Journal of the National Academy of Forensic Engineers®

NATIONAL ACADEMY OF FORENSIC ENGINEERS

http://www.nafe.org
ISSN: 2379-3252

Vol. 35 No. 2 December 2018

Forensic Engineering Application of the Matchmoving Process

By Richard M. Ziernicki, PhD, PE (NAFE 308F), Angelos G. Leiloglou (NAFE 956C), Taylor Spiegelberg, and Kurt Twigg

Abstract

This paper presents a methodology that uses the photogrammetric process of matchmoving for analyzing objects (vehicles, pedestrians, etc.) visible in video captured by moving cameras. Matchmoving is an established scientific process that is used to calibrate a virtual camera to "match" the movement and optic properties of the real-world camera that captured the video. High-definition 3D laser scanning technology makes it possible to accurately perform the matchmoving process and evaluate the results. Once a virtual camera is accurately calibrated, moving objects visible in the video can be tracked or matched to determine their position, orientation, path, speed, and acceleration. Specific applications of the matchmoving methodology are presented and discussed in this paper and include analysis performed on video footage from a metro bus on-board camera, police officer body-worn camera footage, and race track video footage captured by a drone. In all cases, the matchmoving process yielded highly accurate camera calibrations and allowed forensic investigators to accurately determine and evaluate the dynamics of moving objects depicted in the video.

Keywords

Matchmoving, photogrammetry, on-board video, photo-match, high-definition scanning, body-worn cameras, police cruiser cameras, drone video footage, accident reconstruction, lens distortion correction, SynthEyes, PFTrack, Boujou, forensic engineering

Introduction

Matchmoving (also referred to as "camera tracking") is a technique based upon photogrammetry, which is the science of attaining measurements from photographs or images. Accordingly, matchmoving is simply the application of photogrammetry to a sequence of individual images (i.e., video frames).

The purpose or goal of matchmoving is to take 2D information from an image sequence and solve for or "calibrate" a 3D virtual camera, which "matches" the movement and optic properties of the real-world camera that captured a given video. When done correctly, this technique allows computer-generated, 3D virtual objects to be accurately composited into the video footage with correct position, scale, and orientation.

With advancements in matchmoving software programs, high-definition laser scanning (also known as LIDAR - Light Detection and Ranging), and other related technologies, the matchmoving technique can be an effective tool for forensic engineering investigations and accident reconstruction to accurately determine and analyze the orientation, translation, velocity, and acceleration of vehicles, pedestrians, or other objects depicted in video footage.

Background

Photogrammetry (the basis of the matchmoving technique) is rooted in the principles of perspective and projective geometry, which were developed centuries ago by artists and mathematicians to transform 3D (or Euclidian) space into 2D (or projective) space (**Figure 1**). Matchmoving uses reverse projection to transform the 2D image back into 3D space by analyzing the change of perspective (parallax shift) in a sequence of images.

Before dedicated software programs for matchmoving existed, manual hand-tracking methods were used. In hand-tracking, the user makes an approximation as to the camera's position in each frame of the image sequence and then attempts to refine its position over many iterations

Richard M. Ziernicki, PhD, PE, 7185 S. Tucson Way, Centennial, CO 80112, 303-925-1900, rziernicki@knottlab.com

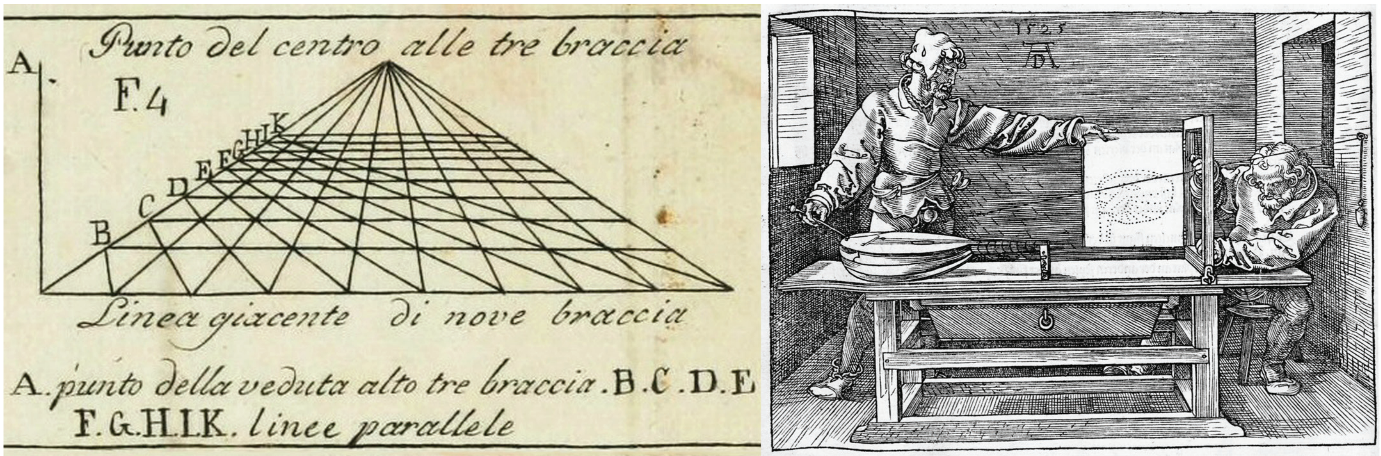

Figure 1
Left: Visual pyramid by Leon Battista Alberti, in "De Pictura," 1448;
Right: One of Albrecht Durer's perspective machines, "Underweysung der Messung," 1525.

until something close to a match is achieved. Hand-tracking methods would be very difficult, if not impossible, to implement in the matchmoving process to achieve an appropriate level of accuracy for forensic engineering. Now, with advanced matchmoving software programs and the accessibility of LIDAR, even videos with complex camera movements can be analyzed with great precision to accurately determine and evaluate the kinematics of moving objects depicted in the video.

