2012 IEEE International Conference on Vehicular Electronics and Safety
July 24-27, 2012. Istanbul, Turkey

Modeling Potential Dangers in Car Video for Collision Alarming

M. Kilicarslan and J.Y. Zheng, Senior Member, IEEE

***Abstract*— This work models various dangerous situations that may happen to a driving vehicle on road in probability, and determines how such events are mapped to the visual field of the camera. Depending on the motion flows detected in the camera, our algorithm will identify the potential dangers and compute the time to collision for alarming. The identification of dangerous events is based on the location-specific motion information modeled in the likelihood probability distributions. The originality of the proposed approach is at the location dependent motion modeling using the knowledge of road environment. This will link the detected motion to the potential danger directly for accident avoidance. The mechanism from visual motion to the dangerous events omits the complex shape recognition so that the system can response without delay.**

I. INTRODUCTION

An in-car camera has a wider angle of view as compared to radar [1] and laser range finder and is good at grasping various happening on the road. There have been many works on tracking vehicles and pedestrians with a vehicle borne camera during the vehicle driving for the purpose of safety. However, most of them are based on the shape recognition [2,3,4,5,6,7,9]. Complicated algorithms are applied to video frames to identify vehicles based on their shape, color and intensity distribution learned through off-line learning. Although such efforts normally have a certain success rate in complex urban environments, they and have not been used for the fast response to the vehicle collision situations. On the other hand, the motion based approach has added information to the shape recognition with the dynamic objection localization. According to the vehicle motion model and extracted image motion properties such as positions and velocities of features, the background and moving vehicles can be classified under normal driving conditions [10]. More importantly, the allocated motions of targets in the field of view help the vehicle following and accident preparations.

Along the line to focus on the motion information [10], we investigate more complicated situations on the road than just separating background scenes and foreground moving vehicles. In this work, we model major potential dangers visible by a facing forward camera on the road when the vehicle is moving. These dangers are caused by other vehicles approaching to the observing vehicle in normal driving scenarios. Then, we map the dangerous situations in terms of location and speed of vehicles to the video space in terms of images position and velocity. Using the image velocity detected in the video, we can identify the types of dangerous events using Bayesian framework and Decision Tree, and estimate the time to collision for early preparation. Our goal is to assist driver to realize the potential dangers at each moment with the in-car video information.

The contribution of this paper is the probability modeling of dangerous events in the video space that allows a fast classification for alarming. In the following, we will start from the motion analysis for the vehicle collision in Sec. II. Then we will describe the potential dangers of collision on the road with the probability in Sec. III. Section IV is devoted to the mapping of probability distribution to the video space. Section V addresses the future work of Bayesian decision and Decision Tree algorithm to alert the potential dangers.

II. COLLISION EVENTS AND IMAGE PROJECTION

Assuming a vehicle mounted with a video camera is moving on a road with a known speed. Surrounding vehicles may collide into it from various direction if their relative motions are towards the observing vehicle as illustrated in Fig. 1. The video camera facing forward can pick up several events. We omit the collision detection from rear and rear-side during a lane change of the observing vehicle because such situations require more cameras facing back. The camera takes video input at the frame rate of 30fr/sec. The significant feature points from corners due to the color discontinuity and intensity peaks due to tail lights are extracted in the images and tracked in real time during the vehicle movement [10]. These points may occur as surface mark on the back or side of the surrounding vehicles or environment so that they can be followed stably.

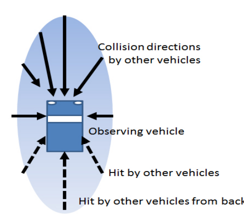


Fig. 1. Collision Directions

Denote the vehicle/camera coordinate systems by $O-XYZ$, X is the horizontal distance of an object point from the camera axis, Y is facing down, and Z is the depth of a point from the camera. The camera is set in a typical way in the vehicle in the forward directions. Assume that $x(t)$ and $y(t)$ are image positions and $u(t)$ and $v(t)$ are the horizontal and the vertical image velocity components, respectively. The image mapping from a feature point $P(X,Y,Z)$ in the 3D

Mehmet Kilicarslan: Department of Computer Science, Indiana University, Bloomington, Indiana, USA

Jiang Yu Zheng: Department of Computer Science, and Transportation Active Safety Institute, Indiana University Purdue University Indianapolis,Indiana, USA space in front of the observing car to the camera frame is

$$x(t) = \frac{fX(t)}{Z(t)}, y(t) = \frac{fY(t)}{Z(t)} \tag{1}$$

where f is the camera focal length.

