Optimal Moving Windows for Real-Time Road Image Processing

Sung Yug Choi and Jang Myung Lee
Intelligent Robot Laboratory, Pusan National University
http://robotics.ee.pusan.ac.kr

Abstract

In this paper, a "Moving Window" scheme for detecting lanes and obstacles from the images captured by a CCD camera in an automobile is proposed. Processing the input dynamic images in real time requires high performance hardware as well as efficient software. In order to relieve these requirements for detecting lanes and obstacles from the images in real time, a "Moving Window" scheme is proposed. The size of the moving window is optimally determined based upon road and automobile conditions to be big enough to detect the lanes and obstacles and to be small enough to process the detection procedure in real time. For each image frame, the "Moving Window" is newly defined and it is moved in a certain direction that is predicted by the Kalman filtering technique. By detecting the left and right lane marks of the driving lane, it becomes possible to search the obstacles within the lane. The obstacle can be verified through the correlation between the stochastic characteristics of the suspected obstacle and the actual obstacle in the database. The feasibility of the proposed algorithm is demonstrated through the simulated experiments of freeway driving.

Keyword : Optimal moving window, Lane detection, Obstacle detection, Kalman filter.

1 Introduction

Recently, several advanced countries are promoting the construction of Intelligent Traffic System(ITS), the aim of which is to provide traffic service forming a part of national welfare improvement[8]. Hereupon, Korea has established ITS Korea and has been putting spurs to the construction of ITS. Advanced Vehicle System(AVS), which is a part of ITS, is to maximize safety through the improvement of facility and manipulation behind the wheel and ultimately to introduce autonomous driving. It is working under different names, for example, AVCS(Advanced Vehicle Control System) in America, Prometheus in Europe, and ASV(Advanced Safety Vehicle) in Japan[1],[2].

Current autonomous driving system has the basic conception that in driving situation, a driver, a vehicle, and road environment are not individual elements but one element combined organically[7].

The detection of lanes and obstacles using image information suggested in this study belongs to the former active system, and therefore can assist with the safety of driving through the transmission of information on a given situation. Various sensors, such as ultrasonic, radar, laser etc. have been used together to provide a driver with information on safety[3],[5],[7].

This paper consists of six sections including the instruction, and proposes the optimal size of window in section II, and discusses detection of lane by moving windows and recognition of obstacles in section III and IV respectively. And then it describes an experiment for detection of lane in section V and deal with conclusion in section VI.

2 Decision of Window Size

Consecutive input images have the size of 640×480 pixel at every sampling time. It can cause the increase of unnecessary computation to apply preprocessing to total image in order to draw some necessary information. Therefore, the research on the decrease of computation time by designating a necessary part as a sub-block is being in progress[5]. This section explains edge detection based on image processing and central moment that is important when determining position of windows. And then it decides the optimal size of windows.
Optimal Window

It is required to minimize the size of windows for image processing in real time. Since the information of lane must be found from windows. The size of a window will be determined according to the condition that at least some part of a lane must be included inside one sub-block, i.e.,

Size of a window: The location of a lane in an input image can be represented by a curvature radius $r$ and the gradient. When a vehicle is moving at the maximum allowable speed $v_{max} = 100$ [km/h] along the road which has the curvature radius $r$ and gradient $0 \leq \text{gradient} \leq 1.718$ [°] that allow maximum safety, the minimum size that does not lose the location of a lane can be designated as a sub-block.

The necessary number, and the initial location of windows: According to the proposed conditions under which the size of a sub-block is determined, the size of subsequent sub-blocks distantly located from a moving vehicle will be determined. Besides, considering the actual distance $d$ [m] corresponding to the size of a sub-block, the number of subsequent sub-blocks needs to be minimized so that a vehicle moving at the maximum admissible speed $v_{max} = 100$ [km/h] can detect the lanes and the obstacles that are located within the minimum safety distance $100$ [m] in front. For at least some part of a lane to be included inside each sub-block, the initial location of a sub-block needs to be determined on the premise that a vehicle should start on a rectilinear road.

Determine of initial position

Any point in real world corresponds to a point in the image obtained through a CCD camera as shown in Fig. 1.