Lens Distortion Correction

Before the matchmoving process can be performed, a major factor that must be addressed is lens distortion, which is attributed to the imperfections due to the physical characteristics of the components that make up the camera lens. The apparent effect causes the image or video to be distorted so that straight lines appear curved or bowed out toward the edges of the image. When the edges tend to bend inward, it is referred to as barrel distortion; when the edges flare outward, it is called pincushion distortion (**Figure 2**).

The amount of lens distortion can vary, but because the virtual cameras in 3D animation programs do not exhibit lens distortion, it must be corrected for accurate matchmoving and photogrammetry to be performed. Although most matchmoving programs are able to solve for and correct lens distortion, it is best to first correct the lens distortion ("undistort") the video footage using a camera calibration process.

Matchmoving software can calculate the type and amount of lens distortion in a video by using a calibration pattern or grid (**Figure 3**). A calibration grid is typically a grid of lines, points, or checkerboards. This grid can be recorded by either the same camera that shot the original video or an identical exemplar camera using the same settings that were used when the original footage was shot.

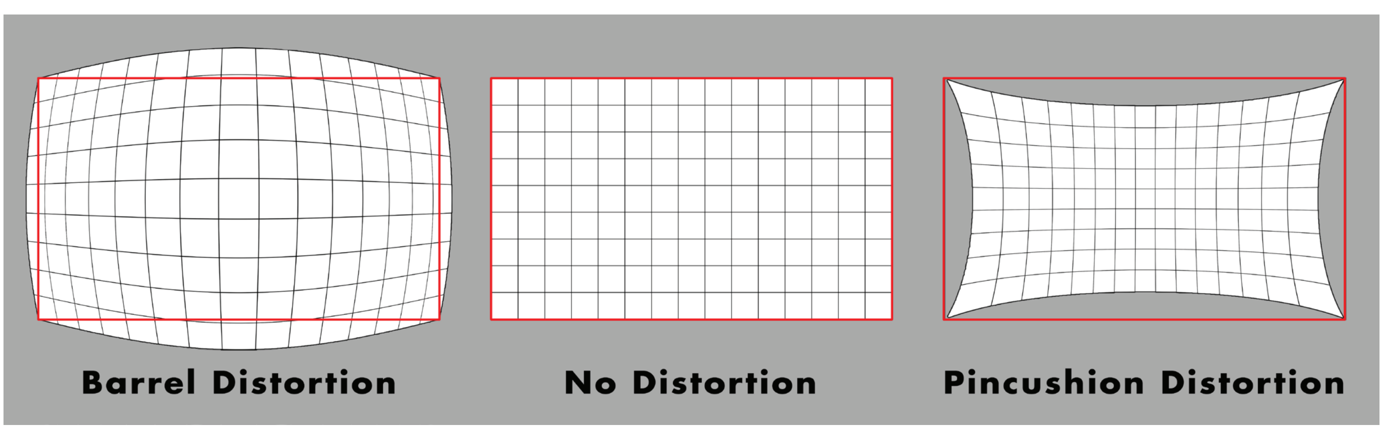

Figure 2
The two most common types of lens distortion are barrel distortion and pincushion distortion.

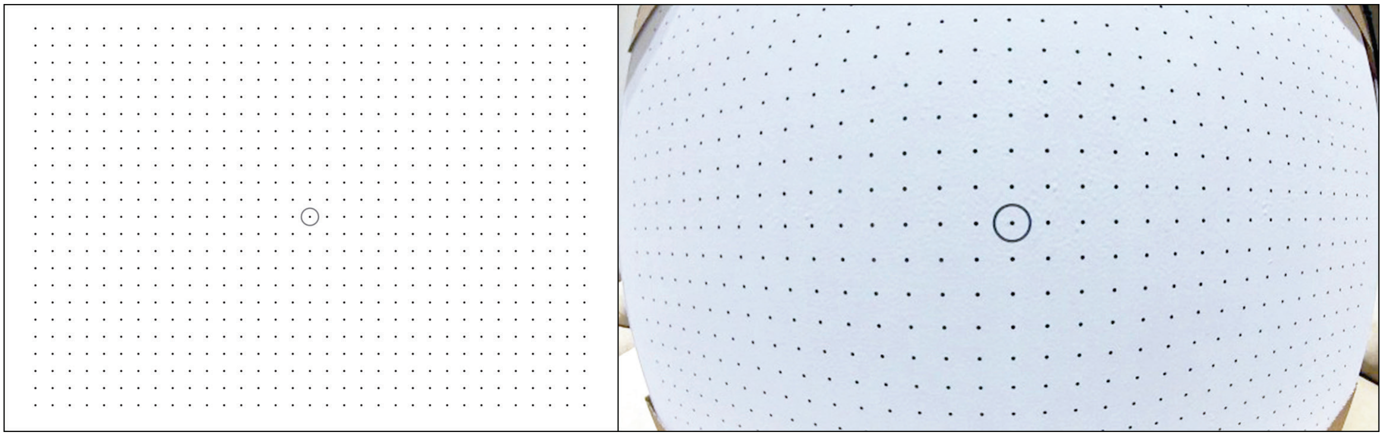

Figure 3
Left: Example of typical calibration grid used to determine the amount of lens distortion produced by a camera lens;
Right: Calibration grid capture using wide-angle lens with barrel lens distortion.

If the subject camera or an exemplar camera are not available, there are other methods or algorithms that can be used to calculate the lens distortion.

