Accident Analysis and Prevention 84 (2015) 1–8

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

How accurate is accident data in road safety research? An application of vehicle black box data regarding pedestrian-to-taxi accidents in Korea

Younshik Chung[a,*], IlJoon Chang[b]

[a] The Korea Transport Institute, Sejong 339-007, Republic of Korea
[b] Gachon University, Sungnam 461-701, Republic of Korea

ARTICLE INFO

Article history:
Received 3 March 2015
Received in revised form 9 July 2015
Accepted 3 August 2015

Keywords:
Police-recorded accident data
Vehicle black box
Accident data accuracy
Crash speed error
Spatiotemporal accident record error

ABSTRACT

Recently, the introduction of vehicle black box systems or in-vehicle video event data recorders enables the driver to use the system to collect more accurate crash information such as location, time, and situation at the pre-crash and crash moment, which can be analyzed to find the crash causal factors more accurately. This study presents the vehicle black box system in brief and its application status in Korea. Based on the crash data obtained from the vehicle black box system, this study analyzes the accuracy of the crash data collected from existing road crash data recording method, which has been recorded by police officers based on accident parties' statements or eyewitness's account. The analysis results show that the crash data observed by the existing method have an average of 84.48 m of spatial difference and standard deviation of 157.75 m as well as average 29.05 min of temporal error and standard deviation of 19.24 min. Additionally, the average and standard deviation of crash speed errors were found to be 9.03 km/h and 7.21 km/h, respectively.

1. Introduction

The process of roadway safety improvement is based on the collection and analysis of traffic accident data. In general, highly sophisticated statistical and mathematical models are applied to identify accident causal factors by using the collected accident data and its associated data such as traffic data, demographical data, and others. Thus, the validity of the modeling results highly depends on the availability and accuracy of accident data (Austin, 1995; Loo, 2006; Tarko et al., 2009). However, in most cities and countries, the collection of road traffic accident data is conducted by the police who record accident information based on estimation through the engineering methods of post-crash trajectories and skid marks at the accident spot and statements of accident-involved drivers and witnesses in the vicinity of an accident (Chung et al., 2014; Loo, 2006). Therefore, such traffic accident data, specifically crash time, crash speed, and crash location, cannot be free from fallacies as long as there is no recording device at the accident spot. Another error in accident data can be caused during the input into the computerized accident data management system. Since the accident data are typically recorded by the pen-and-paper method at the accident scene, another human error can be made by accident data cording staff.

Therefore, there have been various studies carried out to assess the accuracy of accident data collected by the police (Austin, 1995; Conche and Tight, 2006; Ibrahim and Silcock, 1992; Loo, 2006; Shinar and Treat, 1977; Shinar et al., 1983). Of these, the studies by Shinar and Treat (1977) and Shinar et al. (1983) evaluated the validity of the police-reported accident data relative to those by the multi-disciplinary accident investigation teams consisting of an accident reconstruction specialist, an automotive engineer, and a psychologist, and they found that roadway characteristics such as vertical and horizontal properties and road surface composition, and accident severity were the least reliable. Based on the questionnaire method to local authorities, Ibrahim and Silcock (1992) found that the accident location was the most inaccurate. Due to such a locational inaccuracy issue, Austin (1995) validated the accident location by using a geographic information system (GIS) validation system, and he found that the validity varied from more than 90% to less than 80% for different types of highway features. A similar study has been performed by Loo (2006), and it was found that the road names and district boards of the police-recorded accident data in 2004 contained about 12.7% and 9.7% mistakes, respectively. Unlike these studies that have tried to assess the accuracy of

* Corresponding author.
E-mail addresses: tpgist@koti.re.kr (Y. Chung), ijchang@gachon.ac.kr (I. Chang).

http://dx.doi.org/10.1016/j.aap.2015.08.001
0001-4575/

police-recorded accident data without the ground truth data such as video recordings, Conche and Tight used CCTV cameras to assess the feasibility for the accurate accident data collection with respect to the accident causal factors (Conche and Tight, 2006). Although they used CCTV cameras to increase both the quality and quantity of accident information, using an existing CCTV camera system can be limited in recording the accurate crash location, speed, and spatial coverage.

However, in-vehicle data recording devices such as a vehicle black box (VBB) have been recently introduced to be used as legal evidence regarding a traffic accident (Koo et al., 2013). This latest technology enables the collection of internal driving information such as time, GPS (global positioning system)-based location, and speed as well as the collection of video and audio data with respect to external driving situations such as crash severity, speed change, and other accident causal factors. As a result, the VBB system has a great potential in the accuracy assessment of the police-recorded accident data. Based on this background, the objective of this study is to assess the accuracy of the police-recorded accident data by comparing it to the VBB data regarding pedestrian-to-taxi accidents, specifically in terms of space, time, and speed. To perform this study, VBB data for two years collected by taxis operating in Incheon, Korea were used. Then, the police-recorded accident data were assessed based on the VBB data. Lastly, the results of the assessment are discussed, and the conclusions and future research issues are presented.

