Video Analysis of Head Acceleration Events

Content of Paper presented at the March 2022 CHAMP conference.

Consensus statements discussed, revised, and voted upon.

Paper not published in the formal peer reviewed literature.

Authors:
William Neale, MArch, JS Held, wneale@jsheld.com
Ron Jadischke, Ph.D, Xenith, rjadischke@xenith.com
Steve Rundell, Ph. D, Explico Eng., steve@explico.com
Dusty Arrington, M.S., Adamson Engineering, dusty@adamson-engineering.com

The Consensus Head Acceleration Measurement Practices (CHAMP) group was founded to develop and recommend best practices for the collection, analysis, and reporting of head acceleration measurement data in sport.

CHAMP workgroups and key stakeholders convened both in-person and virtually at a consensus conference held in Philadelphia on March 24-25, 2022.

Video Analysis Of Head Acceleration Events

Four Key Statements:

- **Video analysis and Videogrammetry are critical tools to understand the position, orientation, and motion of objects observable in video, such as the head or body of a sports athlete, or the helmet and equipment positioning and orientation. Analysis of these positions and orientations can be used to calculate change in linear and rotational velocities of the head as a result of an impact event. These techniques and procedures expand beyond sports and are applicable to the measuring of motion and position of most anything observable in video.**
- **When applied properly, and validated, current video tracking methods have been shown to accurately estimate pre and post impact velocities. These estimates can be used to calculate characteristics of impact severity, such as change in velocity, using industry standard methodologies.**
- **There are several variables (e.g. frame rate, resolution) that can affect the quality of the results of video analysis. The effect these variables have on the quality of the analysis should be quantified when possible, establishing a range of certainty for the specific set of parameters used in a given analysis.**
- **Videogrammetry provides valuable input when reconstructing head acceleration events in the laboratory or when performing multi-body or finite element modeling.**
- **Videogrammetry should consider and correct, when possible, sources of potential error including lens distortion, interlacing, down sampling, compression, and variance in timestep.**

1. Introduction:

This Video Analysis section details the parameters, limitations, methodologies, uses, general practices and procedures surrounding video analysis, video enhancement, photogrammetry, videogrammetry, and video tracking. The primary focus of this section is to provide best practices for analyzing and tracking of video recordings of sporting events. However, the science, technology and principles are applicable to other industries where analysis of video is utilized. Additionally, the principles surrounding analysis of video are similar to analyzing a single photograph, since video is simply a series of single photographs played back at a particular rate, referred to as a playback rate. There are two primary components of video tracking an object which include spatial tracking and temporal tracking. When combined, these two components can allow for the full characterization of an object’s displacement and motion. Additional information derived from analysis of the video can include the orientation, height, shape, angle, trajectory, or other unique characteristics of the object or its motion. Where the object is, when it moves, how fast it travels, the rate of acceleration, the object’s size, shape, deformation, and other unique characteristics are all potentially measurable attributes that can be determined through processes of video analysis, photogrammetry, videogrammetry, and video tracking. In simple terms, the spatial tracking of objects generally involves determining the location of the object in the video relative to known markers, geometry, or other reliable references for its position, scale, or orientation. Examples of such objects being tracked are the position and orientation of an individual’s body or limb, the location and orientation of an individual’s protective equipment, or other provided equipment such as a ball, disc, or bat. Temporal tracking of an object involves measuring the elapsed time between two or more spatial measurements through accurate forensic analysis of the video files. Together, the spatial and temporal measurements associated with the tracking of an object can provide the information needed to calculate the average velocity (speed), defined by the equation: $v=\Delta d/\Delta t$, where v is velocity, d is distance and t is time. When there is sufficient data the average acceleration of an object, can also be calculated, as defined by the equation $\alpha=\Delta v/\Delta t$. However, limitations do exist when attempting to apply this in an impact event and depend on a suitable frame rate in the video. This paper details the conditions which may limit the accuracy and reliability of video analysis and the positions, orientations, and speeds that result from this analysis. This paper also describes the techniques, concepts and methodologies that can help overcome, improve, or accommodate these limitations.

2. Terminology:

The terminology section generally defines common terms used in photogrammetry, video analysis and video tracking. These terms are further defined, and their application detailed in the remaining sections of this paper.

1) Photogrammetry: The art and science of taking reliable measurements from photographs.

2) Videogrammetry: The art and science of taking reliable measurements from video, or still frames of video.

3) LiDAR (LIDAR, lidar): A portmanteau of Light and Radar. This is a commonly accepted acronym for Light, Detection and Ranging. A measurement technology using laser light to determine distance from objects, by timing the light returning from objects, similar to radar.

4) Digital Video Files: A digital file that is typically a combination of four major parts, a video stream or sequence of images/frames, an audio stream, metadata which can contain information such as the frame rate, and the type of compression, all stored in a digital container.

5) Meta Data: Additional information about a file's image and sound that is stored within the file itself.

6) Exif Data: Acronym for Exchangeable Image File Format

7) Containers and File Formats: Digital organization of data, or wrappers that contain video stream, audio stream and metadata. The video stream, audio stream, and meta data are packaged and delivered together (.avi, .mp4, .mpeg…). Data within video containers varies based on the availability and specific file format.