The relative velocity of a scene point $P(X,Y,Z)$ can be calculated from the relative translation $(T_x(t),T_y(t),T_z(t))$ with respect to the camera, and rotation $(R_x(t),R_y(t),R_z(t))$ of the camera. On a flat road situations we can assume $R_x(t)=0$ and $R_z(t)$. Also, we can assume $V_y(t)=0$ since there is no velocity change in Y direction. The relative speed of the scene point P to the camera is,

$$\begin{aligned}&(V_x(t),V_y(t),V_z(t)) \\ &= (T_x(t),T_y(t),T_z(t)) + (X,Y,Z)\times(R_x(t),R_y(t),R_z(t))\end{aligned} \tag{2}$$

The image velocity (u,v) can be computed as in [10] as

$$u(t) = \frac{fT_x(t) - x(t)T_z(t)}{Z(t)} \quad v(t) = \frac{-yT_z(t)}{Z(t)} \tag{3}$$

with $T_y(t)=0$, $R_x(t)=0$, $R_y(t)=0$ and $R_z(t)=0$. Mildly curved road is omitted here but can be considered as in [10] by adding a common flow component to the above equations. Vehicle turning at a road crossing will not be modeled in this paper, since it could be a more complex scenario that needs to be examined in a different way. Thus, we can consider that, at a short time duration, the vehicles are performing translation only in various directions. Therefore, the image properties such as vanishing point can be employed in describing the Time-to-collision (TOC).

Considering the width of the observing vehicle and a target vehicle as shown in Fig. 2(a), the relative moving direction of the target vehicle outside the arrows will not cause a collision. Otherwise, the target vehicle approaches to the camera in any angle α within the gray region will yield a potential collision. From the camera point of view, the collision direction is equivalently indicated within the yellow region in Fig. 2(b), considering a secure width of the observing vehicle added to the width of the target vehicle. A linear motion in α direction with respect to the camera will have a vanishing point or focus of expansion (FOE) in the image frame for the feature points on the target vehicle. The yellow range in Fig. 2(b) is thus mapped onto the image frame as a scope also indicated in yellow as shown in Fig.2(c) on the projected horizon.

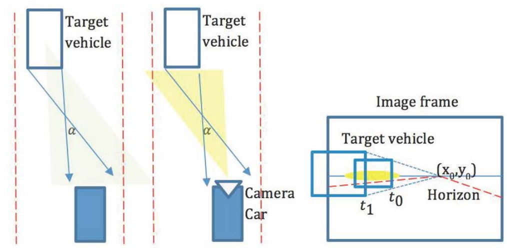
Let us denote this span of the dangerous moving direction as $[x_l,x_r]$ wider than the target vehicle in the image at time t_0. At time t_1, the image velocities (u,v) of points on the target vehicle form a vanishing point (x_0,y_0) under the linear motion during $[t_0,t_1]$. The vanishing point should be on the projection of horizon if the road is horizontal, i.e.,$y_0=0$. If $x_0 \in [x_l,x_r]$, there will have a potential collision because the relative motion vector falls into the collision direction. Otherwise, two cars will not collide.

In a relative speed T_z, distance Z takes the Toc to be reduced to zero. It is not difficult to derive from the computation of image velocity that

$$T_{oc} = \frac{Z}{T_z} \sim \frac{\sqrt{(x-x_o)^2+(y-y_o)^2}}{\sqrt{u^2+v^2}} \tag{4}$$

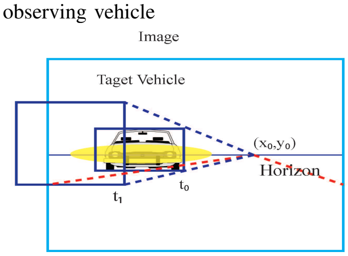
where (x_0,y_0) is the vanishing point in the image plane.

Although the vanishing point can be computed from the intersection of motion vectors (u,v) of many points and the projection of horizon $(y=0)$ in the image frame, the accuracy is not guaranteed if the error in (u,v) is taken into account. More importantly, the width of the vehicle is unknown. The recognition of the vehicle occupied area in the video is another ongoing research task and has not been used in real time video at a high successful rate. Therefore, a more reliable system should be developed for early alarm of collision based on probability.



(a) The widhts of target and observing vehicle

(b) The collision direction



(c) Vanishing point (x_0,y_0) in an image frame.