![Fig. 1. Coordinate systems for a point on the image and on the real road.](image)

When the position vector of the point in real world from the camera is defined $P_c = \begin{bmatrix} p_{x,1} & p_{y,1} & p_{z,1} \end{bmatrix}$, the equation can be denoted as (1),

$$P_c = sR_g P_o + P_g \tag{1}$$

where $P_o$ is the position vector to the object \([O]\) w.r.t the road coordinate system \([G]\), and $P_g$ is the position vector to the road coordinate system \([G]\) w.r.t the camera coordinate system \([C]\). Assuming that the camera is installed horizontally to the road, and directed to the front from the center of the road, we have $R_g = R(x,90^\circ)$, and $P_g = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$. The perspective model of the camera in the relationship between the input image and the pixel coordinate system of an object is shown in Fig. 2.

![Fig. 2. Camera projection model.](image)

where $f$ means a focal length of the camera. The transformation into the pixel-coordinate of the image frame yields (2).

$$s_i = \frac{p_{xi}}{p_{xi}\delta x} + c_x \quad y_i = \frac{p_{yi}}{p_{yi}\delta y} + c_y \tag{2}$$

where $\delta x$ and $\delta y$ represents the horizontal and vertical length of a CCD element respectively and $(c_x, c_y)$ represents the coordinate of the central point in the CCD image frame.

For the calculation of hidden area, assuming that the optical axis of a camera is installed horizontally to the road, the actual length between the two bottom corners which do not appear in the input image can be computed as.

$$s_x = \frac{a_0 \cdot \delta x}{2f} = \frac{a_0 \cdot \delta x}{h_y} d [\text{m}] \tag{3}$$

$$y_y = \frac{2f \cdot c_y}{h_y} = \frac{2f}{h_y} d [\text{m}] \tag{4}$$

where $a_0$ and $h_y$ represents the number of the horizontal and vertical pixels in the input image respectively; $P_g$ represents the height of the camera from the bottom. Fig. 3 shows the modelling of a road into a circular arc to consider the shape of the curvature of a driving road. In case of a freeway, the minimum radius of rotation is limited to 690 [m], and if the radius $r$ becomes indefinite, it means that the road becomes rectilinear.

![Fig. 3. Cross-section modeling of the road.](image)

In Fig. 3, $r_o$ is the distance between the center of the camera and either lane of the road. Thus, it is $r = r_o$ for the left lane and it is $r + r_o$ for the right lane. Putting the center of the arc as a base, $\theta_{r_L}$ and $\theta_{r_R}$
each represent, the arc angle between the starting point of the bottom left lane and the bottom right lane in the input image and the location of the camera. For the right lane, the following two equations can be derived.

\[
\theta_{th} = \sin^{-1} \frac{y_h}{r+e_r} \tag{5}
\]

\[
l_h = (r + e_r) \theta_{th} = (r + e_r) \sin^{-1} \frac{y_h}{r+e_r} \tag{6}
\]

In Fig. 3, the gray area is the hidden layer that does not appear as a camera image; \(y_h\) is the distance for the area and it can be denoted as (4). Using this, \(\theta_{gr}\), which is the arc angle of the \(k\)th right lane, can be denoted as follow:

\[
\theta_{kr} = \frac{y_h + l_{gr} - k}{r} \tag{7}
\]

\[
= r^{-1}\left\{ (r + e_r) \sin^{-1} \left( \frac{y_h}{r+e_r} \right) + l_{gr} - k \right\}.
\]

In addition to the above, the position vector of a lane from the road coordinate system \(\{G\}\) can be computed as.

\[
P_{xk} = (r + e_r) \cos \theta_k - r
\]

\[
P_{yk} = (r + e_r) \sin \theta_k
\]

\[
P_{zk} = 0
\]

Substituting (8) into (1) and then into (2), the coordinate of the lane on the screen can be determined as.

\[
x_{ak} = \frac{(r + e_r) \cos \theta_{kr} - r}{(r + e_r) \sin \theta_{kr}} \frac{f}{dx} + c_x
\]

\[
y_{ak} = \frac{d}{(r + e_r) \sin \theta_{kr}} \frac{f}{dy} + c_y
\]

Based on the above, the starting point of the \(k\)th lane appearing in the input image can be also determined as,

\[
W_{ak}(r) = f_{xk} y_h \int_{0}^{M} W_{ak}(r) \leq M \tag{11}
\]

where \(M\) is the maximum size of the input image and it is 640×480 in this paper.