8) Video Streams and Audio Streams: Packets of data captured and stored in a file or transmitted on the internet.

9) File Extension: A file extension, or filename extension, is a suffix at the end of a file name following a period. The file extension will identify the file type. The common video file extensions are .mp4, .avi, .mov and common audio file extensions are .wav, .mp4 and .aac.

10) Compression: Optimization of an image, video, or audio through a computer-based algorithm to reduce file size.

11) Frame rate: Often expressed in frames per second (fps) is the rate at which consecutive images or frames are captured or displayed.

12) Interlacing: A broadcast standard where images are divided by odd and even rows of pixels referred to as upper and lower fields. Each field represents a separate point in time yet they are displayed together, effectively creating smoother motion by doubling the recording frame rate while maintaining the displayed or viewing frame rate.

13) Resolution: For digital imagery, the resolution is the number of pixels that define or make up the image. Typically measured by the horizontal and vertical axis (width and height) of the image or the sensor recording the image.

14) Pixel Aspect Ratio (PAR): The ratio of the width of a pixel to the height of a pixel. If this ratio is 1:1, the pixels are square.

15) Display Aspect Ratio (DAR): the ratio of the width to height of an image.

16) Storage Aspect Ratio (SAR): the set resolution of a file as it is stored digitally.

17) Lens Distortion: an optical aberration that results from the manner in which a lens is designed and manufactured. A common characteristic is that lines that are straight in the real world, appear bent in the image due to the curved nature of the lens

18) Rolling Shutter: An image artifact that occurs when recording fast moving objects on image sensors that record individual rows of data in sequence, rather than recording all of the data at a single point in time.

19) Global Shutter: An image sensor design where the entire image is captured at a single point in time. This design prevents rolling shutter artifacts.

20) Blur: An image artifact that can be caused by several conditions including low resolution, compression, an object moving fast in the field of view, or the camera moving fast relative to the background or object in the field of view (motion blur).

3. Overview of Video tracking and Literature Review:

Video, which is effectively a series of still images or photographs played back at a specified rate, can provide a valuable record of an event, that can be analyzed using tools that measure the position and orientation over time of objects observable in the video or photographs. Two primary ways of analyzing video include both quantitative and qualitative methods. These two methods refer not only to the analysis of video, but of a sequence of frames within video, or even a still frame or single photograph. Qualitative video analysis can refer to the review of video for purposes other than tracking an objects position to determine its temporal kinematics. In analyzing video of sports events, for instance, qualitative video analysis is beneficial in retrospectively characterizing the circumstances leading up to, and after, an event. This includes generally characterizing the motion, dynamics, location, or circumstance of a play, equipment, players body, or other element in the field of play. This type of analysis has been used in sports such as ice hockey[7], American football[14,23], 7 v 7 non-tackle football[9], Australian Football[15], and Soccer.[24] Qualitative analysis is also useful as a supplemental video analysis tool to confirm the presence or absence of impacts when acquiring data with on-field sensors.[4,22] These qualitative methods may not provide complete information regarding the pre-impact velocities, post-impact velocities or change in velocity, and hence not utilized to assess kinematics of the head or helmet but can nonetheless be useful in evaluating events based on other criteria.

Quantitative video tracking, in contrast, applies methods and techniques for accurately taking measurements of objects observable in the video, such as an object’s location, orientation, velocity or duration. Few studies have applied quantitative video tracking to assess the impact orientations, impact locations, pre-/post-impact velocities to assess the severity of impacts in sport. These studies utilize available game film for retrospective photogrammetry and videogrammetry analysis primarily to measure helmet to helmet, helmet to ground, and helmet to body impacts from the National Football League (NFL).[20,19,14,2] *Newman et al*[20]*.* utilized reverse camera projection and estimated the linear velocity of the helmet prior to and after collisions in a football game. *Lessley et al.*[14] used multibody image matching to track linear and angular positions of the helmet prior to and after impact. Others have taken a prospective approach to generate calibrated, high-quality video, with stationary, fixed lens action cameras to estimate linear helmet kinematics in youth football[6] and linear and rotational head kinematics[9,10] in 7 v 7 football. *Gyemi et al.*[6] tracked linear velocity prior to and after impact using ProAnalyst 3D and estimated positional errors as a scoring mechanism to assess the accuracy of the tracking sequences. *Jadischke et al.*[11] utilized multibody image matching to track head rotation and position in un-helmeted athletes. The best practices presented herein relate to quantitative video tracking, however, studies that are qualitative in nature may also benefit from them.

Regardless of the approach taken, there are several factors that must be considered and accounted for related to physical camera properties, recording properties, video quality and video processing to accurately perform video tracking. The limitations of the case-specific tracking methods employed and utilized in the industry should be assessed and quantified, when possible, through validation studies and/or peer reviewed research to quantify the uncertainty of the video tracking results. This uncertainty may be quantified through staging a video tracking event with known input parameters and has been executed using various approaches.[20,1,9]

Basic Description of the Primary Methods Utilized:
The use of video to determine the kinematics, including position, orientation, trajectory, and some physical properties like size and shape, of an object in space can be divided into two categories: prospective and retrospective.