Fig. 2. The principle to compute Time to Collision

III. Model Potential Collisions on Road in Probability

Because a car visible at distance can rarely causes a collision to the observing vehicle even with its maximum speed, we model potential collisions by focusing on events on adjacent lanes in a relatively close range. As the observing car is moving on the road, we identify six scenarios briefly that may cause potential collision if both cars continue their motion. These situations, as shown in Figure 3, are denoted by events E=$\{V_f,V_o,C_r,C_l,S_r,S_l\}$ as follows.

V_f : The relative distance of a *front* car is reduced due to its slowing down, stopping, or a speed up of the the observing car.

V_o : A vehicle approaches from the opposite lane with a fast speed.

C_r : A vehicle approaches to the driving lane from right without slowing down or stop. This appears at a road crossing or a possible merge of vehicles into the same lane with the observing vehicle.

C_l : A vehicle approaches the driving lane from a left crossing road with the possibility of merging into the same lane of the observing vehicle.

S_r : A vehicle moving in the same direction on the right lane cuts in.

S_l : A vehicle moving in the same direction on the left lane cuts in.

For each of them, we can define the vehicle location in a normal distribution. Also, the vehicle speed is in another normal distribution. Among the speed distribution, only partial probability for the target car approaching the camera may cause danger. For example, the front vehicle in the same direction has a relative speed distributed in normal distribution. Only a negative value that narrows down the relative distance between two cars may cause the danger of collision. Outside these ranges, the vehicles are not considered to be danger because of a far distance or a motion leaving the camera. Besides the position and speed of the target vehicles, the size of vehicle and the road width are also briefly set in terms of probability.

In the camera field of view, there may also have background such as road, buildings, parked vehicles, trees, poles, etc. on road sides. We assume that the driver of the observing vehicle will not bump into roadside objects to cause self-accident, then we can treat the background as static scenes.

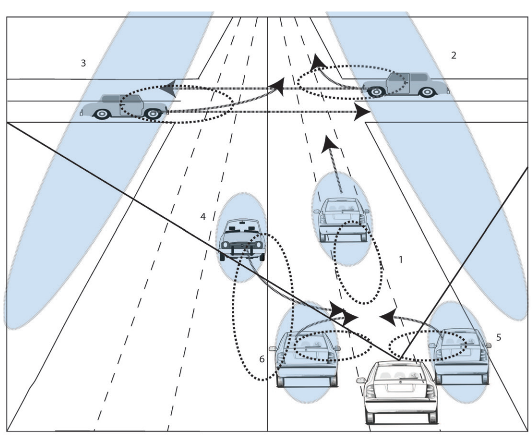


Fig. 3. Possible collisions from different directions with various relative speeds. The collision from rear and side rear happens when the observing vehicle is performing lane change or slowing down. They are not visible by the video camera facing forward. The blue ellipses show the target vehicle position distributions and dotted ellipses indicate their relative speeds that may cause danger.

Assume the vehicle position (X,Z) follows a $2D$ normal distribution. The relative speeds (T_x,T_z) in X and Z directions also follow normal distributions. The height of a vehicle can be set in a probability as $p(Y) \sim (Y+3)/(Y+4)$, where $Y \in [-3m, 1m]$, with the camera at the height of 1m from the ground. In general, the features are more extractable at the bottom part close to road (bumper, shadow, window, etc.) than at high positions (only from large trucks) for all types of vehicles. Because the vehicle location, height, and speed are independent, for each dangerous event $D \in \{V_f, V_o, C_r, C_l, S_r, S_l\}$,

$$p(D) = p((X,Z) \in D)p(Y)p((T_x,T_z) \in D) \qquad (5)$$

As the special setting of those potential dangers of collision, the detailed data for the normal distributions are as in TABLE I.

Sec.	σ_X	μ_X	σ_Z	μ_Z	σ_{Tx}	μ_{Tx}	σ_{Tz}	μ_{Tz}
V_f	0	6	20	20	0	5	0	10
V_o	20	10	10	60	-5	10	-15	5
C_r	-20	10	10	60	5	10	-15	5
C_l	-5	5	20	60	0	5	-30	15
S_r	30	10	10	10	-10	5	5	5
S_l	-30	10	10	10	10	5	5	5
Background	0	6	-	-	-	-	-15	5

TABLE I

VARIABLES OF 3D NORMAL DISTRIBUTIONS FOR THE POSITIONS AND VELOCITIES OF VEHICLES(UNIT METER AND METER/SECOND).

