March 28, 2019

Determination of Vehicle Speed from Recorded Video Using Reverse Projection Photogrammetry and File Metadata

Brandon Epstein, *Middlesex County, NJ Prosecutor's Office*
Bryce Garreth Westlake, *San Jose State University*

Determination of vehicle speed from recorded video using reverse projection photogrammetry and file metadata

Abstract

The prevalence of security and in-car video has increased the number of motor vehicle accidents captured on digital video. However, inconsistencies in how to accurately determine time and distance for vehicle speed has led to examinations with varying results. A potential solution for calculating time intervals is to use frame timing, accurate to 0.0001 seconds, contained within many digital video file metadata. This paper examines a fatal motor vehicle accident where frame timing information was used with distance measurements from reverse projection photogrammetry to calculate vehicle speed. A margin of error was then calculated based on the accuracy in performing reverse projection photogrammetry distance measurements. The resulting speed calculation was then compared to Event Data Recorder data and found to be within an average of +/- 1.43538 MPH. Using specific time intervals may lead investigators to more accurate speed calculations, specifically those involving variable frame rate video.

Keywords: Forensic science; speed calculation; forensic video analysis; frame timing; reverse projection photogrammetry

The prevalence of closed-circuit television (CCTV), or video surveillance, in crime detection and prevention has necessitated investigators to be able to accurately analyze attributes present within digital recordings. A commonly conducted analysis is to determine the speed of a recorded object. In law enforcement use of force cases, this may mean the rate/speed an object/person is moving towards/away from an officer (1). It can also be used to determine the amount of time that has elapsed between events, such as how long after a person reached for an object was a gun fired. In motor vehicle accident cases, analysis of video may be needed to determine a vehicle's speed, when an accident reconstruction was unable to be conducted or there was insufficient evidence at the scene (e.g., skid marks) to draw a conclusion (2). In determining speed within a digital video recording (DVR), it is imperative that methods for accurately calculating both the time interval and distance travelled be accurately identified by investigators.

Case Report

On July 25, 2016 a fatal accident occurred in front of the Wonder Seafood Restaurant at 1984 RT-27 in Edison, NJ. At the time no accident reconstruction was conducted, however one of the vehicles, a blue Dodge Ram, was equipped with an event data recorder (EDR) - essentially a vehicle's black box - that recorded the vehicle's speed, and other parameters, prior to impact. The restaurant's exterior security video camera also captured the collision; that recorded video was used to calculate the Dodge Ram's speed prior to impact. The Middlesex County New Jersey Prosecutor's Office (MCPO) requested that the speed of the blue Dodge pickup truck, prior to impact, be determined from the video footage of the accident. It should be noted that although the examiner was informed that investigators possessed speed information from the EDR, those speeds were not disclosed until after speed was determined from the security video.

The examiner was provided with a .MP4 video file from the Wonder Seafood Restaurant DVR by the MCPO investigator. Upon initial review of the file it was found to have a resolution of 1280x720 pixels encoded in a H.264 format. Initial metadata analysis in FFmpeg (v 3.0.6) (3) and MediaInfo (v 0.7.90) reported the video to have a frame rate of 25 frames per second (FPS). The examiner believed this to be incorrect, as 25 FPS is known to be reported in various media interrogation tools when frame rate cannot be determined. Additionally, the DVR system's settings and owner's manual both indicated that the DVR was recording at 30 FPS. A visual inspection of the video file found that the incident occurred almost directly in front of the camera, with the vehicle in question perpendicular to the camera travelling from left to right in the horizontal axis. Furthermore, the incident occurred during the day, in good lighting conditions, and without any obstruction of the vehicles.

An initial analysis of the video file was conducted to identify video frames to be utilized in the examination. This was accomplished using a visual macroblock analysis of the file generated by FFmpeg (3). The analysis was conducted to identify frames of the video prior to impact that consisted of newly encoded information in order to accurately place the Dodge Ram at a position on the roadway. A total of seven frames were identified for use in the examination (Figure 1). The frame immediately prior to impact was not utilized as it consisted of too much predictive information for the examiner to accurately determined the location of each vehicle (4).

A combination of reverse projection photogrammetry and frame timing analysis was utilized to calculate speed as distance over time. The reverse projection and frame timing analysis were viewed as two separate parts of the examination whose results would be used to calculate speed. That calculated speed would be verified against the vehicle's EDR data.