Prospective Video Tracking relies on highly specialized systems with multiple, linked cameras along with sophisticated, proprietary software, Motion Capture technology, or depth mapping sensors. A critical limitation to these systems is the small amount of space that can be reasonably covered by the numerous cameras, as well as the requirement to utilize markers to track the motion of the subject. These systems use either passive retroreflective markers or active LED markers that emit light. The markers enable each camera to essentially take two-dimensional images, which are aligned to create the three-dimensional positions, i.e., triangulation. Since these systems rely heavily on the controlled reflection, or generation, of light, outdoor environments can pose some challenges. These systems are used extensively in laboratories and studios across the world and are particularly valuable in biomechanical gait analysis research. The results of these methods have been benchmarked against more traditional means of collecting kinematic data (accelerometers) and have been shown to be in good agreement.[5,13] The results from these systems can be incredibly accurate and well-suited for nuanced 3D kinematics. However, due to their high cost, spatial constraints, inability to be used outdoors, and setup time, they do not represent a practical means for determining kinematics of objects during many real-world scenarios. Data and information determined from these methods can help inform best practices associated with other techniques of object tracking from video. However, this paper will focus mostly on the use, accuracy, and limitations of marker-less motion tracking from more traditional and less purpose specific cameras, i.e., retrospective video analysis.

Retrospective Video Tracking Analysis includes video that is collected for purposes other than tracking the 3D motion of an object in space and time. Photogrammetry techniques like reverse projection camera matching can be used to track the location of the camera, whether fixed or moving. Typically, specialized software, such as SynthEyes, PFTrack, Houdini, and Nuke, can be used to determine the position and orientation of the camera and export it as a generic 3D file type. *Figure 1* depicts a sample of video that has been tracked into a three-dimensional environment, with the location of the camera that obtained the original video footage, located at each frame in the three-dimensional environment.

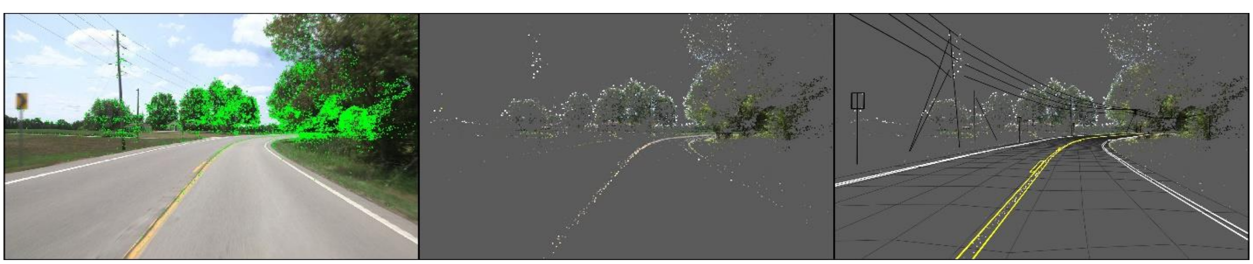

Figure 1 – Automated video tracking and resulting 3D computer geometry

3D scanning can be used to capture the 3D geometry of the scene and any corresponding objects that can serve as stationary reference markers. In the absence of a 3D scan of a subject scene or object, the captured video can be utilized to determine the relative motion of the subject camera and/or object being tracked through additional testing of the camera system.[18,16] For example, if the 3D geometry of an element in the video sequence is fully characterized, then, the relative location and orientation of the camera can be uniquely determined. The process of accurately achieving these results can be expediated if the physical characteristics of the camera and lens are known. Once the camera's position is known, then the 3D kinematics of the camera or a tracked object can be determined. The basic methodology for determining the 3D position of an object as a function of time from video footage typically requires the following steps:

1. An accurate 3D virtual environment of the scene associated with the footage
2. Tracking the position and orientation of the camera
3. Creation of a virtual camera within the 3D virtual environment
4. Creation of a 3D model of the object for which the kinematics are desired
5. Placement of the 3D model of the object into the 3D virtual environment based on the video footage

This methodology, while able to provide extremely valuable information during the reconstruction of a real-world event, can have limitations and challenges. The quality of the video footage is often less than ideal. For example, security footage can be heavily distorted and not always collected at consistent temporal intervals. The resolution of the footage may limit the ability to line-up the 3D orientation of an object. The lack of geometrical anchors, or control points, may limit the quality of the virtual camera's placement. Some or all of these limitations can be improved, fixed, or accounted for using appropriate techniques. The specifics of these limitations will be discussed in greater detail in subsequent sections of this paper.

With respect to video tracking with the intention of determining head accelerations during impact, most retrospective video analyses will not have the temporal or spatial resolution necessary to capture impact-related accelerations. Therefore, determining the input velocities provides an effective workaround in which the impact-induced accelerations can be approximated by either a) performing laboratory experiments utilizing the determined input speeds, or b) performing calculations or simulations based on a known or determined impulse duration. Ultimately, the use of retrospective video tracking analyses can be used to determine input velocities, which can be used as the basis for additional work to determine impact-induced head acceleration.