The parameters of the normal distributions are selected according to the real environment of roads. We have selected the sufficiently large standard deviations to cover wide ranges of possible cases if no specific data are available. The potential dangerous scopes are not set extremely far to avoid false-alarming and are not set at a too close range because of insufficient time to response to such accidents.

IV. MAPPING DANGEROUS EVENTS INTO VISUAL FIELD

In real-time vehicle detection, not all the parameters in (X,Y,Z,T_x,T_z) are extractable from the target vehicle. What we can obtain are only the image properties of the objects of interest. We therefore need a probability model to describe the image properties such as its position (x,y) and velocity (u,v) on the target. It is clear from Eqs. 1, 3, 4 that u is determined from X,Z,T_x,T_z and v is determined from Y,Z,T_z, respectively. Using positions and velocities in the 3D space, we can now map both background and vehicle probabilities into the camera space (x,y,u,v) in terms of probability $p(x,y,u,v)$. Further, we will examine the probability distributions in their subspaces (x,u) and (y,v) separately, because a four dimensional space is not easy to classify in real time collision computation.

A. Background Probability Maps

On a flat road, background is spread on both sides of the road. Assuming the vehicle speed, V, is known from the vehicle encoder or GPS, for the background, its $V_x = 0$ and $V_z = -V$ with respect to the camera, and its X is thus fixed. The horizontal component of the image velocity in Eq.3 becomes

$$u(t) = \frac{fXV}{Z^2(t)} = \frac{Vx^2(t)}{fX} \qquad \text{for V>0} \qquad (6)$$

In general, if $p_\chi(\chi)$ is the PDF of a random variable χ, and β is a monotonic function of χ, then the PDF of β can be calculated as

$$p_\beta(\beta) = p_\chi(f^{-1}(\beta))|\frac{\partial f^{-1}(\beta)}{\partial \beta}| = \frac{p_\chi(f^{-1}(\beta)}{|\frac{\partial f^{-1}(\chi)}{\partial \chi}|} \qquad (7)$$

according to [9]. Now, the image motion behavior for a background point can be computed by the likelihood $p(x,u|B)$. We use Bayes theorem in order to map the background B to the image since there is no one-to-one mapping from X,Z,V to x,v. The likelihood of background is as follows,

$$
\begin{aligned}
p(x,u|B) &= p(x,u|(X,Z)\in B) && (8)\\
&= \int_X p(X)p(x,u|X)\,dX && \mathit{Bayesian}\\
&= \int_X p(X)p(Z(x,u)|X)p(V(x,u)|X)\,dX && Z,V\ \mathit{independent}\\
&= \int_X p(X)p(Z=\frac{fX}{x}|X)p(V=\frac{vfX}{x^2}|X)\,dX && Eqs.1,7
\end{aligned}
$$

In order to separate background into two parts, right and left side, we define the $3D$ background distribution $p(X)$ as follows.

$$
p(X) = (1-e^{(-\frac{X^2}{2\sigma_x{}^2})})/(1+|X|) \qquad (9)
$$

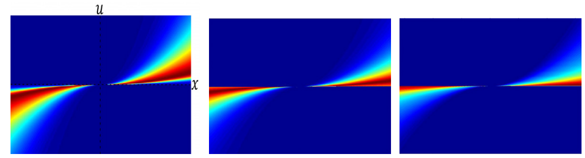
with the homogenous distribution in Z direction. Figure 4 shows background PDFs for different relative speed values.

In addition to the background scenes on road sides, we further identify the road area. Road has its own position and velocity informations that help us to model PDFs in terms of $p(x,u)$ and $p(y,v)$. For road surface, $T_z = -V$ and $T_x = 0$. In addition to these, X follows a normal distribution centered at 0, and Y is exactly at $1m$ according to the fixed height of the camera on the observing vehicle, i.e., $p(Y) = 1$ and $p(Y \neq 1) = 0$. With the values set in Table I, the road surfaces is mapped into the video space as

$$
\begin{aligned}
p(x,u) &= p(x=\frac{fX}{Z}, u=\frac{xV}{Z}) && (10)\\
&= p(Z=\frac{fX}{x}, V=\frac{uXf}{x^2}) && \text{Eq.7},\ Z=\tfrac{fX}{x}\\
&= \int_X p(X)p(Z=\frac{fX}{x}, V=\frac{uXf}{x^2}|X)\,dX && \mathit{Bayesian}\\
&= \int_X p(X)p(Z=\frac{fX}{x}|X)p(V=\frac{uXf}{x^2}|X)\,dX\\
& && Z,V\ \mathit{independent}\\
&= \int_X p(X)p(V=\frac{uXf}{x^2}|X)\,dX && \mathit{Invariant\ in}\ Z
\end{aligned}
$$

