Hierarchical Probabilistic Network-based System for Traffic Accident Detection at Intersections

Ju-Won Hwang, Young-Seol Lee, and Sung-Bae Cho  
Department of Computer Science  
Yonsei University  
Seoul, Republic of Korea  
{juwon,tiras}@sclab.yonsei.ac.kr, sbcho@cs.yonsei.ac.kr

Abstract—Every year, traffic congestion and traffic accidents have been rapidly increasing in proportion to increasing number of vehicles. Although the roadway design and signal system have been improved to relieve traffic congestion, traffic casualties and property damage do not decrease. The traffic accident is a serious issue of society because vehicle is a primary means of transportation. This paper develops a real-time traffic accident detection system (RTADS): This system helps us to cope with accidents and discover the causes of traffic accident by detecting the accident. We gathered video data recorded at several intersections and used them to detect accidents at different intersections which have different traffic flow and intersection design. However, because the data gathered from intersections have incompleteness, uncertainty and complicated causal dependency between them, we construct probability-based networks which calculate based on the probability for correct accident detection. This system instantly sends the detected result to managers using accident alarm system. RTADS features real time accident detection and analysis of the cause of accidents. In performance evaluation, the proposed system achieved a detection rate of 97% with a correct detection rate of 92% and a false alarm rate of 0.77%.

Keywords—traffic accident at intersections, accident detection system, real-time traffic accident detection system (RTADS), dynamic Bayesian networks

I. INTRODUCTION

Every year, traffic casualties and property damage are increasing in proportion to increasing the traffic accidents and the number of vehicles. Regardless of expansion of road safety facilities and improved traffic conditions such as signal system at road, roadway design, and performance of vehicle, a decrease in the rate of traffic accidents falls short of our expectation and is now recognized that the problem should be addressed as soon as possible. Especially, the rate of road intersection accidents has not significantly changed in more than two decades [1]. In Japan, more than half of all traffic accident occurs at intersections [2], because each intersection has the diversity of the complexity and characteristics such as different intersection design, traffic flow, traffic volume, signalized/unsignalized and average traffic speed. So drivers need to concentrate on several intersection designs. Intersections are among the most hazardous sites on U.S. roads [2]. The statistic of accidents in the year 2002 in the USA reported that 50 percents of all reported accidents, approximately 3.2 million accidents, were intersection-related [1]. Also, when traffic accident occurred at intersections, it can damage vehicles in the other direction and have a bad effect on the subsequent accidents with increasing traffic congestion because of the above reasons such as the complexity and characteristics. So, several studies have been researching for resolution above the problem. However, operators of many systems have to remotely monitor accidents using closed-circuit television (CCTVs) to handle accident promptly. This method is not efficient because it requires much time and effort. Therefore, development of system was required to prevent the subsequent accident by detecting accident promptly.

This paper aims to develop a system that helps us to cope with the accident by rapidly detecting a traffic accident at intersections. We consider the time-series information about situations and analyze and define complicated relationships between states of each feature in kinds of accidents and degree of accident at intersections. Bayesian networks are used to design complicated relationships. We infer traffic accident types as well as occurrence of the accident. The accidents such as a broadside collision, a head-on collision, a rear-end collision, and a collision between vehicle and pedestrian are extracted by using many features. We evaluate the real-time traffic accident detection system (RTADS) using real data. The real-time traffic accident detection system (RTADS) consists of data gathering unit, preprocessing unit, accident inference unit, and accident alarm unit.

This paper is organized as follows. In Section II, we will outline the related work and the previous work in the field and its differences to our proposed method. In Section III, we will describe the architecture of the real-time traffic accident detection system. In Section IV, the performance of the RTADS is evaluated. Section V wraps up the paper with a conclusion and suggestions for future work.
II. RELATED WORKS

There are many studies for detecting traffic accident at an intersection in the past several decades. Pattern recognition techniques such as image processing, K-means clustering, fuzzy logic, Bayesian networks, neural networks, and decision trees have been attempted [4]–[14].

In 2007, Ki et al. [2] suggested the accident recording and reporting system (ARRS) by using image processing. The accident detection algorithm is summarized as follows. Step 1, extracts vehicle objects from the video frame and step 2 tracks moving vehicles (MV) by the tracking algorithms. Step 3 extracts feature such as acceleration, position, area, and direction of the MV, step 4 calculates VF (rapid velocity variation), PF (variation rate of the position), SF (variation rate of the area), and DF (variation rate of the direction) and step 5 estimates the summation of the accident indices (VF+PF+SF+DF) and identifies the accident.