Once the matchmoving software successfully calculates the lens distortion, it then applies the correct type and amount of "warping" to "undistort" or correct the footage so that there is no longer any lens distortion (i.e., straight line in the scene appears straight in the video footage), as shown in **Figure 4**.

Matchmoving Process

Matchmoving software programs analyze 2D information (x, y) and convert it into 3D information (x, y, z) about the camera and scene. There are many matchmoving software programs available today, such as SynthEyes, PFTrack, and Boujou. While these programs can vary in some areas and features, they all generally follow the same matchmoving procedure, which can be broken down into two basic steps: 2D tracking and 3D calibration.

2D Tracking

The first step in the matchmoving process is identifying 2D points (commonly referred to as "features") in the video frames or image sequence and then tracking them throughout the image sequence using 2D trackers. Features are specific points in an image that can be easily identified (i.e., corners of objects or high-contrast spots) and represent real-world 3D objects in the scene that are static. For most matchmoving software to solve for a calibration, a minimum number of 2D features must be tracked in each frame of the image sequence. There are generally two basic methods of 2D tracking: automatic tracking and manual ("supervised") tracking. Matchmoving projects will typically require a combination of these methods.

Automatic Tracking

Most matchmoving programs now have the capability to do automatic 2D tracking, which means the software searches for and tracks features with minimal user intervention.

Figure 4
Left: Original dash camera video footage with significant barrel distortion; Right: Same frame after being corrected for lens distortion.

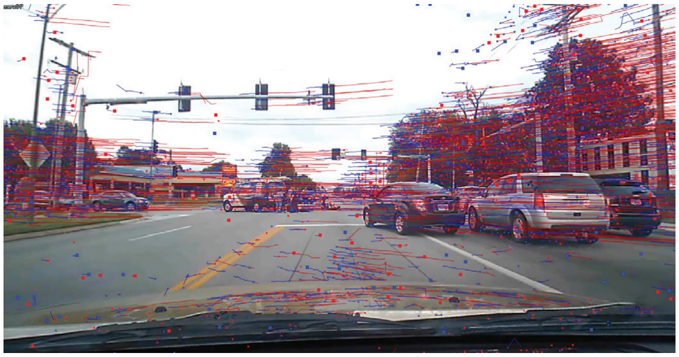

Figure 5
Blips and potential 2D tracks produced by automatic tracking method of dash camera video.

In automatic tracking, the software goes through the image sequence and identifies unique features in each individual frame and marks them as potential 2D tracking features (often referred to as blips). Then the software program tries to match or join these blips together into tracks that span several sequential frames (**Figure 5**). Finally, the software analyzes the 2D tracks to determine which tracks are valid and potentially useful for 3D calibration and eliminates those that are not.

Supervised Tracking

Supervised tracking is a "manual" method used to perform 2D tracking, only in the sense that it is the user who decides what feature he wants to track, instead of leaving it up to the software to find features to track automatically. In this method, the software uses its searching or tracking algorithms to automatically search for and track the features defined by the user, while the user "supervises" the tracking, intervening if or when the software loses track of the feature.

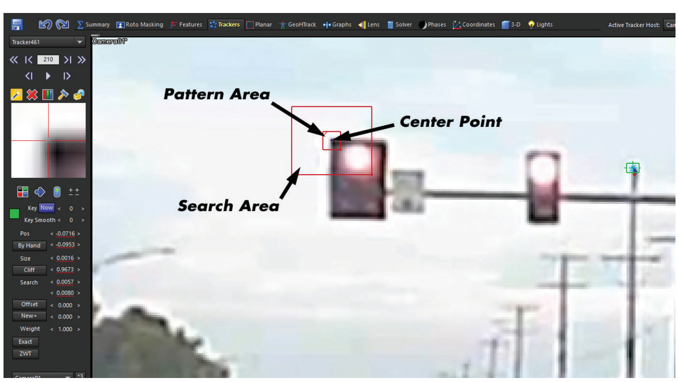

Figure 6
Using supervised tracking, the user places a tracker by defining the center point of the feature to be tracked (in this case, the upper left corner of the traffic signal), the pattern to search for, and the area in which to search for that pattern in subsequent frames.

Figure 7
2D tracks produced by manual (or "supervised") tracking. The blue/red sweeping lines represent the 2D paths of the tracked features.

To initiate the supervised tracking process, the user places a 2D tracker on a feature in a frame, defining the feature's center point. The user then defines the pattern area and search area to inform the software what pattern to search for and what part of the image it should search to find that pattern in subsequent frame (**Figure 6**).

If the software finds a similar pattern in that frame, it will automatically move the tracker's center point to match the center point of the pattern; and then the software will move on to the next frame and repeat the search process, and so on, resulting in a 2D path of that tracke feature (**Figure 7**).

If the software cannot find a similar pattern within the search area, the tracker will "slip" off-track. When this occurs, the user can go back to the frame where the tracker had slipped and "help" the software by moving the tracker to the proper position (and setting a keyframe). The user then tells the software to resume the tracking process as before.

Constrained Points

Once the 2D tracking step has been completed, the software technically has enough information to attempt to solve for or calibrate a virtual camera that matches the real-world camera that recorded the video. However, to increase the likelihood of an accurate calibration, most matchmoving software programs allow for the use of constraints — ways of forcing the software to calibrate a solution based on known 3D information. One very powerful type of constraint is using LIDAR data, which a user can use to force individual tracks to be solved to fit their corresponding real-world xyz coordinates (**Figure 8**).

Using this type of constraint assures the user that the calibration (if successful) will be accurate and in line with

Figure 8
Sample of the 3D (XYZ) coordinates data (from accident scene point cloud) used to constrain the corresponding 2D trackers.

the real-world scene as defined by the point cloud captured by the high-definition 3D laser scanning. The accuracy of the calibration can also be assessed by comparing the difference between the constrained points' positions versus the solved 3D position. The smaller the difference, the greater the accuracy of the calibration (**Figure 9)**.