TWO CASE STUDIES OF A RETROSPECTIVE ANALYSIS

The following case study illustrates the basic concepts of the methodology described above, and a retrospective analysis of a video recording of a helmet-to-helmet collision event. In this example, the three-dimensional geometry of a field was scanned, digitized, and converted to a fully scaled three- dimensional computer model, complete with features such as field markings, that are visible in the videorecording of an event that occurred on the same field. *Figure 2* depicts the three-dimensional environment.

Figure 2 – Computer model of the field used in analyzing a sports event that was recorded with video

This computer environment was used in the camera matching process, where the video of the event was used as a background behind the computer environment, and both the video and computer model overlay were viewed from the same perspective. This example included a rig that allowed two "players" helmets to come into contact. Using photogrammetry and video analysis, the position, rotation and orientation of the helmet observed in the video was tracked at each frame. This process included a three-dimensional model of the helmet being overlayed, or tracked, such that it matches the location and orientation of the helmet observed in the video. Figure 3 depicts the process, showing the computer modeled helmet, properly tracking the position and orientation of the corresponding helmet in the video.

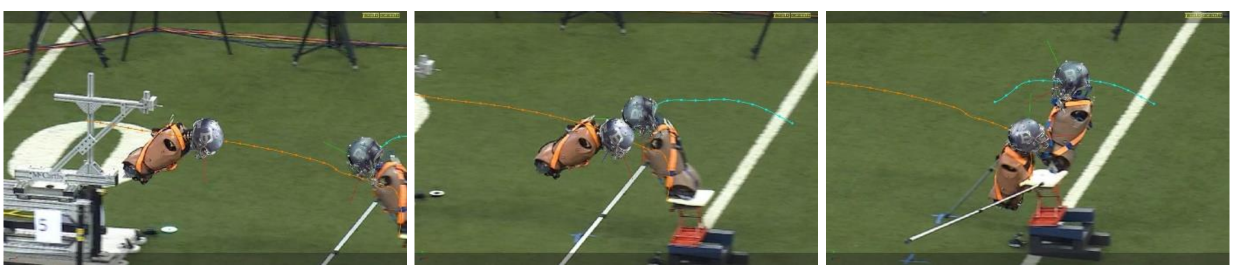

Figure 3 – Tracking of a helmet in video – the computer model of the helmet matches the position and orientation of the helmet in the video at each frame.

The results of this analysis provide accurate data of the position of the helmet and its rotation at a high frame rate. Research on this method[1] reported the results of this research, showing that not only the velocity of the helmet tracking was accurate, but also the rotation rate of the helmet.

A second example is used here, to illustrate how relevant objects in video, in this case the shadow cast by a moving object, can be analyzed to determine the objects position and velocity. In this example, a scooter and its rider experienced a single-vehicle accident resulting in injuries. Traffic camera footage was obtained from the local police department in a generic video format (Apple MPEG-4, H.264, 1440x1080) (Figure 4 and 5). The specific information related to the camera system and original file format was not provided. In some cases, this information can be determined from the metadata found within the video file itself.

Timestamps were embedded into the available video which indicated the time of each frame out to one thousandth of a second.

Figure 4 – Annotated frame of the subject video indicating the timestamp information (blue arrow) and approximate location of the subject incident (green box)


Figure 5 – Zoomed footage depicting the basic sequence of events associated with the incident

An inspection of the site was conducted. During the inspection, a combination of photographs and laser scans were collected. Data from the laser scans was used to build a simple 3D model of the location. Additionally, an exemplar scooter was also inspected, and a corresponding 3D model was constructed. During the 3D scanning of the site, the actual position of the camera was determined. Commercially available software (PFTrack) was used to precisely determine the camera's position and remove the distortion (*Figure 6*).

Figure 6: Original merged frame (left), lens distortion corrected frame (middle), and three-dimensional scan data overlaid on lens distortion corrected frame (right).

The 3D model of the scooter was placed into the scene at various positions based on the overlay with the distortion-corrected footage (Figure 7). Given the resolution of the available footage, when video tracking was performed to determine speeds, the data was found to be unreliable. Specifically, the tracked location of the scooter was predicting variations in speed between 5 and 30 mph during its approach. These values appeared grossly inconsistent with the video and predicted a maximum speed well beyond the physical capabilities of the scooter.

Figure 7: Three-dimensional model of the scooter/rider at various times based on available footage and tracked camera position

Careful investigation of the video indicated the clear presence of a shadow being cast by the rider on the nearby curb (Figure 8). The position of the sun was determined based on meteorological data for the location and time. A light source for the sun was added to the scene and the placement and location of shadows within the scene were verified form the video. The position of the shadow was determined in each frame of which it was visible and was used to refine the position of the scooter and rider. The resulting speed values, approximately between. 18 mph and 22 mph were consistent with the physical limitations of the scooter as well as the available telemetry data and provided a reasonable range for the scooter's speed prior to the accident (Figure 9). Essentially, the contrast of the shadow, combined with the fact that it was inherently constrained to the location of a known surface, i.e., the curb, provided greater precision when compared with tracking the 3D position of the scooter alone. This case study provides a good example of the general process, capabilities, and inherent limitations associated with retrospective kinematic analysis from video.