Results

Reverse Projection Distance Measurements

Reverse projection photogrammetry involves the positioning of a camera and recording in the perspective and aspect ratio duplicating the original imagery. A calibrated measuring device may then be used to complete the requested analysis (5). Reverse projection photogrammetry has been shown to effectively measure objects and distances in images (6, 7, 8). In this case, investigators returned to the scene of the incident to employ reverse projection, to determine the distance the Dodge Ram travelled prior to impact.

Upon returning to the scene, investigators found a LTS model LTD8308T-FT digital video recorder. Building management stated that the DVR settings and all associated cameras have not been changed or accessed since the initial incident. Furthermore, the system settings displayed by the DVR matched the same resolution and codec found on the initially acquired video file.

The DVR system's HDMI out signal was then connected to an Epiphan DVI2USB video grabber in order to capture live video from the system. This allowed the investigator to view the video stream as it was displayed, from the exact same camera that recorded the initial incident, using the same system settings. The live video displayed from the restaurant's DVR system was then opened within Amped FIVE forensic video software (revision 8678). Using the software's video mixer function that overlays the live video feed from the Epiphan DVI2USB video grabber over the recorded footage, it was also confirmed by visual inspection that static elements of the images (i.e. parking and lane line striping, curbs, and exterior signage) remained in the same locations; indicating that the camera had not moved since the time of the incident. The live feed from the Epiphan DVI2USB video grabber was then overlaid on still images of the seven identified pertinent frames.

Using the overlaid recorded images and the live feed from the Epiphan DVI2USB video grabber as a guide, an investigator with a reference marker was given instructions by two-way radio to mark the roadway in locations that corresponded to the same point on the vehicle's front bumper in each identified frame (Figure 2) (5). Seven marks identified as points 1-5 and impact points 1 and 2 were used to determine the distance that the vehicle travelled. Distances were then calculated between points 1-5 and impact points 1 and 2 as well as points 1-3 and points 4 and 5 (Table 1).

Frame Timing

The .MP4 video file from the Wonder Seafood Restaurant DVR was analyzed to determine specific frame timing differences. FFprobe was utilized to create a frame analysis spreadsheet consisting of the individual presentation times for each video frame (3). While the frame analysis spreadsheet consists of many different metadata values for video files, the packet presentation time (pkt_pts_time) was utilized, as it is derived from the encoding time of a frame to the 0.000001 second.

The FFprobe frame analysis derives the decode and presentation times from the video file's container, whereas the initial metadata analysis looked at the video stream or codec for this information. As decode and presentation time in the file container is an integral part of the ITU-T H.264 standard (9) it may be more reliable than any attempts to decode it within the video stream. FFprobe reports packet decode time as "pkt_dts_time" and packet presentation time as "pkt_pts_time". Packet decode time is the specific time at which the frame is intended to be decoded, whereas packet presentation time is the time at which it is meant to be displayed. In video files that contain bidirectional frames, p-frames that are displayed after b-frames are actually decoded before those b-frames as the information needed to display the b-frame is contained in the p-frame after it. In this case investigation, the examiner found a variable frame rate video, yet was able to determine elapsed time using the time each frame was intended to be decoded and presented using the FFprobe report (Table 2). The packet presentation and packet decode times were also identical as there were no bidirectional frames present in the file.

Although the initial metadata analysis using FFmpeg and MediaInfo found the file to be 25 FPS, the FFprobe spreadsheet found that the difference in frame display time varied from 0.03136 seconds to 0.06666 seconds as opposed to a constant 0.04000 seconds for 25 FPS video or 0.03333 seconds for 30 FPS video. The differences between pkt_pts_time for each frame used in the reverse projection was noted and then used to calculate the time it took for the vehicle to travel between identified frames (Table 3). For example, in Table 2, the pkt_pts_time for Point 1 was 86995.12060 and 86995.58731 for Point 4. Therefore, the time it took for the vehicle to travel from Point 1 to Point 4, noted in Table 3, was the difference between each's pkt_pts_time: 0.46671 seconds.

Speed Determination

Having identified the distances the vehicle traveled (in feet) in Table 1, and the time that it took to travel those distances (in seconds) in Table 2, speed was calculated as distance over time. That feet per second measurement was then converted to miles per hour (Table 4). The final calculated speed represents the average speed traveled between two points, not at any exact point in time. Using the calculated speed between multiple points, acceleration and deceleration can be evaluated.