3D Calibration

The second and final step in the matchmoving process (after 2D tracking) is 3D calibration. The goal of 3D calibration is to determine the exact camera movement and optic properties (e.g., field of view [FOV], focal length, lens distortion, optical center, etc.) of the real-world camera that was used to record the video of the scene, and then to reproduce a 3D virtual camera that "matches" it.

To calibrate a virtual camera, the software analyzes the 2D tracking information from the first step (2D tracking) of the matchmoving process and uses triangulation between corresponding points/features in multiple frames of the image sequence (video) to solve for the virtual camera position. In addition to generating a virtual camera, which matches the real-world camera, the calibration process also generates virtual 3D markers that represent the 3D locations of the features that were tracked in the 2D tracking step.

Evaluating Error in 3D Calibration

When a matchmoving process is complete and solved correctly, the 3D virtual camera should accurately match the real-world camera. The simplest way to evaluate how

Axes	Tracker	Locked To			Distance	Solved Position (x y z)			Error (ft)
XYZ	#298	x=133.051	y=313.03	z=33.7734		133.051	313.030	33.773	0.000
XYZ	#299	x=55.4804	y=338.583	z=32.7554		55.480	338.583	32.756	0.001
XYZ	#300	x=55.065	y=316.18	z=33.3131		55.065	316.180	33.313	0.000
XYZ	#301	x=367.287	y=256.397	z=61.4309		367.287	256.397	61.431	0.000
XYZ	#302	x=201.716	y=362.838	z=33.4569		201.716	362.838	33.457	0.000
XYZ	#303	x=144.756	y=384.769	z=50.4419		144.756	384.769	50.442	0.000
XYZ	#304	x=51.3948	y=261.215	z=25.7705		51.395	261.215	25.771	0.000
XYZ	#305	x=64.553	y=388.033	z=27.8675		64.553	388.033	27.867	0.000
XYZ	#306	x=62.5352	y=275.258	z=53.8683		62.535	275.258	53.868	0.000
XYZ	#309	x=43.5314	y=205.455	z=37.896		43.531	205.455	37.896	0.000
XYZ	#310	x=44.6667	y=197.61	z=31.6698		44.667	197.610	31.670	0.000

Figure 9
Error (far right column) of the constrained point position versus the solved or calibrated 3D positions.

accurately a calibrated virtual camera matches the real-world camera is to look through the virtual camera and evaluate the alignment between the 2D tracked features with the 3D (calibrated) markers or features (**Figure 10**). In a good calibration, the 3D markers should be aligned with the feature they represent in the image. Most match-moving software programs conveniently feature the ability to visually evaluate the error of each 3D marker's position versus the 2D tracker position in each frame of the video.

The difference between the alignment of the solved 3D marker and the 2D track is typically referred to as the "solution's error." Error values are usually expressed in pixels, which correlate to a unit of measure relative to the resolution of the video and the scale of the scene.

Once the 3D calibration step is done, the matchmoving process is complete. The virtual camera is then exported from the matchmoving software program and imported into a virtual scene within a 3D animation software program. Since the virtual camera was calibrated using constrained points from the LIDAR point cloud data, the virtual camera is accurately positioned, scaled, and oriented relative to the point cloud within the virtual scene.

In the case of body-worn cameras or cameras attached to vehicles (i.e., police cruiser camera, bus, etc.), the virtual camera's movement directly correlates to the movement of the pedestrian wearing the camera or the vehicle the camera is attached to. Therefore, the path, speed, and acceleration of the pedestrian wearing the camera or the vehicle that the camera is in can be attained from the virtual camera itself.

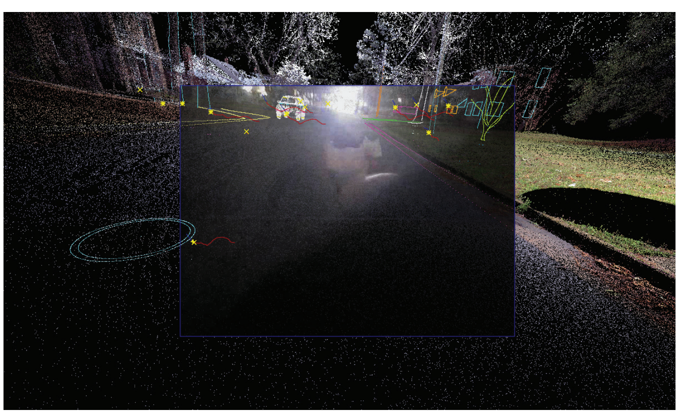

Figure 10
View through virtual camera within virtual scene with view from real-world camera (video frame) composited together to analyze the accuracy of the calibration process.

Object Tracking/Matching

The motion of vehicles, pedestrians, or other objects depicted in the video can also be determined by using a process called object tracking or object matching.

The process of object matching involves viewing the video footage through the virtual camera in the virtual scene, and manually positioning a surrogate virtual model, which has the same size and geometry as the real-world object in the video, so that it matches the object's position relative to the point cloud as depicted in each frame of the video (**Figure 11**).

Determining and Analyzing Object Dynamics

Once an object has been tracked/matched, the 3D translation (x,y,z) and orientation (roll, pitch, yaw) data of that object, for each frame, is exported from the 3D animation program and imported into an Excel spreadsheet where the object's motion data (i.e., speed, acceleration, heading angle, etc.) is calculated and graphed. The object's motion data is then evaluated to confirm that its motion is in line with the laws of physics.