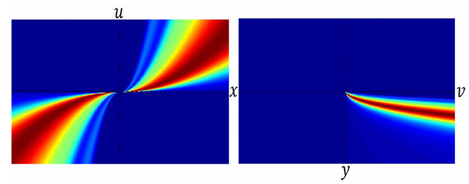
Similarly $p(y,v)$ can be computed as

$$
\begin{aligned}
p(y,v) &= p(y=\frac{fY}{Z}, v=\frac{yV}{Z}) && (11)\\
&= p(Z=\frac{fY}{y}, V=\frac{vYf}{y^2}) && \text{Eq.7},\ Z=\tfrac{fY}{y}\\
&= \int_Y p(Y)p(Z=\frac{fY}{y}, V=\frac{vYf}{y^2}|Y)\,dY && \mathit{Bayesian}\\
&= \int_Y p(Y)p(Z=\frac{fY}{y}|Y)p(V=\frac{vYf}{y^2}|Y)\,dY\\
& && Z,V\ \mathit{independent}\\
&= \int_Y p(V=\frac{vYf}{y^2}|Y)\,dY && \mathit{Invariant\ in}\ Z, p(Y)=1
\end{aligned}
$$



(a) V=50 km/h (b) V=25 km/h (c) V=5 km/h

Fig. 4. Background PDFs for various relative speed, V.



(a) $p(x,u)$ (b) $p(y,v)$

Fig. 5. Ground(road) surface PDFs in (x,u) and (y,v) spaces. V=50km/h

B. Modeling Dangerous Vehicles in Video

For the probability distributions of potential dangerous vehicles on the road, their probabilities in terms of image properties (x,y,u,v) are computed. Assume a dangerous event $D \in E$, the general form next is computed as

$$
\begin{aligned}
& p(x,y,u,v|D) = p(x,y,u,v\ |(X,Y,Z,T_x,T_y,T_z)\in D) \qquad (12)\\
&= p(x=\frac{fX}{Z}, y=\frac{fY}{Z}, u=\frac{fT_x - xT_z}{Z}, v=\frac{-yT_z}{Z})\\
&= \int_Z p(Z)p(x=\frac{fX}{Z}, y=\frac{fY}{Z}, u=\frac{fT_x - xT_z}{Z},\\
&\qquad v=\frac{-yT_z}{Z}|Z)\,dZ \qquad \mathit{Bayesian}\\
&= \int_Z p(Z)p(X=\frac{xZ}{f}, Y=\frac{yZ}{f}, T_z, T_x=\frac{Zu+xT_z}{f},\\
&\qquad T_z=\frac{-vZ}{y}|Z)\,dZ \qquad Eq.7\\
&= \int_Z p(Z)p(X=\frac{xZ}{f}|Z)p(Y=\frac{yZ}{f}|Z)p(T_z=\frac{-vZ}{y}|Z)p(T_z,\\
&\qquad T_x=\frac{Zu+xT_z}{f}|Z)\,dZ\ \ (X,Z),Y,(T_x,T_z)\ \mathit{are\ independent}\\
&= \int_Z p(Z)p(X=\frac{xZ}{f}|Z)p(Y=\frac{yZ}{f}|Z)p(T_z=\frac{-vZ}{y}|Z)\\
&\qquad \int_{T_Z} p(T_Z)p(T_x=\frac{Zu+xT_z}{f}|T_z,Z)\,dT_Z\,dZ \qquad \mathit{Bayesian}
\end{aligned}
$$

Although the likelihood probability distribution $p(x,y,u,v|D)$ can be computed offline through integration, its storage as a lookup table in a four dimensional space is not feasible for online retrieval. The classification in such a space is more time consuming. Therefore, in real practice, we compute the subspace probability distribution $p(x,u)$ and then $p(y,v)$ to discriminate the potential dangerous in a stepwise way.