2. Vehicle black box

The black box was firstly introduced in the middle of the 1950s by an Australian scientist David Warren to investigate and preserve the cause of aircraft accidents (Suddath, 2009). He developed a prototype for a flight-memory recorder that would track basic information of an aircraft such as altitude and direction. Although airlines were using black boxes by the end of the 1950s, they became a mandatory feature in 1960, when the Federal Aviation Administration (FAA) required all commercial planes to carry them. The initial version of the aircraft black box was vulnerable to the crash impact and fire, but recent ones can maintain the status up to 100 gravity forces and up to 1000 °C temperatures.

As for the road traffic accident, accident data is also critical to understanding the cause leading to the accident, occupant kinematics and vehicle performance during a crash, and post-crash events (IEEE, 2005). Specifically, engineers are interested in road traffic accident analysis and prevention and vehicle manufactures rely on accident information to reduce accident occurrence and their severity. Therefore, the Institute of Electrical and Electronics Engineers (IEEE) launched the first universal standard for motor vehicle event data recorders (MVEDR) in 2004 (IEEE, 2005). It is a device installed in vehicles to record information related to vehicle accidents much like the aircraft black box. Therefore, it is also called an event data recorder (EDR) or vehicle black box (VBB).

However, the VBB used in this study is slightly different from the most advanced EDR. The EDR was developed for collecting, recording, storing and exporting all data related to a vehicle accident, such as engine temperature, vehicle speed and acceleration, airbag and driver belt status, vehicle identification number (VIN), anti-braking system (ABS) status, location based on a global positioning system (GPS) device, and driver's eye view video (Gabler et al., 2004). On the other hand, most common VBBs are equipped with a high resolution camera and wide lens that can capture clear, color images in the front of the vehicle and save them on memory, and they have a GPS device that provides location and time information anywhere on the Earth. Therefore, other data of EDRs are limited in common VBBs. Additionally, since the VBBs are generally installed in the inside of the vehicle windshield, when fault-related events occur or events are triggered, the data stored in memory can be easily downloaded using a computer to analyze and verify the status of the vehicle during the events. Although the price of VBBs depends on their specification such as memory size and camera image resolution, the average is about $350 including installation fee. Moreover, installation takes at most about 20 min. However, most of the time is dedicated to hiding cables. In fact, the VBB is installed by attaching it to the inside top of the windshield (or under the rear-view mirror), connecting the power through cables and then hiding the cables.

Table 1
Collectable data list by the VBB.

Information	Information type
Current time	GPS-based number
Vehicle location	GPS-based number and image
Vehicle speed	GPS-based number
Weather condition	Image
Roadway condition	Image
Roadside facility type	Image
Vehicle's control status	Voice and GPS-logged data
Vehicle type in front of the VBB-installed vehicle	Image
Vehicle maneuver in front of the VBB-installed vehicle	Image and voice
Any objects in front of the VBB-installed vehicle	Image and voice
Etc.	Image

In general, the VBB system is composed of three main components: driving information, automotive black box, user devices (see Fig. 1). Driving information is considered as generated data while the vehicle is in active mode or when any moving objects in the vicinity of the stopped/parked vehicle are detected. As shown in Fig. 1, driving information includes image, voice, GPS-based vehicle location, and speed. Based on this information, the VBB can extract various information regarding its internal and external driving environment such as time, location, speed, environmental condition, roadway facility type, and drivers' vehicle control status. Table 1 shows the list of data collected by the VBB used in this study.

The automotive black box of the VBB system includes a tri-axial accelerometer, which can be used to sense orientation, coordinate acceleration (so long as it produces g-force or a change in g-force), vibration, shock, and falling in a resistive medium. By using this device, the VBB records various shock events such as crash and even small shocks while driving on bumpy roads or over potholes. Additionally, once a VBB-installed vehicle is parked, the VBB begins with an image recording. Then, if two consecutively recorded images are different, the VBB recognizes that an event occurs. Although the VBB can record all scenes while driving as well as parked, it records the image into its internal memory from 10 s before a triggered event to 10 s after the event, to save and efficiently manage its memory capacity. Lastly, the generated data is stored in an automotive black box, then users can check the stored driving information through the view program. Due to the fact that the VBB has various important information with vivid images, it has been accepted as legal evidence regarding traffic accidents. Recently, it has been extended as a tool to find criminals. As described, the VBB can record all scenes while parked. Thus, when investigating a crime, criminal investigators often uses the VBBs of vehicles parked in the vicinity of a crime scene.

As seen in Fig. 2, accident information with respect to all of the accident scenes from pre-accident to post-accident, crash speed, accident time and location, and vehicle control status can be collected. These data are very different from the police-recorded data, which can inevitably include human errors. Based on the