In 2003, Bruce et al. [3] proposed a system for automated traffic accident detection at intersection. It was designed using a 3-segment of audio signal. The system consists of the two-class and multiclass modes. The system has three main signal processing stages which are feature extraction, features reduction, and classification. The system consistently results in accident detection accuracies of 95% to 100%.

In 2000, Kamijo et al. [15] proposed an algorithm based on spatio-temporal Markov random field. The algorithm tracks the state of each pixel in an image, state transiting along time axis as well as x and y axes in an image. The overlapping portion of the vehicles passing the intersection causes the difficulty in tracking vehicles. Also, they proposed an HMM-based model by learning various patterns of each vehicle and showed good performance in the awareness of traffic situations at the intersection. The system can recognize bumping, passing, and jamming.

In 2006, Zhang et al. [16] proposed a decision-making model for effective reasoning and traffic knowledge management near the intersection. The automated accident detection system consists of two modules that are data processing module and accident detection module. The accident detection module use Bayesian networks to extract accidents probabilistically. The networks consist of three traffic events at both upstream and downstream intersections and five traffic parameters such as turning count at the upstream intersection, volumes of both intersections, occupancy of both intersections. Total 40 different types of arterial road accident are simulated to test the performance of the algorithm.

In 2007, Salim et al. [17] developed an intersection collision detection system which is able to adapt to different types of intersections by acquiring the collision patterns of the intersection through data mining techniques. The patterns are used for matching vehicle pairs to be calculated for the possibility of future collision events. To evaluate the system, they used a computer based simulation of two different scenarios such as an intersection with traffic lights and that without traffic lights.

These systems have not considered complex and causal relationships between features related to traffic accident. Actually the state values of features have an immediate connection with several situations such as passing, bumping, jamming, etc. Also, context awareness in outdoor environment such as traffic accidents on road has to deal with the time-series (t-2, t-1, and t) information because we cannot understand a situation for inference using information at the current time (t).

III. PROPOSED SYSTEM

In this section, we describe the RTADS, which consists of data collection unit, preprocessing unit, accident inference unit, and accident alarm unit. In data collection unit, system gathers video from a camera located on the intersections. Videos are sent to preprocessing unit in a server. The preprocessing unit extracts basic information of objects and features of the objects such as coordinate, moving vector, velocity and direction. The accident inference unit infers types and probability of the accident using above features and relationships between features. Finally, accident alarm unit reports the result of inference to operators so that they can rapidly handle the accidents.

An overview of the system is presented in Fig. 1. The technical details of each component are in the following subsections.

A. Preprocessing

Video data are sent to a server from a camera located at intersections. Preprocessing unit converts the video data to a usable form. This preprocessing unit extracts the basic information of objects such as overlapping between objects, stopping, velocity, differential motion vector of objects, and direction of each object. We select objects to extract features. If one object exists in a frame, we do not extract features of one object. If two or more objects exist in a...
frame, we select every a pair which consist of two objects and extract features of objects and calculate probability of accident.

1) The basic information of objects: First of all, the basic information of objects must be extracted for feature extraction. We empirically determine suitable frame rate as one per 0.5 second which is analyzed collected videos at many intersections that have many different factors related to environment and features of objects. If we use the frame rate more than 0.5 second, we often miss major information about flow of objects. Whereas, if we use frame rate of less than 0.5 second, it tends to extract too detailed information. For example, it can extract movement of branches in a tree. The objects have features that are active movements at intersections. We extracted objects by using tracking algorithm based on inter-frame and a difference of pixel by comparing the current frame and previous frame. The basic information consists of object ids, left, top coordinates, right, bottom coordinates, and center(x, y) coordinates.

2) The overlap of objects: Generally, in the event of a traffic accident, we can check the overlap of objects. Hence, we extract feature about the overlap using coordinate of objects. The number of cases of overlap is nine. However, when objects move close to each other, we can extract overlap information. This feature is used in the accident inference unit. Also, we consider a case that all coordinate of an object were included all coordinate of other object. The following equation is used for the overlap of objects:

\[
 CO(T) = \begin{cases} 
 Yes, & \text{if } O_2 \subseteq O_{1r} \text{ or } r \leq O_2^r \\
 No, & \text{otherwise}
\end{cases}
\]

where CO (T) represents the overlap of coordinate of objects at this point of time (T). l, r, t, and b mean left, right, top, and bottom, respectively. Yes and No are state values of feature. O1 and O2 mean object1 and objects2.

3) The stop of objects: In general, we can check the stop of each object because objects usually stop after the accident. Therefore, we extract stop of objects using difference between current coordinates and previous coordinates. The following equation shows the stop of objects:

\[
 CS(T) = \begin{cases} 
 Yes, & \text{if } O_{1l} \subseteq O_{1r} \text{ and } O_{1t} \subseteq O_{1b} \\
 No, & \text{otherwise}
\end{cases}
\]

where CS (T) is stop of coordinates of each object at the point of time (T). Yes and No are state values of the feature. We consider the information as an independent feature because a stop of coordinates on all objects is a very important feature related to accident in analysis result of video.