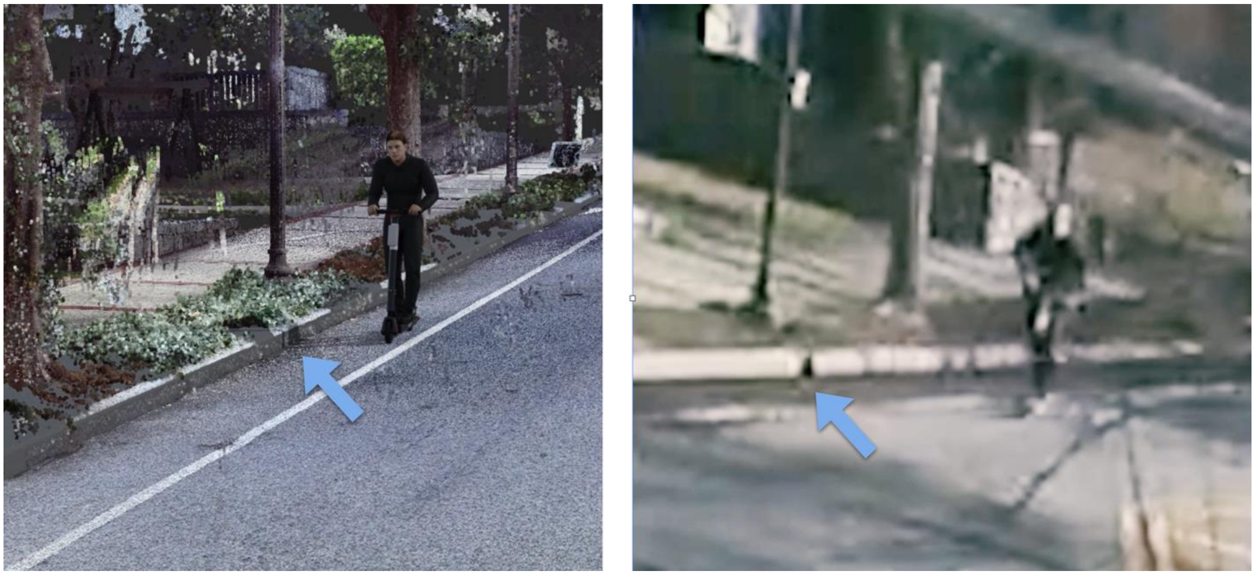

Figure 8: Presence of shadow in the 3D model based on sun position (left) as well as the presence of the shadow depicted in the available video footage (right)

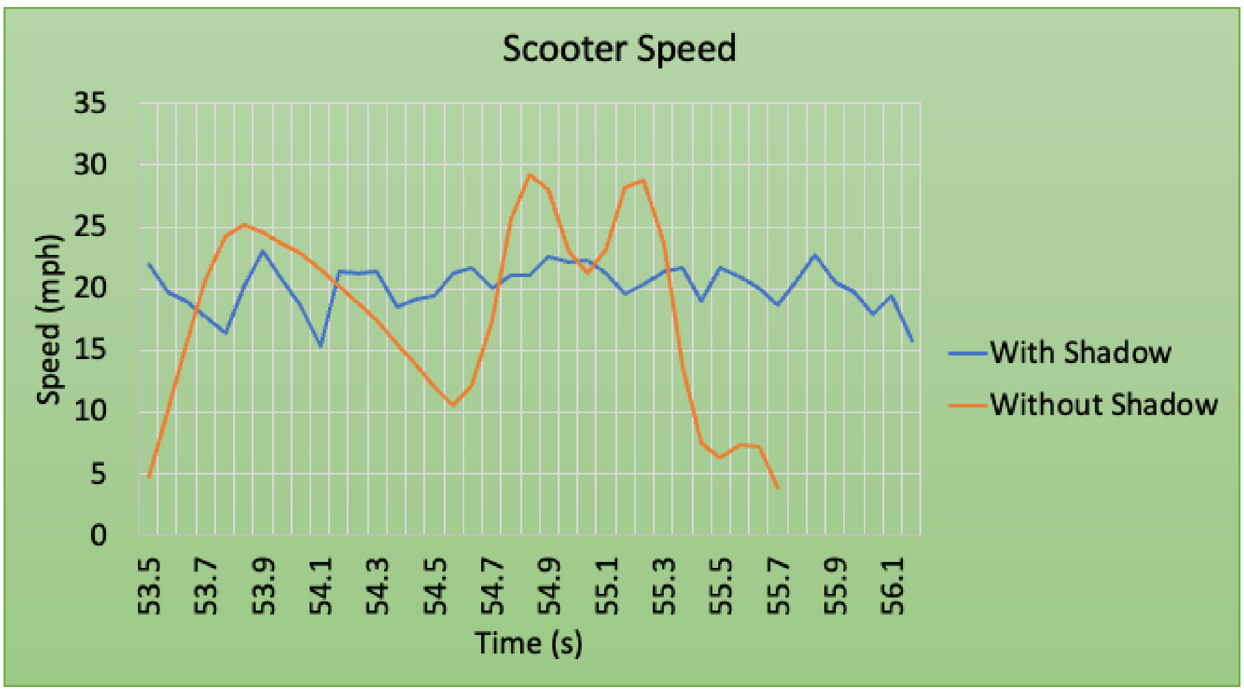

Figure 9: Line chart showing the difference in predicted speed values when utilizing the shadow (blue line) versus without (orange line)

4. Factors that affect accuracy in measuring space and time

Factors that can affect the accuracy in measuring time, distance, position, or orientation of an object that is being tracked in video or a photograph can be described, generally, in four distinct categories: physical properties, recording properties, quality of the video, post processing (Table 1).