Case Studies

Pedestrian vs. Bus

This case involved a female pedestrian crossing in the crosswalk with a walk signal when she was struck by a right-turning bus that failed to yield the right-of-way. The bus impact with the pedestrian can be seen in one of the four on-board video cameras (**Figure 12**). The authors were able to attain the path, speed, and acceleration of the bus using the matchmoving technique to match the on-board dash camera of the bus to the 3D point cloud of the environment (**Figure 13**). The placement of the on-board dash camera within the 3D point cloud of the bus was then aligned with the matchmoved camera to move

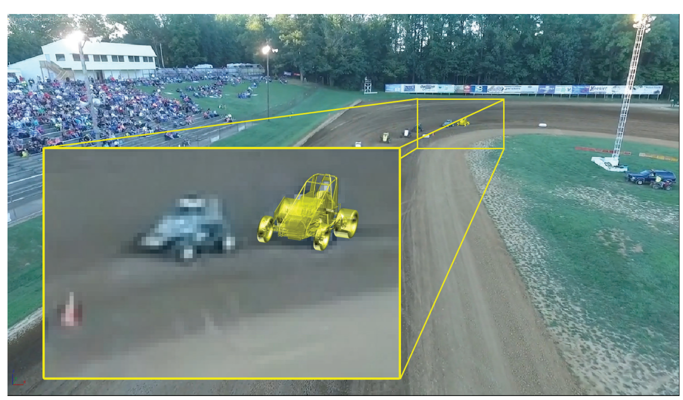

Figure 11
Virtual model of sprint car used to match the position of drifting sprint car depicted in video.

Figure 12
View of the bus's four on-board video cameras at the time of impact with the pedestrian.
Top left: Dash Cam; Top right: Step Cam — showing the impact with the pedestrian (circled in red).

a point cloud model of the bus along the appropriate motion path.

The analysis revealed that the maximum speed of the bus during the turn was 7.3 miles per hour, and the bus was traveling 6.5 miles per hour at the point of impact with the pedestrian. The relevant standard operating procedures of this commercial bus requires drivers to travel 3 to 5 miles per hour during a turn.

Further, the interior "step" camera footage was camera matched to the bus 3D point cloud that showed the passenger door of the bus (**Figure 14)**. Using contrast and color correction filters within Adobe After Effects, it was

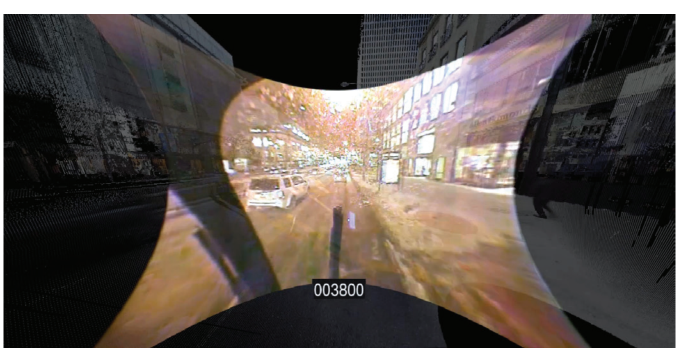

Figure 13
View through matchmoved virtual camera of the bus dash camera video footage, with point cloud in the virtual scene.

possible to view the pedestrian and other features through the bus door in the interior camera footage.

With this virtual interior "step" camera matched to the bus point cloud — and the bus point cloud parented to the virtual dash camera — it was possible to determine information regarding the motion of the pedestrian in the same 3D space, relative to the bus motion and 3D environment point cloud, as follows:

- The pedestrian, was visible to the bus driver for 6-plus seconds, but the bus had a dirty window and door. The driver failed to recognize the moving pedestrian. According to standard operating procedures, drivers are not allowed to operate a bus with dirty windows/mirrors.

- There were two signs present at the intersection: "turning vehicles YIELD to pedestrians." The bus driver failed to yield to the pedestrian.

The utilization of matchmoving revealed that the bus driver's failure to follow the standard operating procedures and posted signs at the intersection was the probable cause of the incident. Had the bus driver complied with the standard operating procedures by traveling within the designated turn speed, keeping windows/mirrors clean, and keeping proper lookout, it was opined that

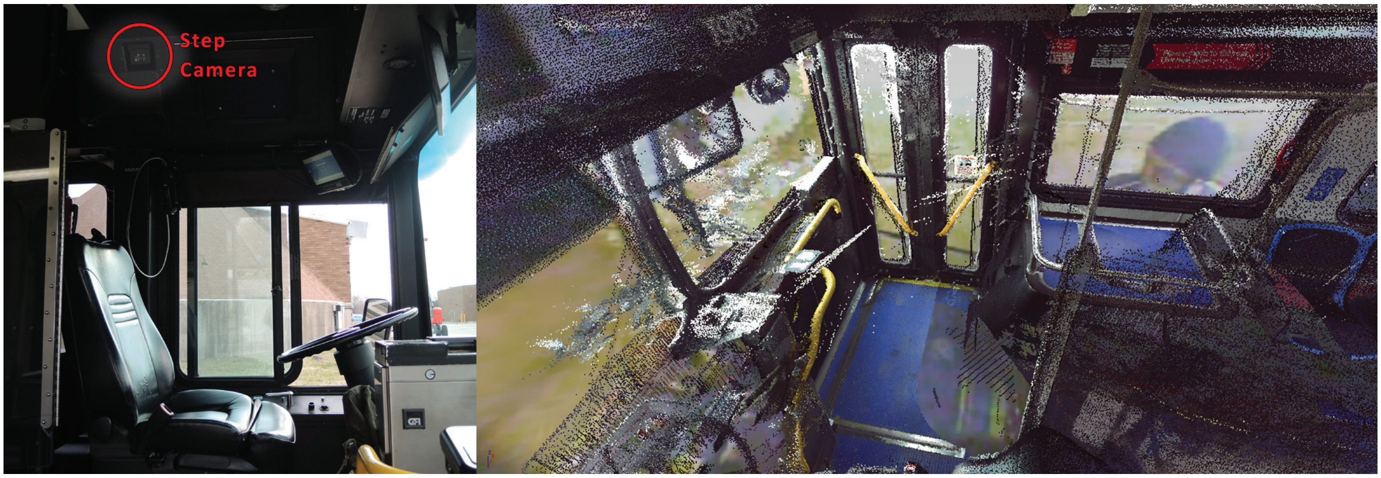

Figure 14
Left: Photograph showing location of step camera; Right: View through virtual camera match with point cloud of bus in virtual scene.