4) Velocity: In general, we can check the variation of velocity of each object before/at/after an accident. If a driver predicts dangers, velocity of the driver’s vehicle decreases because of stepping on the brake. Therefore, we calculate velocities of objects using current coordinates and previous coordinates. The following equation represents it:

\[
 V(T) = \begin{cases} 
 \text{Fast,} & \text{if } 90 \text{ km/h} < v < \frac{v}{t} \\
 \text{Normal,} & \text{if } 60 \text{ km/h} < v \leq 90 \text{ km/h} \\
 \text{Slow,} & v < 60 \text{ km/h}
\end{cases}
\]

where V (T) is velocity of an object at the point of time (T). As the average speed of a vehicle at general intersection is 60–90km/h, we classify the velocity into three states such as fast, normal, and slow.

5) Differential motion vector: Differential motion vector of objects is very important to detect traffic accident at intersection. Generally variations of velocity and directions before and at accident have to be considered. After collision, direction and velocity of objects are changed, because an external force. Speed of objects decrease according to driver’s braking and collision at the accident. So we consider differential motion vector of objects and use in accident inference unit. The equation (4) is used for differential motion vector:

\[
 VR(T) = |(v_{x(t-1)}^2 + v_{y(t-1)}^2) - (v_{x(t)}^2 + v_{y(t)}^2)|
\]

where VR (T) is differential motion vector at the point of time (T). Vx(t-1) and Vy(t-1) mean the difference between velocity of x-coordinate and y-coordinate of objects, respectively, at time (T). States of the feature have three discrete values such as Up-Down, Up and Rest that represent differential motion vectors between previous and current times. Up state means difference motion vectors larger than the previous times. Up-Down state means reducing the difference motion vectors after Up state and Rest state means difference motion vectors more scarce than the previous times. Fig. 2 compares the differential motion vector of two objects and accident. The differential motion vector and accident expressed dotted line and line respectively. A sudden upward curve means variation of moving vector and accident. We can check differential motion vector before and at accident. The feature was selected as most important.
6) Direction: The direction between objects is used to classify accident types such as a broadside collision, a head-on collision, and a rear-end collision. State values of feature are discretized into eight values. This feature was detected by calculating difference of center (x, y) of an object at previous frame and current frame.

Accident Inference

In order to recognize the accident at intersections, the system should be able to distinguish situations between accident and similar situations such as traffic jamming and to be widely applied in many different environments. It is not a simple problem to recognize occurrence of a traffic accident because of complexity and diversity of the environmental factors such as the shape of roads, number of lanes, and traffic characteristics.

In this paper, in order to design a system to overcome above the problems, we analyze video data collected from intersections with different environmental factors. The video data include a head on collision, a broadside collision, a rear end collision, and a collision between a vehicle and pedestrian. The proposed system is able to cope with diverse situations at many intersections. We defined relationships between accident and feature such as the overlap of objects, the stop of objects, the velocity, differential motion vector, direction, and so on, between features, and between states of features through training of collected accident videos.

The proposed method uses extracted features in frames, but extracted features may be incomplete and uncertain because of the limit of sensors and the complexity of environment. We use the Bayesian networks which have advantages to solve problems with uncertainty to define the dependencies between features mathematically and to solve that problem with a probabilistic method. The Bayesian networks have been used to handle uncertainty in many fields such as artificial intelligence, pattern recognition, and decision making. Bayesian network is one of the best methods to construct a probabilistic network using expert knowledge and to design process of inference and expert’s decision based on cause [18]. It also visualizes probabilistic relationships between variables using DAG (directed acyclic graph) which consists of nodes and arcs [19]. The arcs are defined with conditional probability tables at each node. It effectively expresses probability relationship between nodes. We can design a network that requires to deal with the time-series information using the dynamic Bayesian networks (DBNs). As the time-series relationship between nodes was expressed by arc, we can design the relationships between nodes at current time and previous time. DBNs require more nodes and arcs than general Bayesian networks. We consider time series analysis to recognize traffic accidents accurately. So, we model dynamic Bayesian networks using current time (T) and previous time (T-1, T-2). If we would consider more time of features, the performance of accident detection could be better. However, more time increases the number of nodes and size of conditional probability tables exponentially. Thus, we use virtual nodes to improve result of inference and to reduce the complexity of the network. This method can consider more time using a virtual node which is a result of inference in previous time (Fig. 3). TABLE 1 explains process of inference using virtual node and sequence number of TABLE 1 is shown in Fig. 3.