Categories			
1) Physical Properties	**2) Recording Properties**	**3) Quality of the Video**	**4) Post Processing**
Camera locations	File format	Lens disortion	File conversion
Camera orientation	Compressions	Color distortion	Data reduction
Camera FOV (field of view)	Frame rate variability	Motion blur	Stabilization
Environmental Conditions	Proprietary viewers	Rolling shutter	Enhancement
	Resolution	Unstable footage	Color correction
	Aspect Ratio	Object to frame ratio	

Table 1 –Categories and relevant factors that influence the accuracy of video analysis

The first category is the physical properties surrounding the camera equipment and the environment. The second category addresses the characteristics and properties of the camera recording devices and the parameters for recording a digital file. The third category involves the overall quality of the footage, and inherent distortion, aberrations, or deterioration that may be present in the video file. The fourth category is how the video file is edited by a user in post-processing procedures that may result in changes to the original file. Some factors that affect accuracy can blend between these categories, as they are all related.

Category 1: Physical Properties

The location of the camera that is recording an event can vary, influencing the ability to analyze the footage depending on the goals of that analysis. For instance, a camera positioned where movement of an object is left to right across the frame may be easier to track than an object moving towards or away from the camera. Further, the height of the camera off the ground can impact the accuracy of measurements. The angle of the camera relative to the ground, that results from the camera's height off the ground is referred to as the angle of incident. Research by Terpstra et al.[31,29,28] found that the angle of incident plays a role in the accuracy of evidence placement, where a high angle of incidence had the effect of lowering the accuracy of measuring the location of the object determined through single image camera matching photogrammetry. This relationship is not linear. With multiple angles of incidence in the study, little to no improvement was shown when compared to object placement with an angle of incidence lower than 80°. The same relationship held true for two-dimensional evidence or evidence flat to the ground surface, such as tire marks, and three-dimensional objects such as vehicles. The angular accuracy for vehicles located in the study had a different relationship to angle of incidence, where the vehicle roll and pitch angles showed lower accuracy when the camera had a lower angle of incidence to the ground surface, such that it was more vertical or located above the vehicle. The overall accuracy for roadway evidence placement across all the camera matches was 1.1 in (2.8 cm) with a standard deviation of 1 in (2.5 cm). The overall accuracy for vehicle position placements across all the camera matches was 4.1 in (10.4 cm) with a standard deviation of 5.3 in (13.5 cm). The overall accuracy for vehicle orientation placement across all the camera matches was 0.5° with a standard deviation of 0.4°.

In contrast, an aerial image, or top-down camera location can be more advantageous when measuring objects flat to the ground plane. The distance a camera is located from the object of interest, the smaller that object will be in the field of view, assuming no change in the focal length of the camera lens. The smaller the object appears in the field of view, the smaller that object will be on the sensor, and thus the number of pixels representing that object in an image will likewise be smaller. Smaller objects on the image may prove more difficult to measure than if it appeared larger and was represented by more pixels.

When evaluating the number of cameras that recorded an event, in general the more coverage the cameras have, the greater opportunity there is for an analyst to measure and track an object in the video. However, even one camera can be sufficient depending on the purpose of the analysis and the range of results being reported. Research on the accuracy of single images has been explored in photogrammetry applications.[21,27] Camera matching photogrammetry with more than one camera perspective has been shown to increase the accuracy with which objects within the media can be located. This is especially true in instances where minimal corresponding landmarks exist between the media and the three-dimensional environment used to solve for the camera matching alignment. Research by Terpstra et al.[29] demonstrated this in an environment with limited landmarks. Five objects were located within a field of grass more than 150 feet away from where the photographs were taken. When locating these objects from three camera locations, a 33% increase in accuracy was achieved over placing the objects using a single photograph. Further research by Terpstra et al.[28] further demonstrated this principle in showing how accurately evidence could be located without visiting the incident site. Using a single camera to place roadway evidence and a vehicle at rest, placements were an average of 5.8 inches (15cm) from known locations with a standard deviation of 1.2 inches (3cm). When photographs taken from three different camera locations were used, placements were an average of 3.0 inches (8cm) from known locations with a standard deviation of 1.7 inches (4cm). This was also shown to be true with evidence orientation. The average orientation (all axes) of the vehicle with a single camera location was found to be 0.5° with a standard deviation of 0.03°. The average orientation (all axes) of the vehicle determined with three camera locations was found to be 0.3° with a standard deviation of 0.07°.

While one photograph, or single still frame from a video, may be sufficient for analyzing the position or orientation of an object or subject in the image, multiple images, from varying angles provides better opportunity to accurately determine the location and orientation of the subject or object being tracked. Where possible, obtaining more camera positions, and at varying positions will generally provide better results. What is more relevant, is that the goals and ultimate results of the analysis, and the accuracy reported, is commensurate with the level of data provided, the time step interval available, the number of camera views available, and quality of the footage in terms of resolution and clarity.