Margin of Error

Due to the spatial (interframe) compression within the frames of the recorded video, as well as a potential for human error during the reverse projection distance determination, a margin of error for the calculated speed was determined. This was accomplished by calculating the physical dimensions of a pixel found approximately the same distance from the camera as the Dodge Ram. In this case, a roadway lane stripe was measured on scene and found to be 10’1” long; this same stripe was then measured in Amped FIVE to be 115 pixels. The resultant calculation found the physical length of each pixel to be $1^{3}/_{64}$ inches. Given that the selected stripe was slightly further from the camera than the Dodge Ram, the pixel measurement is actually larger than where the vehicle travelled, resulting in a more conservative margin of error.

Considering the inability to place a specific real-world point within a compressed image, a six-pixel ($6^{5}/_{16}$ -inch) margin of error was utilized in the reverse projection distance calculations. This margin of error was used to calculate a maximum and minimum speed for calculated points (Figure 4). This led to an error margin of +/- 1.10672 MPH ($p<0.05$). However, it is worth noting that because speed calculation is an average of distance over time, the greater the distance measured between two points, the smaller the margin of error becomes.

Comparison with EDR Data

After completion of all calculations, the examiner received the EDR data from the vehicle (Table 5). The calculated speed from the Dodge Ram was then compared to the acquired Bosch EDR data. The Bosch EDR records a speed at 0.1 second intervals prior to impact; it does not record distance. Using the visual display of the vehicles’ point of impact, time was counted back to correlate the EDR data with the calculated speed. Additionally, the EDR data reports speed at a specific point in time, whereas the calculated speed is an average over a distance. For this reason, the EDR speeds for the matching distances were determined and compared to the calculated speed as well. Put another way, because the EDR speed is recorded every 0.1 seconds, the closest two visual displayed values were averaged to show average speed over time. The comparison of EDR average speed and calculated vehicle speed to impact point 2 found an average difference of 1.43538 MPH (Table 5), with a margin of error of +/- 1.17114 MPH ($p<0.05$).

Discussion

Traffic accident reconstructionists and video practitioners are often tasked with calculating a vehicle's speed from a recorded video. Generally, this involves determining the time elapsed for a vehicle to travel between two points. Various methods have been used for determining distance travelled, with varying levels of precision and accuracy (10, 11). However, calculating time is a more complex endeavor because of challenges in properly identifying a video's frame rate.

When using video playback or editing applications, a common method for calculating time differences is to overlay the recorded video's timecode on the footage, to display the time of individual frames. This timecode can be derived from the video file itself or derived from settings within the software used for video playback. Due to the large number of software applications available to view and edit digital video, this means that there are countless ways in which timecodes can be derived. As a result, there are inherent issues when attempting to determine frame timing using the video timecode. Often, security digital video recorders (DVRs), and other camera systems, record at frame rates other than the standard 29.97 frames per second found in NTSC television. Many video playback and editing software applications are not equipped to properly understand this timecode, and simply default to 29.97 FPS timecode (or another frame rate) (12, 13). Therefore, the resultant timecode displayed could lead to a misinterpretation of the time elapsed between frames.

A common method for determining frame rate is to evaluate the video file's metadata. There are open-source software tools that will read the metadata of proprietary video files and report several file attributes, including frame rate. While the reported frame rate is sufficient to enable playback, it may not account for the exact timing between displayed frames, particularly with variable frame rate video files. This inaccuracy can be important when, for example, attempting to determine speed of a vehicle in a fatal accident.

Neither playback applications nor metadata frame rate evaluation accounts for the specific elapsed time between video frames. Rather they identify the number of frames displayed over time, reported as frames per second. While a video file may have a frame rate of 30 FPS, each individual frame may not be displayed at equal 0.03333 second intervals. Given the nature of DVRs, many may not be able to encode or store every frame of video, resulting in dropped frames or repeated (padded) frames. Identification of missing or repeated video frames has been addressed in previous research using visual cues to calculate frame timing in recorded video (14, 15). However, this method may not account for specific frame differences in variable frame rate video, common within CCTV video, resulting in a regular or irregular pattern of frame time differences throughout a video file. It may also lack precision due to visual limitations when examining those cues, like smaller resolutions, motion blur, and high levels of compression, common within security DVRs.