Fig. 4 shows the hierarchical network using generic network and private network. Generic network considers only general features without direction about traffic accident. We can detect kinds of accidents using the direction of objects and the private network. The generic network and four private networks have similar structures but different conditional probability table. Our network uses reverse arcs from accident node (parent node) to feature

![Diagram of Bayesian network](image-url)
nodes (child node) to decrease the size of conditional probability table (CPT). TABLE 2 shows the comparison of CPT of the generic network by two design method. We check efficiency of used a design method in this system.

TABLE 1 Process of inference using virtual node in Fig. 3

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>①</td>
<td>Input extracted features into a probabilistic network</td>
</tr>
<tr>
<td>②</td>
<td>Calculate the rate of accidents using extracted features</td>
</tr>
<tr>
<td>③, ④</td>
<td>Save the inference result in the virtual node</td>
</tr>
<tr>
<td>⑤, ⑥</td>
<td>Input extracted features and inference result in previous time into a probabilistic network</td>
</tr>
<tr>
<td>⑦</td>
<td>Calculate the rate of accidents using extracted features and inference result in previous time</td>
</tr>
</tbody>
</table>

TABLE 2 Comparison of CPT

<table>
<thead>
<tr>
<th>Parent Node</th>
<th>Accident node</th>
<th>Feature node</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child Node</td>
<td>Feature node</td>
<td>Accident node</td>
</tr>
<tr>
<td>The number of CPTs</td>
<td>398</td>
<td>191116993</td>
</tr>
</tbody>
</table>

We have designed dynamic Bayesian networks by using SMILE engine which is a popular Bayesian networks library (http://genie.sis.pittedu). We construct a set of dynamic Bayesian networks in a hierarchical method, where DBNs are one of the most efficient models of inferring uncertain situations.

B. Accident Alarm

The result of accident inference unit was visualized to the operators using our accident alarm interface. Operators can rapidly handle accident using the proposed accident alarm system. The system was designed using C# based on .Net Framework 3.5. Fig. 5 shows the accident alarm system. Operators can check video frame at every 0.5 second and process the inference using the hierarchical networks.

Fig. 4 Hierarchical networks based on probability

Fig. 5 Accident alarm system

IV. EXPERIMENTS

A. Conditions

Video scripts were directly gathered at an intersection and used to evaluate the performance of the real-time traffic accident detection system. Because directly acquired videos include a small case of accidents for experimentation, we can acquire relatively higher detection rate (DR), correct detection rate (CDR) and lower false alarm rate (FAR).

Total 70 video scripts are used in the experiment, which contain 33 accidents, 10 situations similar to accidents, and 27 normal situations. DR, CDR, and FAR [2] are defined.

B. Result

This system as a real time accident detection should act to the accident very fast and accurately. The Result of speed of detection is 0.0069 seconds which is very fast. The experiment is conducted with different thresholds for inference of traffic accidents. If the inference result of traffic accident is above the threshold, we decide an occurrence of traffic accidents.

Fig. 6 represents ROC curve of the first experiment. When the threshold is 70, the best TPF and FPR are obtained. The ROC curve is near the 100% of TPR and near the 0% of FPR. Threshold 70 was used for the following experiments.

Fig. 6 ROC curve using frame data
This research was supported by the IT R&D program of MKE/KEIT (10033776).

ACKNOWLEDGMENT

REFERENCES


Table 5 and Fig. 7 show the comparison of the accident detection system (ARRS) [2]. The proposed system produces relatively higher DR and CDR, very low FAR. Results can be analyzed feature extraction based on a radical analysis of accident videos. We use the probability based networks and consider time-series of relationships between features. However, DR and CDR are not 100% because the system cannot detect accidents from 0.5 seconds to one second, because of result of inference in previous time using virtual node, but virtual node was more usefully used for the detection of accidents.

V. CONCLUSIONS AND FUTURE WORKS

In this paper, we present a system that prevents following accident rapidly by detecting traffic accident at intersections. The proposed system has three units, such as preprocessing unit, accident inference unit, and alarm unit. In preprocessing unit, we extracted basic information and features of objects using frames which were extracted per 0.5 seconds. In accident inference unit, we designed hierarchical probability networks using generic network and private networks. The Generic network considers only general features without the direction about traffic accidents, while private networks consider features with the direction about traffic accident. Operators can handle the accident and analyze the cause of accidents by sending the inference result using the accident alarm system to operators in alarm unit. Finally, we checked the higher evaluation result. Because we considered the time series features, states of features, and time-series of relationships between the features.

For our future work, we will consider more diverse features such as sound, road conditions and the multi-view and additional information using multi-camera and various sensors. We also plan to compare other systems.