Bailey et al. developed techniques for tracking helmet location and orientation, comparing the accuracy of tracking the helmet from one or more cameras.[1,2] Cameras located closer to 90 degrees from each other, where their field of view is essentially perpendicular, and with higher frame rates, achieved better accuracy than single camera locations with lower frame rates. Environmental factors, such as physical obstructions, severe or low light levels, glare, reflections, and particles in the air can affect the ability for a camera to record. Where possible, avoiding or eliminating these conditions can improve the results of the analysis when performing an experiment. If analyzing stock footage, these conditions may degrade the analysis results or in extreme cases invalidate aspects of the analysis.

Category 2: Camera Recording Properties

Video Formats:

This section primarily deals with digital video files, rather than analog video since the former is the prominent current recording platform. However, analog files converted to digital formats would fall under the same criteria. Many sporting events, particularly professional and collegiate sports, are broadcast live. Broadcast cameras will send the video signal through optical cables to a coaxial broadcasting system such as satellite or cable. While the optical signal can have extremely high frame rates (240 frames per second (fps) for instance), the final broadcast signal frequency will likely be different. The optical signal with higher frames rates enables playback in slow motion. There are a several main broadcasting types: NTSC, SECAM and PAL. These acronyms are as follows:

Video Formats	Un-abbreviated	Frame Rate	Interlacing Fields
NTSC	National Television Standards Committee	29.97	525
PAL	Phase Alternating Line	25	625
SECAM	From French-Sequential Color with Memory	25	625

Table 2 – General Video Formats for Broadcast Footage

NTSC is a standard playback rate and defined as 29.97 fps. Nominally, this frame rate has evolved to 30fps for digital recordings. The 3/100th of a second offset from the nominal 30fps occurred as a result of broadcast footage playback rate needed to be based on a timing circuit (as silicon chips where not yet invented). In the US, electricity cycles at 60 times per second (60hz.) Half of that cycle sequence yield an even frame rate of 30fps. However, with the advent of color in the broadcasted footage, the color carrying signal was phasing in and out with the sound carrying signal, resulting in an undesirable out of phase experience by the user. Broadcasting standards made a small adjustment of .03fps to put the color signal and audio signal out of phase, to make it the broadcast easier to watch. This resulted in a NTSC standard playback rate of 29.97fps. Europe electricity flows at 50hz, and hence, the PAL playback rate a standard 25 fps. Film, or movies, as comparison, play at 24fps – which is effective a rate where the individual frames played in sequence appear in smooth motion. NTSC is primarily in North America, Central America, parts of South America and Japan. SECAM is in Russia and parts of Africa, and PAL is spread throughout Europe, Asia and Australia.

One of the factors that affect accuracy in analyzing video is the video's frame rate. In broadcast footage, as well as many commercially available video cameras, frame rates can vary from a fairly typical standard rate of 30 frames per second (fps) to higher rates of 240 fps, which include enough data to play back in slow motion yet still appear smooth. Error associated with using different combinations of low (60-60 Hz), intermediate (240-60 Hz) and high (240-240 Hz) frame rate scenarios for tracking helmet motion has been performed.[2,1] It was reported that resultant translational and rotational pre-impact helmet velocities were the only kinematic parameters that maintained an acceptable level of accuracy across all frame rates, with absolute errors of less than 0.4 m/s and 0.9 rad/s, respectively; estimations of peak and change in helmet velocity required at least one high frame rate camera view of the impact event for accurate measurement. Jadischke et al. utilized stationary action cameras (41 deg. FOV, 2.7 K resolution, 120 fps) within 31 m from all locations on the field and applied MBIM while tracking player heads and reported errors in change in linear speed (ΔV) of 0.24 m/s and change in rotational speed ($\Delta\omega$) of 3.4 rad/s. Figure 10 The camera was solved and calibrated by matching to a 3D Color Laser Scan of the field. Head location and rotation of a surrogate headform was matched to the athletes head in multiple camera views using PFTrack. Linear and rotational positional data was exported in a global coordinate system and resolved to a local head coordinate system to estimated headlinear and rotational velocities as a result of the head to ground impact.

Figure 10 – Example of multibody image matching

One common occurrence in broadcast footage, and with other camera devices that record in multiple fields, is that the video is interlaced. An interlaced signal contains two fields of video, referred to as Upper and Lower fields. Upper and Lower fields are horizontal rows of data captured at two different times, essentially doubling the sample rate of captured video from 30 images per second to 60 images per second. In other words, the television broadcast footage is recording at a rate of 60 images per second, each 1/60 of a second being an image composed of Upper fields or Lower fields. When the fields are combined in playback the playback rate is 30 frames per second. The benefit of interlaced footage is that using industry standard video editing software, each of the 60 images per second of video can be accessed, reviewed, and analyzed. By deinterlacing original footage, an analyst can access the information in a frame at a sample rate of 60 images each second rather than 30 images per second. If the video is not de-interlaced when analyzed, and upper and lower fields are instead combined, there can be a 50% reduction in the number of images for analysis. *Figure 11* depicts sports video that is de-interlaced to extract twice as many frames of movement. The interlaced, 30 frames per second image is on top, which is made up of two separate images at 60 frames per second shown on the bottom. In other words, the top image is a composite of the lower two images, that have been separated as lower and upper field images. As shown, if trying to measure the position of an object that is moving in interlaced frames, there will be two different possible positions, whereas the de-interlaced frames correctly provide both positions in sequence. Because the position measured may be inaccurate when only using interlaced footage, all calculations related to those measurements, including velocity for instance, will likewise be inaccurate.