$$
\begin{aligned}
p(x,u|D) &= p(x,u\ |(X,Z,T_x,T_z)\in D) \qquad (13)\\
&= p(x=\frac{fX}{Z}, u=\frac{fT_x - xT_z}{Z})
\end{aligned}
$$

$$= \int_Z p(Z) p(x = \frac{fX}{Z}, u = \frac{fT_x - xT_z}{Z} | Z)\, dZ \qquad Bayesian$$

$$= \int_Z p(Z) p(X = \frac{xZ}{f}, T_z, T_x = \frac{Zu + xT_z}{f} | Z)\, dZ \qquad Eq.7$$

$$= \int_Z p(Z) p(X = \frac{xZ}{f} | Z) p(T_z, T_x = \frac{Zu + xT_z}{f} | Z)\, dZ$$

$$(X,Z), (T_x, T_z)\ independent$$

$$= \int_Z p(Z) p(X = \frac{xZ}{f} | Z) \int_{T_Z} p(T_Z) p(T_x = \frac{Zu + xT_z}{f} | T_z, Z)\, dT_Z\, dZ$$

$$Bayesian$$

$$= \int_Z \int_{T_z} p(Z) p(X = \frac{xZ}{f} | Z) p(T_Z) p(T_x = \frac{Zu + xT_z}{f} | T_z, Z)\, dT_Z\, dZ$$

Now, we calculate $p(y,v)$,

$$p(y,v|D) = p(y,v\ |(Y,Z,T_z) \in D) \qquad (14)$$

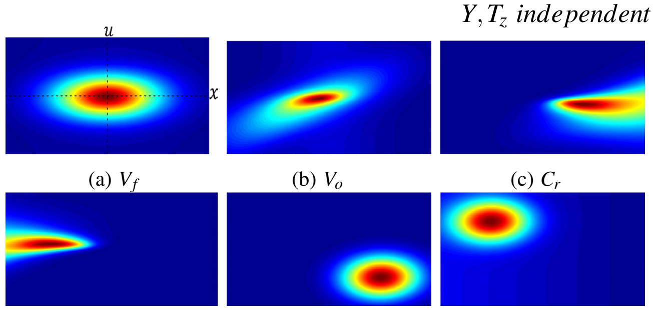
$$= p(y = \frac{fY}{Z}, v = \frac{-yT_z}{Z})$$

$$= \int_Z p(Z) p(y = \frac{fY}{Z}, v = \frac{-yT_z}{Z} | Z)\, dZ \qquad Bayesian$$

$$= \int_Z p(Z) p(Y = \frac{yZ}{f}, T_z = \frac{-vZ}{y} | Z)\, dZ \qquad Eq.7$$

$$= \int_Z p(Z) p(Y = \frac{yZ}{f} | Z) p(T_z = \frac{-vZ}{y} | Z)\, dZ$$

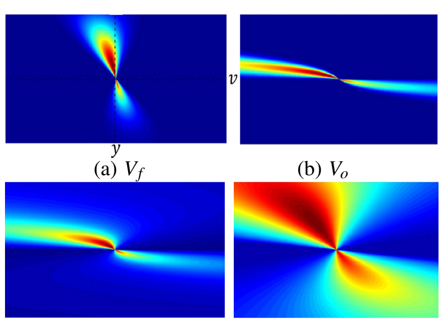
$$Y, T_z\ independent$$



(a) V_f (b) V_o (c) C_r

(d) C_l (e) S_r (f) S_l

Fig. 6. All possible scenarios probability distribution maps in (x,u) space.



(a) V_f (b) V_o

(c) C_l and C_r (d) S_l and S_r

Fig. 7. All possible scenarios probability distribution maps in (y,v) space.

In general road environment, (x,u) reflects the location of target vehicles and possible movements, whereas, (y,v) depends on the height of vehicles. Its scaling in the vertical direction provides an important clue of distance to a target vehicle, which can be used in computing time to collision. In addition, as shown from Eq.14, $p(y,v)$ does not depend on x. Thus dangerous cases C_r, C_l and S_r, S_l have the same probability maps in (y,v) space.

Now, we compute $p(x,y,u,v|V_f)$, $p(x,y,u,v|V_o)$, $p(x,y,u,v|C_r)$, $p(x,y,u,v|C_l)$, $p(x,y,u,v|S_r)$, and $p(x,y,u,v|S_l)$ for each dangerous case by filling Eqs.13 and 14 with the parameters in Table I. These results serve as the likelihood probability for detecting dangerous events in the video. Figures 6 and 7 show plots of all the probabilities $p(x,y,u,v|D)$ in the subspace of (x,u) and (y,v).