**Determination of window size**

In (9) and (10), the position of a lane is the function of the radius of rotation \(r\). When \(r\) becomes infinite from the minimum radius of rotation 11690 [m] on a freeway in case of a rectilinear road, the maximum range of information on a lane being acquired inside a block is defined as the size of a sub-block. It is shown in (12).

\[
W_{ak} = \max\left\{W_{ak} = W_{ak}(r) - W_{ak}(r) \mid 690 < r < \infty, k = 1,2 \ldots n\right\} \tag{12}
\]

The slope a road may influence the vertical length of a sub-block. Since the slope is very smooth in the freeway, we ignore the effect in this paper.

3 Definition of Moving Window

This section deals with the definition of moving window and explains of lane detection using the information from the driving vehicle.

3.1 Determination of check-point and detection of lane

After labeling by preprocessing the area that is marked as edge inside a moving window, a central moment is obtained. The position of a lane in the input image with the sampling time 95 [ms], does not change drastically, compared to the size of a moving window. Based on position of central moment, a check point is determined, which will the center of a moving window.

**Decision of check point**

The central moment for the lane area shown in Fig. 4 can be marked as Fig. 5.

![Fig. 4. Input image. Fig. 5. Central moment.](image)

The detected central moment can be moved to a certain position along the Y-axis as shown in Fig. 6 with the aid of the curve fitting method. The moved central moment is renamed as "Check point" in this paper.

![Fig. 6. Decision of check point.](image)

**Detection of lane**

The lane areas can be expressed in \(N\) th order polynomial as (13),

\[
g(y) = a_0 + a_1 y + a_2 y^2 + \cdots + a_N y^N . \tag{13}
\]

the coefficient \(a_n\) can be found in the equation (14) derived from the partial derivative which minimizes the summation of the square deviation[9].

\[
\sum_{i=0}^{N} \sum_{i=0}^{L} a_n + \sum_{i=0}^{L} x_n y_i \quad \text{for } k = 0,1,2,\ldots,N \tag{14}
\]

3.2 Expectation of moving window position

This subsection proposes the decision of searching area at the time \((k+1)\). Searching area is called as "Moving window" in this paper.

![Fig. 7. Predicted value, \(x_k\)](image)

The check point located in the center position along the Y-axis can move to a certain position along the X-axis denoted as \(x_k\) in Fig. 7. A fixed size of window is newly determined around the check point, that is defined as a moving window[10].

The state \(x_k\) is estimated as following equation (15).
\[ x_k^* = F x_{k-1}^* + \omega_{k-1} \]  

where, \( x_{k-1}^* \) means a measurement value in instance \((k-1)\) and \( \omega_{k-1} \) presents the Gaussian noise from the system model. And its covariance matrix is denoted as \( Q_{k-1} \). A system matrix, \( F \), is defined as:

\[ F = \begin{pmatrix} 1 & \Delta_i \\ 0 & 1 \end{pmatrix} \]  

where, \( \Delta_i \) represents the measurement time of the check point. Besides, in the measurement step, the measurement vector \( y_k \) is denoted as \((17)\) and \( v_{k-1} \) is the measurement error caused by the Gaussian noise with zero mean.

\[ y_k = H x_{k-1}^* + v_{k-1} \]  

Note that \( v_{k-1} \) is uncorrelated with \( \omega_{k-1} \) in the estimation step and its covariance matrix is \( R_{k-1} \) in the measurement step. Since the image itself does not show the change of distance per image frame, the measurement matrix can be represented as \((18)\).

\[ H = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \]  

The recursive Kalman filter to compensate the state vector by adding measurement deviation to the estimated can be described as follow:

\[ x_k^* = x_k^t + G_k (y_k - H x_k^t) \]  

4 Obstacle Detection Using Moving windows

4.1 Distinction between obstacle and road marks

When the position of a right line and a left line is denoted as \( F_R(y) \) and \( F_L(y) \) respectively the middle line can be calculated as \((20)\).

\[ F_M(y) = F_R(y) - F_L(y) \]  

The intensity of a middle line is \( f_M(x, y) \) within 0 to 256 representing gray level. An obstacle ahead of the vehicle (the obstacle here is actually another vehicle) has the characteristics of horizontal constituents such as a rear bumper and a trunk. So if an edge gradient is detected in the horizontal direction, the expected obstacle can be characterized. Using \( f_{MB}(x, y) \) that is the binary horizontal edge, the estimated area of the obstacle, that is \( f_{OB}(x, y) \), can be extracted.