In order to determine the precise difference in time between displayed frames in video files, an analysis of each frame's presentation time can be conducted using file metadata. Denoted within the video file metadata as "packet presentation time", the presentation time of each frame is an integral part of the ITU-T H.264 encoding standard, as such it will be present in CCTV video with H.264 encoding (8). By the nature of the standard, specific timestamps for the time each frame is to be presented (displayed) is recorded in the file's metadata; the ability to decode this presentation timestamp is dependent on the file container. This packet presentation time will denote the time for when each specific frame is intended to be displayed, often to the 0.00001 second. The analysis of packet presentation time can be more accurate and precise than the existing methods as it is specific to each individual frame. The presentation time can also account for variable frame rates, dropped or padded frames, as well as small timing differences in the encoding process.

Conclusion

Understanding frame timing within video files is critical to accurately determining vehicle speed. Using FFmpeg to identify specific frame intervals from the recorded video allowed the examiner to calculate speed from variable frame rate video more accurately than using an average of frames per second. Had the examination been conducted using an incorrect frame rate of 25 or 30 FPS, the resulting speed calculation could have been as much as 13.91442 MPH incorrect (Table 6). This variance further demonstrates the need to effectively determine frame timing using packet presentation time in these examinations. The availability of EDR data in this investigation was integral in validating the calculated speed, demonstrating that the frame timing and speed measurements in this examination were accurate.

An additional method to validate these findings can be conducted by driving a vehicle at a known speed through the same scene and calculating that speed from the recorded video. This was not completed in this examination as the lowest calculated speed was still 11 MPH greater than the speed limit of the roadway. Because the roadway in question is highly travelled, and a major route to two hospitals, a decision was made to minimize any potential traffic delays by closing the road to perform this task. It is also a challenge to accurately record the known speed of the vehicle at the time it is recorded. Even with a calibrated speedometer, most vehicles have an analog speed display which makes determining exact speed troublesome. Even when recorded, it is only precise to one MPH. The use of a radar/laser speed measurement device or other GPS/GLONASS device to measure speed of the known vehicle may help increase accuracy.

There is need for additional research in determining the elapsed time between video frames using reported packet presentation times in various digital media files. This applies not only to H.264 compressed files in general, but also to different implementations of the standard. A wider study of different manufacturer's DVRs, using different frame rate setting and variable frame rates, could provide additional data to gain a better understanding of how reliable packet presentation time is.

Research on reverse projection photogrammetry and the accuracy, as compared to LiDAR crime scene scanning, could also help to improve accuracy (16). This could help to reduce the margin of error in these calculations. Additional black box studies into an examiner's ability to accurately use LiDAR crime scene scanning, reverse projection, and/or other methods of photogrammetry could also assist in determining the accuracy of vehicle speed calculations.

Additional considerations should be given to the ability of examiners to complete an examination of this nature. It is no small task to perform a reverse projection and understand the intricacies in frame timing differences. These projects are often attempted by traffic accident reconstructionists with little to no digital video training or experience. Future studies regarding the ability for personnel to conduct this examination with varying levels of training would provide insight as to the discipline or amount of training required to accurately complete the task.

References

1. State of Texas v. Derick Wiley, 2018.
2. State of New Hampshire v. Witty, 2018.
3. SWGDE. SWGDE Technical Notes on FFmpeg; https://www.swgde.org/documents/Current%20Documents/SWGDE%20Technical%20Notes%20on%20FFmpeg (accessed November 20, 2018).
4. SWGDE. SWGDE Technical Overview of Digital Video Files; https://www.swgde.org/documents/Current%20Documents/SWGDE%20Technical%20Overview%20of%20Digital%20Video%20Files (accessed November 20, 2018).
5. SWGDE. SWGDE Best Practices for the Forensic Use of Photogrammetry; https://www.swgde.org/documents/Current%20Documents/SWGDE%20Best%20Practices%20for%20the%20Forensic%20Use%20of%20Photogrammetry (accessed November 20, 2018).
6. Hoogeboom B, Alberink I, Vrijdak, D. Photogrammetry in Digital Forensics. In: Ho TS, Shujun L, editors. Handbook of Digital Forensics of Multimedia Data and Devices. Chichester, UK: Wiley, 2015; 183-218.
7. Hoogeboom B, Alberink I. Measurement Uncertainty When Estimating the Velocity of an Allegedly Speeding Vehicle from Images. J Forensic Sci 2010;55(5): 1347-51.
8. Gebruik van 3D-modellen in forensisch onderzoek [Use of 3D models in forensic investigation]. Den Haag, Nederland: Ministerie van Veiligheid en Justite, 2014 Jun.
9. Series H: Audiovisual and Multimedia Systems: Infrastructure of audiovisual services–Coding of moving video. Geneva, CH: International Telecommunications Union, 2017 Apr.