the bus driver would have been able to have ample time to observe and avoid hitting the pedestrian.

Police Officer Body-Worn Camera

This was an officer-involved shooting (OIS) case that resulted in the fatality of a 28-year-old male. The moments leading up to the shooting (and the shooting itself) were captured by the officer's body-worn camera. At the time of the shooting, the victim was sitting on the ground with his back to the officer.

By matchmoving the officer's body-worn camera and matching the pedestrian in the video, the authors were able to determine the movement of the officer (**Figure 15**) and the victim. From the analysis, the authors determined that the officer was approximately 23 feet away from the victim when he fatally shot the victim in the back and killed him (**Figure 16**).

Sprint Car Race

In this case, aerial video footage captured by a drone during a sprint car race was used to verify whether a

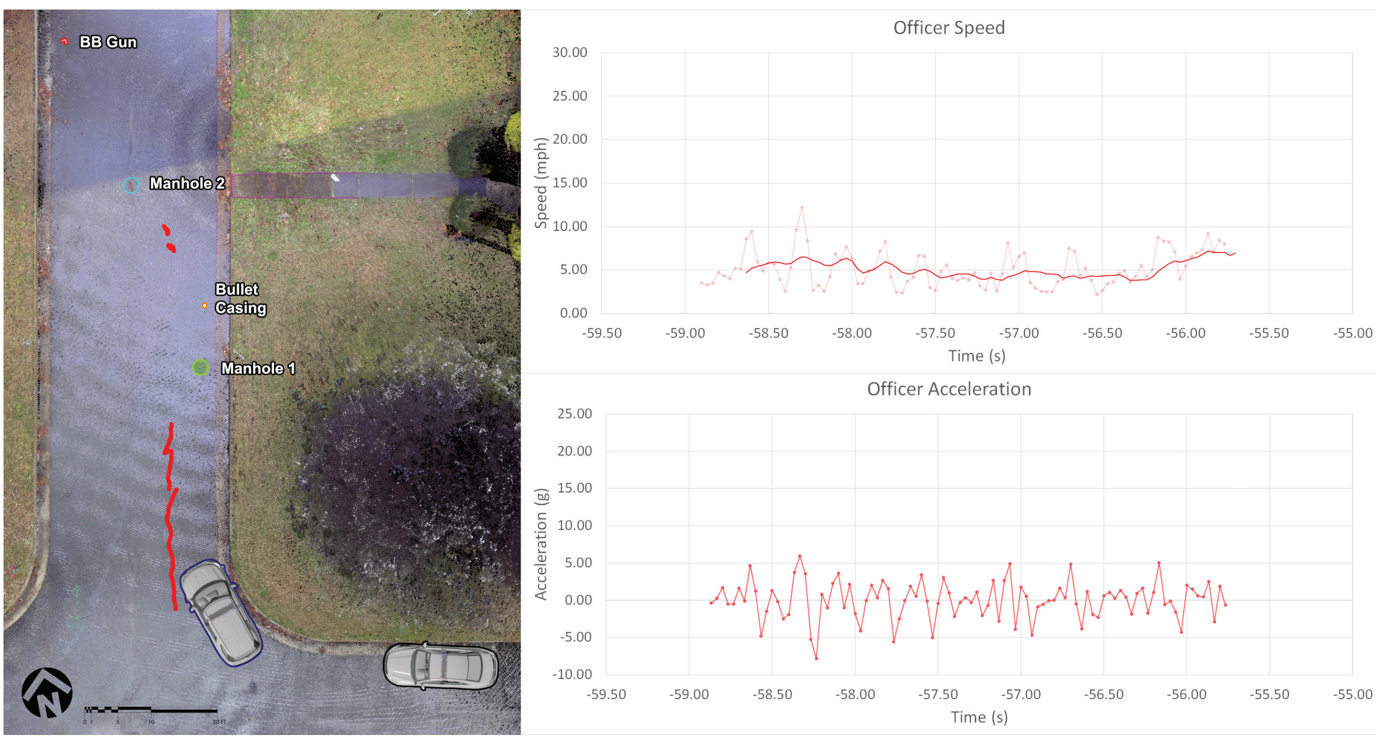

Figure 15
The path, speed, and acceleration of a calibrated virtual camera that directly correlates to the motion of the officer wearing a body camera.

sprint car could, in fact, drift at relatively low speeds. Aerial Imagery and LIDAR data attained from the United States Geological Survey (USGS) were used to accurately matchmove the video footage (**Figure 17**). Using object

Figure 16
Left: View through the virtual matchmoved camera with zoomed out to show shooting scene point cloud; Right: Top view of shooting scene point cloud depicting the location of the officer and the distance between him and the victim. * The names used are fictional.

