant This comparison is done by pixels per pixel. As seen in figure 7. If pixel i satisfies the equation, so the pixel is the background, another that not the background. Comparison of these two backgrounds is done to take further action if it is detected as an object or background.

Update the background pixel value. The stage of updating pixel background values aims to adjust the background to the change that occur. This update is applied to GMM parameters which will be used to process the next input. The parameters updated in this study are the mean and variant. After that blob analysis automatic detection, the number of vehicle.

Figure 8. Result of detection object

Figure 9. Result of count vehicle
B. Count Speed Vehicle

After blob analysis on object detection process, so the pixel coordinates of the objects detected in each frame will be known. The value of the Blob analysis is obtained from equations (3) and (4). Then this coordinate value is used to calculate the distance between one frame and one previous frame in pixel.

In giving a bounding, a white box object will be calculated by using the Gaussian mixture model method. So that the object is determined when a white image is detected, this is due to noise reduction. The use of Region of interest to limit the frame area for object detection, and as a determinant of starting and ending points of speed calculation, the condition of the vehicle start and speed detection. The coordinate midpoint of the object are obtained in each frame that is used to find the value of the distance displacement between frames in pixels. The trick is to find the difference in resultant from center point of detection object. However, to detect more than one object, is needed Euclidean measurements by looking for the difference between the resultant distance of the nearest pixel between the coordinates of the current frame and the previous coordinate with eq (5). After that, find the distance value using the formula

\[ \sum_{i} d_{i} \]  

(13)

Where
- \( d_{i} \) = result from eq (6)
- \( I \) = the initial value of the number of pixels
- \( n \) = the final value of the number of pixels

The input needed for the speed detection process are:
- a. Horizontal distance between the camera and the road used when first entering (D)
- b. The mounting height of the camera used for recording (H)
- c. The angle of view covered by the camera (teth(1))
- d. Regional viewpoints that are not captured by cameras(teth(2))
- e. Framerate video(frame)
- f. Size video in pixel (pixel)

![Figure10. Parameter for calculate speed detection](image)

After inputting the parameters in the program, speed obtained from a comparison of distance and time. The distance in program is the actual distance. Where the actual distance value is obtained by multiplying the number of distances in pixels multiplied by skala at result program. The value of visual distance based on pixels is obtained using eq (6) and sum of them by eq (13). and the skala value can be obtained from the comparison between eq (1) and eq (1).
C. Relationship of density, speed and number of vehicles.

Based on the equation $q = k \times v$, the results show that the density of the vehicle is directly proportional to the number of vehicles and inversely proportional to the speed of the vehicle. Evidenced by several experiments taken.

| No. | Jam            | V (smp) | S (km) | D = V/S (smp/km) |
|-----|----------------|---------|--------|------------------|
| 1   | Movie1.*avi    | 18      | 0.24   | 75               |
| 2   | Movie2.*avi    | 22      | 0.24   | 91.66666667      |
| 3   | Movie3.*avi    | 20      | 0.24   | 83.33333333      |
| 4   | Movie4.*avi    | 17      | 0.24   | 70.83333333      |
| 5   | Movie5.*avi    | 21      | 0.24   | 87.5             |
| 6   | Movie6.*avi    | 20      | 0.24   | 83.33333333      |
| 7   | Movie7.*avi    | 18      | 0.24   | 75               |
| 8   | Movie8.*avi    | 25      | 0.24   | 104.1666667      |
| 9   | Movie9.*avi    | 19      | 0.24   | 79.16666667      |
| 10  | Movie10.*avi   | 19      | 0.24   | 79.16666667      |

Figure 11. Result of speed detection

Figure 12. Linier Regresion
So Base on that equation density can estimate

\[
\frac{S_{ff}}{D_j} = 1.0136
\]

\[
D_j = \frac{42,413}{1,0136} = 40,938
\]

5. Conclusion

Vehicle density can be known based on several parameters including the speed and number of vehicles obtained from the equation \( q = k \cdot v \). Density is directly proportional to the number of vehicles passing and inversely proportional to vehicle speed. So if the value of the density greater, then speed of vehicle have smaller value and number of vehicle is getting bigger and the value of density can estimate that 40,938 km/jam.

6. References

[1] Harsha, S., Sri and Anne, K.R. A Highly Robust Vehicle Detection, Tracking and Speed Measurement Model for Intelligent Transport System, pp. 3733-3742, 2016
[2] Abhijeet, Singh, Abhijeet, Kumar and R.H, Goudar, Online Traffic Density Estimation and Vehicle Classification Management System, Vol 7(4), 508-516, 2014
[3] Sanguesa, Julio A, Naranjo Fernando, Torres-Sanz Vicente, Fogue Manuel, Garrido Piedad, J.Martinez Francisco, On the study of Vehicle Density in Intelligent Transportation System, 2016
[4] M.Reza, Pooshideh, Javad, Haddadnia, Tahere, Royani, A New Prediction Algorithm: Flexible Switch for Combining LMS and RBF Adaptive Filters (CLMF), ISSN: 1792-4618, 2010
[5] Saurav Barua, Anik, Das, Hossain Julfiker, Estimation Of Traffic Density to Compare Speed Density Models with Moving Observer Data, International Journal of Research in Engineering and Technology, 2015
[6] Celil, Azkurt and Fatih, Camci, Automatic Traffic Density Estimation and Vehicle Classification for Traffic Surveillance Systems Using Neural Networks, Vol 14 No 3 pp187-196, 2009
[7] Widodo, Wahyu, Wicaksono, Nur and Harwin, “Analysis of The Volume, Speed, and Density of Traffic with Greenshields and Greenberg Methods” Vol 15 No 2, 178-184, 2012
[8] Mukhlis, Relation of Speed, Density and Volume at Traffic with Greenshields, Vol 3 No 2, 2013
[9] Samhitha, Harsha Pulli, Jyothi, Allu Naga, Vesapogu, Ramana, Mannem, Manasa, Harsha, ill.