C. Modeling Positions of Dangerous Vehicles in Video

In the similar deduction as above, we can also obtain the probability distribution of dangerous vehicles in the image frame as shown in Fig. 8. For the possible appearing positions described by the image subspace $p(x,y)$, it is not sufficient to discriminate which type of events and background are there. However, the space provides the attention regions for assigning the computing power in feature detection and tracking. For example, the points extracted the trees and high buildings as well as the ground are processed with lower priorities than the points in the hot regions where dangerous vehicles appear move Figure 8 below this line, above next section.

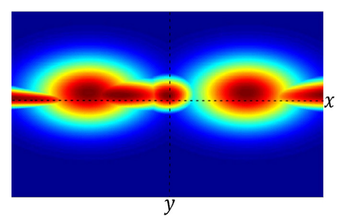


Fig. 8. Visualization of dangerous vehicle positions in (x,y) image space. From left to right, the high probability regions correspond to $C_l, S_l, V_o, V_f, S_r, C_r$ respectively.

For the varying vehicle velocity V, the dangerous cases V_f, S_r, and S_l did not change. Since, vehicle velocity takes only in T_z distributions. In all cases, we used relative speed to target vehicle, thus $p(T_z)$ distributions are invariant for case $\{V_f, S_r, S_l\}$. However, for the remaining cases of $\{C_l, C_r, V_o\}$, their $p(x,u|D)$ are affected by the speed of observing vehicle. As shown in Figs.11 and 12, events V_f through S_l are separable from background and road. Only V_o is merged into V_f at a high speed of observing vehicle, which can be considered at highway. They are not separable with the image velocity (x,u), but should be based on the vehicle size change related to (y,v).

	Physical Environment	Values
$\int_X$	Wider than a road to include all backgrounds in video	-40 $\sim$ 40m
$\int_Y$	Higher than vehicle heights	-3 $\sim$ 1m
$\int_Z$	From camera position to distance close to infinity	0 $\sim$ 200m
$\int_{T_z}$	Range for relative speed	-60 $\sim$ 0m/s

TABLE II

PARAMETER SELECTION IN PROBABILITY COMPUTATION

V. Alert Collision Events

In this section, we discuss the possible approaches to alert the potential events by computing the motion signals in the video frame and the likelihood of the dangerous events. We will identify if the points extracted have their motion behaviors that are predicted as the potential dangerous events.

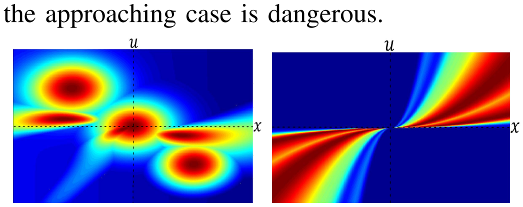
Denote the background event B, and non-dangerous vehicle event $\neg E$, we can use Bayesian rule

$$
\begin{aligned}
p(E|x,y,u,v) &\propto p(x,y,u,v|E)p(E) \quad (15)\\
p(\neg E|x,y,u,v) &\propto p(x,y,u,v|\neg E)p(\neg E)\\
p(B|x,y,u,v) &\propto p(x,y,u,v|B)p(B)
\end{aligned}
$$

to classify the feature points. Further, using the condition probability, we have the posterior probability of potential collision as

$$
\begin{aligned}
p(E|x,y,u,v) &\propto p(x,y,u,v|E)p(E) \quad (16)\\
&= p(x,y|E)p(u,v|x,y,E)p(E)\\
&= p(x,y)p(u|x,y,E)p(v|x,y,u,v,E)p(E)
\end{aligned}
$$

It is noticeable that the dangerous events are partially separable from the non-dangerous motions and the motion of background in the subspaces of $p(x,u)$ and $p(y,v)$. To identify dangerous events with the image output of (x,y,u,v), a decision tree can be constructed on subspaces of $p(x,y)$, $p(x,u)$, and $p(y,v)$. The root is separated to two subtrees i.e., potential target vehicle areas or background (sky, highrise, etc.) by using the distribution of vehicle appearing positions $p(x,y|E)$ given in Fig. 8. The feature points classified in the background are not processed anymore. In the possible locations of the collision vehicles, background and non-danger vehicles $\{B \cup \neg E\}$ can be separated from dangerous vehicles E by examining $p(x,u|D)$. To identify whether a captured point might possibly be on a collision vehicle, we can examine $p(x,u|E) > \delta_d$ as indicated by Fig. 9(a), where δ_d is a threshold. This limit a small u so that the target vehicle is almost moving along the line of sight towards the camera. However, this still leaves some ambiguity as depicted in Fig. 10. Finally, we can check $p(y,v|D)$ as in Fig. 7 to classify if the point is on an object that is approaching to the camera, stays at the same distance, or leaves away. Only the approaching case is dangerous.