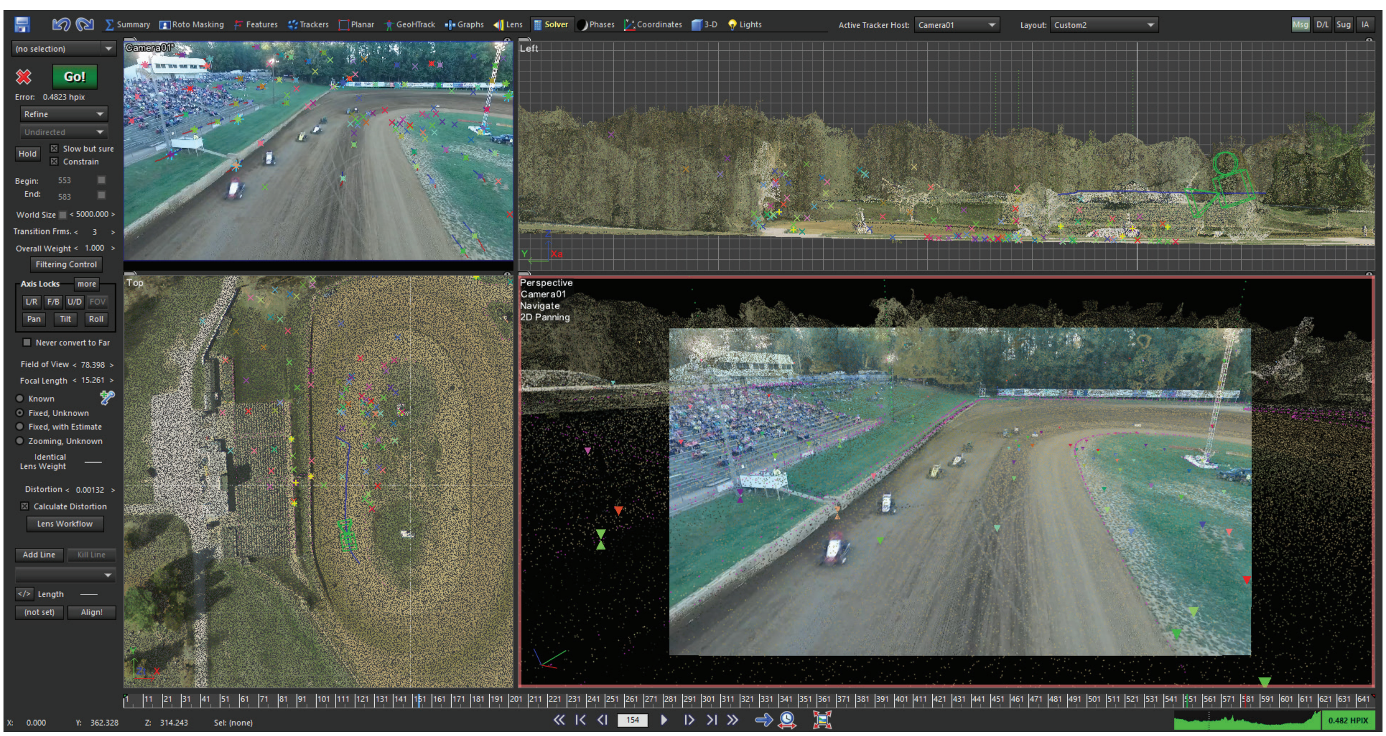

Figure 17
USGS LIDAR data used as constraint to accurately matchmove aerial video from a drone during sprint car race.

matching, the authors were able to match an exemplar sprint car model to one of the sprint cars depicted in the video (**Figure 18**). The position and rotational data of the sprint car was analyzed, and showed that the sprint car was drifting at speeds below 50 miles per hour.

Conclusion

The matchmoving process is based on the science of photogrammetry, which provides a solid foundation for forensic engineering investigations. The use of laser scanning technology (accurate to within a few millimeters) assures the accuracy and validity of the matchmoving process and the resulting analysis.

It is important to recognize that the matchmoving process must be done correctly to yield accurate results. The simplest way to verify whether this was done is to look through the virtual camera and evaluate the alignment between the 2D tracked features with the 3D (calibrated) markers. In a good calibration, the 3D markers should be aligned with the feature they represent in the image. In addition to visual verification, most matchmoving software programs conveniently feature the ability to evaluate the mathematical error of each 3D marker's position versus its corresponding 2D tracker position in each frame of the video.

With advancements in matchmoving software programs, high-definition laser scanning, and other related technologies, the matchmoving technique can be effective in forensic engineering investigations and accident reconstruction to accurately analyze video to determine the orientation, translation, velocity, and acceleration of vehicles, pedestrians, or other objects depicted in video footage.

Bibliography

1. Breen K, Anderson C. The Application of Photogrammetry to Accident Reconstruction. SAE Technical Paper 861422. Warrendale, PA: Society of Automotive Engineers; 1986.

2. Callahan M, LeBlanc B, Vreeland R, Bretting G. Close-Range Photogrammetry with Laser Scan Point Clouds. SAE Technical Paper 2012-01-0607. Warrendale, PA: Society of Automotive Engineers; 2012.