![Fig. 8. Expected area for obstacle.](image8.png)

![Fig. 9. Input image of a hexagon.](image9.png)

Obstacle and Road-mark

The objects expected from the change of intensity ahead of a moving vehicle can be largely divided into two: a three-dimensional object and two-dimensional road marks. When represented as the coordinate systems between a camera and actual object, they can be distinguished through the vertical and horizontal lines of the hexahedron like in Fig. 9.

To analyze the image inside the expected obstacle area \((m \times m)\), the values in gray level in the area is normalized and smoothed to exclude the noise constituent of the input image as follows respectively:

\[ S_i = \sum_{j=1}^{N} f_\omega(x_i, y_j) \]  

\[ Histo = \sup_{i} ||R_{x_j}|| i = 1, ..., n, R_{x_j} = 50 \]  

\[ Smooth = \frac{1}{5} \sum_{i=1}^{5} (Histo_i + Histo_{i+1} + \cdots + Histo_{i+4}) \]  

where \( R_{x_j} \) is required for normalization.

4.2 Recognition of an obstacle

A neural network has frequently been used in the previous studies on obstacle detection through an input image\[4\].

In this study, a neural network is introduced as a pattern classifier to make a distinction by the proposed method between vehicles and road marks. The structure is forward multi-layer neural network as shown in Fig. 10 and it consists of 35 input layer, 16 hidden layers, and 6 output layers. The sigmoid activation function is used for each node.

![Fig. 10. Structure of a neural network.](image10.png)

![Fig. 11. Histogram by normalization for neural network](image11.png)

![Fig. 12. Sub-collection](image12.png)

When the values of the histogram normalized in equation \((22)\) are represented like Fig. 11, the input of a neural network can be used, divided into the subsets of a certain area as in Fig. 12.

\[ (Input)_i = \sum_{j=1}^{N} x_j \]  

When the pixel of 98×50 is divided into 35 subsets of 7×5, one subset consists of 14 pixels in width and 10 pixels in length. The input value of the \( i \)th neuron in the input layer of a neural network is determined by the equation \((23)\) through the normalization of the pixel values which belong to the \( i \)th subset.

\[ (Input_i) = \sum_{j=1}^{N} x_j \]  

where \( N \) is 1,249,500(=4,900×255). The numerator is the total of the pixels that belong to the subsets, so it reflects the attributes of the pixels. Since
the sigmoid function is made up of exponential functions, the normalization is made with the value between 0 and 1.

In Fig. 13, the output of the neural network consists of six neurons, the horizontal axis meaning the number of the output layer neuron and the vertical axis meaning the value of the output layer.

The reason for using the decrease by the exponential function, instead of putting two output layers, is to give some flexibility in case of the input of images which are not learned. The following is the output of the neural network for an input image by the suggested method.

5 Experiments

In this section, the environment for the experiment will be described and the excellence of the moving window technique in the aspect of computation speed and accuracy will be demonstrated, compared with the other cases without this technique. Besides, it will be also shown that the real time process becomes possible if the proposed method is employed.

5.1 Preliminary experiment

In the preliminary experiment, obstacle detection via moving window, a method proposed in this paper, has the advantage of acquiring information, compared to laser and radar sensors.

Falcon radar system was used as a radar sensor. The following Fig. 18 shows the measurement configuration by sensor specification.

Meanwhile, in terms of road conditions, since a three-dimensional obstacle is always in contact with a road, moving window, which is created at regular intervals based on perspective projection, can estimate the distance between a vehicle and an obstacle. That is, when an input image is like Fig. 19, an obstacle is detected in the area between the 3rd moving windows. If we consider 20[m] for the actual distance between the moving window and the hidden layer, which is designated in Fig. 3, section II, we can tell that the obstacle is located approximately 80[m] ahead.