(a) $p(x,u|E)$ where V_f and V_o are hard to discriminate

(b) Background and road $p(x,u|B)$

Fig. 9. The projected subspace of $p(x,u|E)$ and $p(x,u|B)$

VI. Conclusion

This paper aims at real time collision alarming for safety driving. The main approach is to detect motion information in car video in combined with the knowledge of road environment and traffic flow. We model the location specific motion in the in-car video that gives more precise prediction

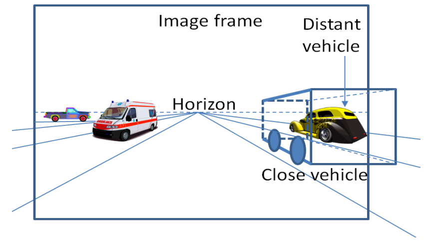
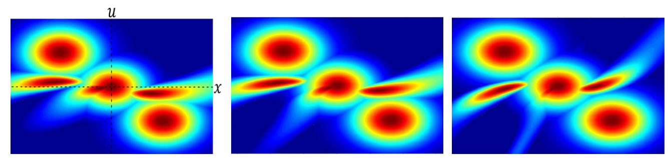
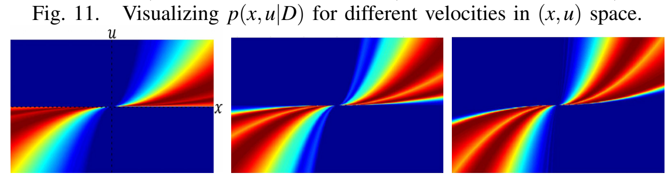


Fig. 10. Image frame showing two cars that may appear at the same location with similar horizontal velocities. The close one is performing Sr action while the distant car is not a danger because it is moving ahead in a fast speed need a little bit space between two figures. One line after this line.



(a) $V = 25km/h$ (b) $V = 50km/h$ (c) $V = 100km/h$

Fig. 11. Visualizing $p(x,u|D)$ for different velocities in (x,u) space.



(a) $V = 25km/h$ (b) $V = 50km/h$ (c) $V = 100km/h$

Fig. 12. Visualizing $p(x,u|B)$ for different velocities in (x,u) space.

of dangerous events than just using motions. Bayesian rule is used in finding and alarming potential dangers based on their precomputed likelihood PDFs in look-up tables. We will further extend this results to various situations to verify our approach. In the future, we will refine our models so that the classification of the potential dangers will be more precise and compliance to real data. We will also extend the road and environment model to curved ones and more complicated scenarios.

References

[1] H. Takizawa, K. Yamada, T. Ito, (2004), Vehicles detection using sensor fusion, IEEE Intelligent Vehicle 04, 238- 243

[2] C. Hoffman, T. Dang, C. Stiller, (2004), Vehicle detection fusing 2D visual features, IEEE Intelligent Vehicle, 280-285

[3] L. Gao, C. Li, T. Fang, Z. Xiong, (2008), Vehicle detection based on color and edge information, LNCS, Image Analysis and Recognition, Vol. 5112, 1611-3349

[4] J. Dubuisson, S. Lakshmanan, and A.K. Jain, (1996 March), Vehicle segmentation and classification using deformable templates, IEEE Trans. on PAMI, Vol.18, (3), 293-308.

[5] T. N. Tan, K. D. Baker, (2000 Aug.), Efficient image gradient based vehicle localization, IEEE Trans. on Image Processing, 9(8), 1343-1356.

[6] Y. Guo, Y. Shan, H. Sawhney, R. Kumar, (2007), PEET: Prototype embedding and embedding transition for matching vehicles over disparate viewpoints, IEEE CVPR 2007, 17-22.

[7] C. R. Wang, J.-J. Lien, (2008), Automatic vehicle detection using local features a statistical approach, IEEE ITS, 9(1), 83-96.

[8] D. Stirzaker (2003). Elementary Probability. Cambridge.

[9] S. Sivaraman, M. Trivedi, (2010) A general active-learning framework for on-road vehicle recognition and tracking, IEEE ITS, 11(2), 267-276

[10] A. Jazayeri, H. Cai, J. Y. Zheng, M. Tuceryan (2011), Vehicle Detection and Tracking in Car Video Based on Motion Model, IEEE ITS, 12(2), 583- 595