3. Coleman C, Tandy D, Colborn J, Ault N. Applying Camera Matching Methods to Laser Scanned

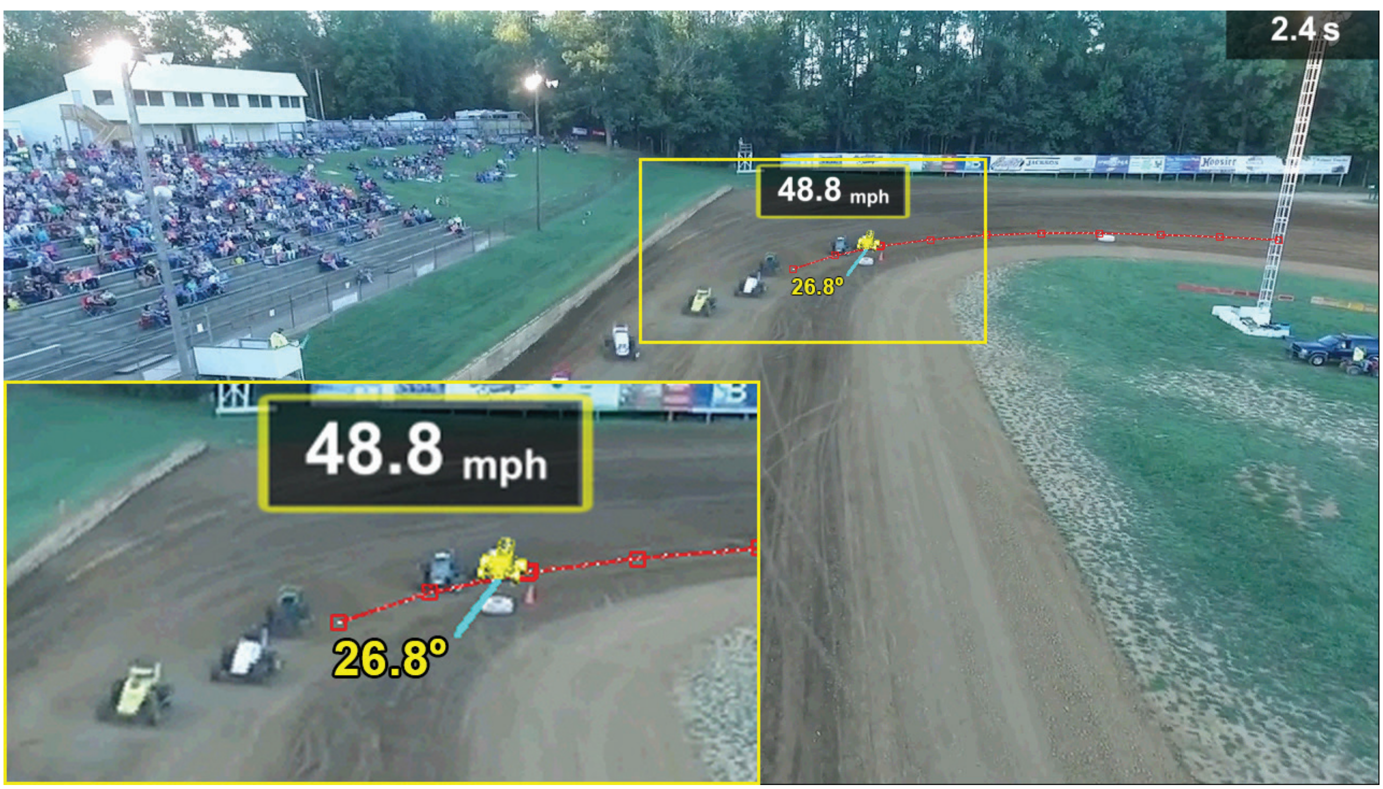

Figure 18
3D virtual model of sprint car matched to sprint car depicted in the video.

Three Dimensional Scene Data with Comparisons to Other Methods. SAE Technical Paper 2015-01-1416. Warrendale, PA: Society of Automotive Engineers; 2015.

4. Fenton S, Kerr R. Accident Scene Diagramming Using New Photogrammetric Technique. SAE Technical Paper 970944. Warrendale, PA: Society of Automotive Engineers; 1997.

5. Fenton S, Neale W, Rose N, Hughes W. Determining Crash Data Using Camera Matching Photogrammetric Technique. SAE Technical Paper 2001-01-3313. Warrendale, PA: Society of Automotive Engineers; 2001.

6. Dobbert, T. Matchmoving- The Invisible Art of Camera Tracking. 2nd Edition. Indianapolis, Indiana: John Wiley & Sons, Inc.; 2012..

7. Neale W, Fenton S, McFadden S, Rose N. A Video Tracking Photogrammetry Technique to Survey Roadways for Accident Reconstruction. SAE Technical Paper 2004-01-1221. Warrendale, PA: Society of Automotive Engineers; 2004.

8. Ziernicki R, Danaher D. Forensic Engineering Use of Computer Animations and Graphics. Journal of the National Academy of Forensic Engineers. 2006;23(2):1-9.

9. Ziernicki R, Danaher D, Ball J. Forensic Engineering Evaluation of Physical Evidence in Accident Reconstruction. Journal of the National Academy of Forensic Engineers. 2007;24(2):1-10.

10. Ziernicki R, Leiloglou A. Advanced Technologies Utilized in the Reconstruction of an Officer-Involved Shooting Incident. Journal of the National Academy of Forensic Engineers. 2017;34(2):1-12.

11. Ziernicki R, Pierce W, Leiloglou A. Forensic Engineering Usage of Surveillance Video in Accident Reconstruction. Journal of the National Academy of Forensic Engineers. 2014;31(2):9-19.

12. Ssontech.com. (2019). SynthEyes - Andersson Technologies LLC. [online] Available at: https://www.ssontech.com/ [Accessed 29 Apr. 2019].

13. The Pixel Farm Ltd. (2019). PFTrack - The Pixel Farm Ltd.. [online] Available at: https://www.thepixelfarm.co.uk/pftrack/ [Accessed 29 Apr. 2019].

14. VICON. (2019). Boujou Matchmoving Software by Vicon. [online] Available at: https://www.vicon.com/products/software/boujou [Accessed 29 Apr. 2019].

15. Adobe.com. (2019). Buy Adobe After Effects Visual effects and motion graphics software. [online] Available at: https://www.adobe.com/products/aftereffects.html [Accessed 29 Apr. 2019].