The following are the results of the measurement about a driving situation. In Fig. 20, since the beam width of the radar sensor is set at 12[°], objects are considered obstacles if they are located 60[m] ahead and within the width of 14.9[m]. In case of highways since the width of a lane is 3.5[m], vehicles on the next lane are recognized as obstacles.

Besides, in case of the measurement distance is over 250[m], although there are no obstacles in the driving direction, false obstacles are detected as shown in Fig. 21 by the laser sensor. The measurement angle of vision sensor is 20[°] and they measure the wide range. However, with the aid of the moving window only obstacles ahead of a vehicle can be detected accurately as shown in Table 1.

### Table 1. Obstacle detection

<table>
<thead>
<tr>
<th>Driving situation</th>
<th>Radar sensor</th>
<th>Laser sensor</th>
<th>Vision sensor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed of vehicle [km/h]</td>
<td>Speed of vehicle [km/h]</td>
<td>Speed of vehicle [km/h]</td>
<td>Speed of vehicle [km/h]</td>
</tr>
<tr>
<td>Realative speed [km/h]</td>
<td>Realative speed [km/h]</td>
<td>Realative speed [km/h]</td>
<td>Realative speed [km/h]</td>
</tr>
<tr>
<td>Distance [m]</td>
<td>Distance [m]</td>
<td>Distance [m]</td>
<td>Distance [m]</td>
</tr>
<tr>
<td>Obstacle</td>
<td>Obstacle</td>
<td>Obstacle</td>
<td>Obstacle</td>
</tr>
<tr>
<td>Fig. 20</td>
<td>84</td>
<td>0</td>
<td>Detection</td>
</tr>
<tr>
<td>Fig. 21</td>
<td>89</td>
<td>0</td>
<td>Detection</td>
</tr>
</tbody>
</table>
5.2 Optimality test

The following Fig. 22 shows the detection of a lane and the result of Kalman filtering by using the change in the X-axis’s coordinate of the nearest moving window to the vehicle. The vehicle is moving from the rectilinear road to the curved road.

And the size of a moving window is adjusted by ±15 [%] The detected position in Fig. 22 does not show a big difference for the rectilinear road compared to the previous experiment, as long as the initial positions of the two moving windows are the same. However, when the vehicle enters the curved road, a big error occurs in the detected position of a lane. That is, in case of a smaller moving window, information is deficient due to the empty space between the lanes, and in case of a bigger moving window, the interference of the vehicles and the noise from the next lane decreases accuracy in detecting lanes. The proceeding direction of the vehicle based upon the position of the lane detected by the adjusted moving windows is shown in Fig. 23 and we can see that the vehicle deviates from the driving lane.

Fig. 22. Detection of the lane by different sizes of the moving window.

![Fig. 22. Detection of the lane by different sizes of the moving window.](image)

Fig. 23. Driving direction of a vehicle.

5.3. Discussion

When the moving window was adjusted to the optimal size according to the road design conditions, accuracy in detecting lanes was improved. Since the latest updated image in memory was used for the sampling image at the instance of time passing 95[mS], which is the processing time, the error of the maximum 4 frames could occur between the processing time and the input image. This is a valid time, considering the fact that a collision accident can be reduced by 60[%] with 0.5[sec] advanced warning when a vehicle moves at the speed of 100[km/h] on a freeway[6]

6 Conclusion

In this paper, the technique of selecting sub-blocks and optimizing the sizes was proposed. This new moving window technique showed excellence of detecting lanes and obstacle according to the road design conditions. The proposed size of a moving window was big enough to include some part of a lane at any time, considering the maximum admissible speed of a vehicle and the change of road conditions for maximum safety. Besides, the number of moving windows was minimized so that the area within the maximum safety distance could be searched for obstacles. The optimal size of a moving window at the current instance moved to the estimated position at the next instance based upon curve fitting and Kalman filtering techniques. By locating the middle line of the detected lane and searching obstacles, the front vehicles were considered as obstacles. This method is especially proper for the Korean freeways, the road shape of which changes frequently. It is also applicable to AGV(Autonomous Guided Vehicle), which moves slowly following land marks in various industrial fields.

References