10. Lue S, Yang X, Cui J, Yin Z. A Novel Pixel-Based Method to Estimate the Instantaneous Velocity of a Vehicle from CCTV Images. J Forensic Sci 2017;62(4):1071-4.
11. Han I. Car speed estimation based on cross-ratio using video data of car-mounted camera (black box). Forensic Sci Int 2016;269:89-96.
12. Adobe. Adobe Premiere Pro CC Help; https://helpx.adobe.com/pdf/premiere_pro_reference.pdf (accessed December 15, 2018)
13. Apple. Final Cut Pro X User Guide https://manuals.info.apple.com/MANUALS/1000/MA1666/en_US/final_cut_pro_x-10.1.2-user_guide.pdf (accessed December 15, 2018).
14. Cheng YK, Wong KH, Tao CH, Tam YY, Tsang CN, Poon KC. Calibration of dashboard cameras for speed determination from video recording. Impact 2016;24(3):18-25.
15. Lower S, Stevens R, Crouch M, Cash S. Collision investigation: CCTV playback and validation using a lightboard. Impact 2017;25(3):20-7.
16. Meline K, Bruehs W. A Comparison of Reverse Projection and Laser Scanning Photogrammetry. Journal of Forensic Identification 2018;68(2):281-92.

Table 1 -Reverse Projection Distance Measurements

	Point 4	Point 5	Impact Point 1	Impact Point 2
Point 1	44’ $^{1}/_{8}$”	46’ 8½”	73’ 1½”	75’ 8”
Point 2	31’ $^{1}/_{8}$”	33’ 7”	60’ 1”	62’ 8”
Point 3	21’ 6”	24’ 2”	50’ 7½”	53’ 3½”
Point 4			29’ 1”	29’ 8”
Point 5			26’ 6”	29’ 0”

Table 2 -Pertinent frames with packet presentation time (pkt_pts_time) and packet decode time (pkt_dts_time)

frame	pkt_pts_time	pkt_dts_time
Point 1	86995.12060	86995.12060
Point 2	86995.25396	86995.25396
Point 3	86995.35396	86995.35396
Point 4	86995.58731	86995.58731
Point 5	86995.62064	86995.62064
Impact 1	86995.92064	86995.92064
Impact 2	86995.95400	86995.95400

Table 3 -Pertinent frame time differences

	Point 4	Point 5	Impact 1	Impact 2
Point 1	0.46671	0.50004	0.80004	0.83340
Point 2	0.33335	0.36668	0.66668	0.70004
Point 3	0.23335	0.26668	0.56668	0.60004
Point 4			0.33333	0.36668
Point 5			0.30000	0.33335

Table 4 -Calculated average speed

	Point 4	Point 5	Impact 1	Impact 2
Point 1	64.29495	63.68807	62.31931	61.90412
Point 2	63.42722	62.44605	61.44761	59.08752
Point 3	62.82017	61.78668	60.91098	60.55466
Point 4			59.48921	55.16177
Point 5			60.22726	59.31342

Table 5-Calculated average speed to impact point 2 with EDR speed (in MPH)

	Calculated	EDR Average	Delta
Point 1-Impact 2	61.90412	60.25	1.65412
Point 2-Impact 2	59.08752	59.571	0.48348
Point 3-Impact 2	60.55466	59	1.55466
Point 4-Impact 2	55.16177	58	2.83823
Point 5-Impact 2	59.31342	58.667	0.64642

Table 6-Average calculated speed to impact point 2 with frame rates 25 FPS and 30 FPS (in MPH)

	Calculated	25 FPS	30 FPS
Point 1-Impact 2	61.90412	61.41774	73.70865
Point 2-Impact 2	59.08752	60.82886	73.00194
Point 3-Impact 2	60.55466	60.55870	72.67771
Point 4-Impact 2	55.16177	56.18686	67.43098
Point 5-Impact 2	59.31342	61.78976	74.15512

Figure Legend

FIG. 1–*Seven frames selected from macroblock analysis*

FIG. 2–*Reverse projection roadway marking*

FIG. 3–*Average calculated speed from point 1 with margin of error (in MPH)*

FIG. 4–*Average calculated speed from point 2 with margin of error (in MPH)*

FIG. 5–*Average calculated speed from point 3 with margin of error (in MPH)*