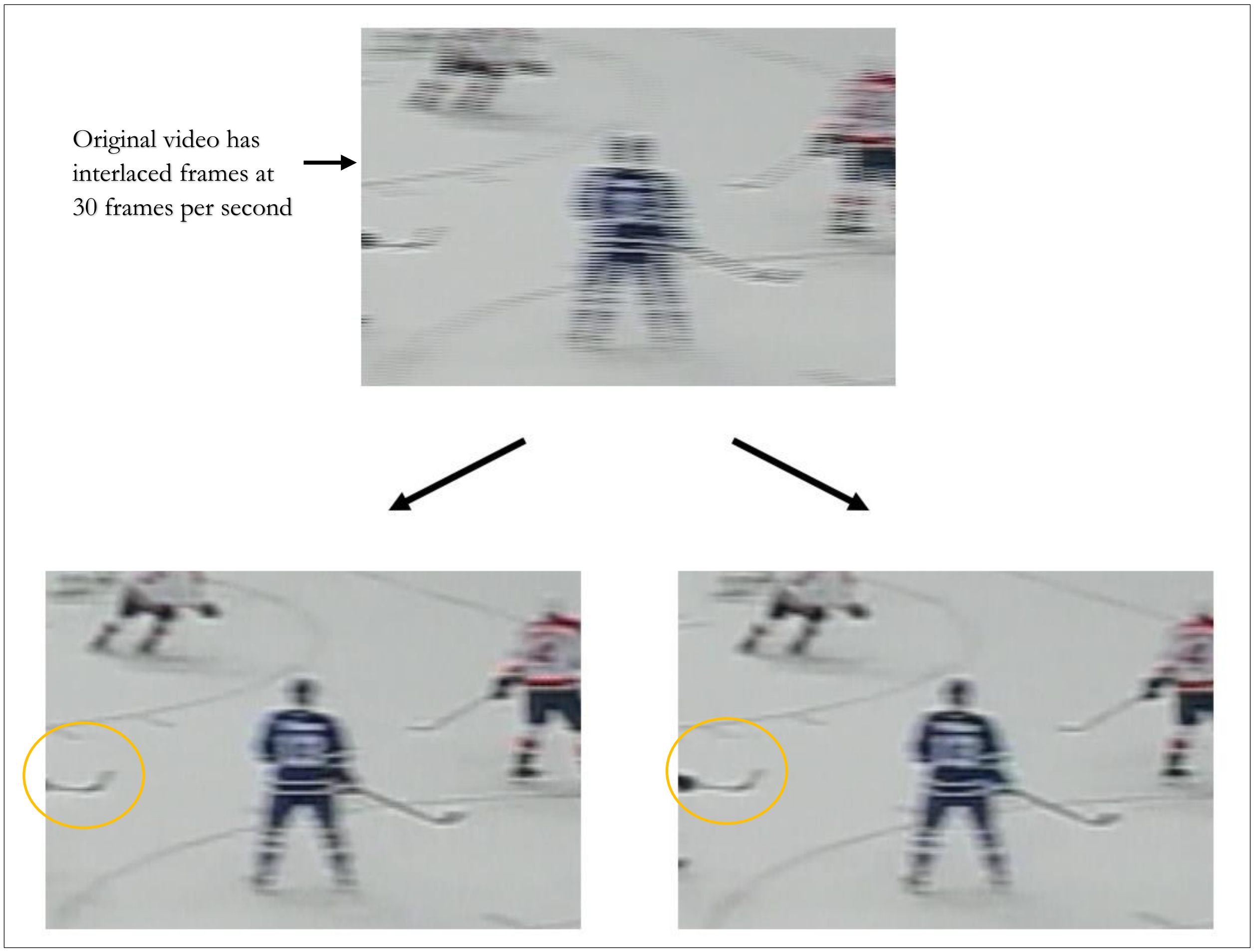

Figure 11 Example of interlaced footage

Compression:

Data Compression is used to reduce file size and save storage space. Compression is done by different software algorithms depending on the data type and exits in two different types: Loss-less and Lossy. Loss-less compression is a compression (i.e., encoding) that when it is decompressed (i.e., decoded) data is exactly the same as uncompressed file. This type of compression works by removing redundancies within the data and is most effective when similar data repeats in a file. Typical examples where loss-less compression can be utilized are text documents, source codes and image file formats like .BMP, .TIFF and .PNG, when pixel colors in a large block of pixels are identical.

In contrast, Lossy compression uses data approximation methods that generally modify original data to efficiently reduce the data size. Unlike lossless compression, the original data cannot be retrieved from a file with lossy compression as original data has been altered. Lossy compression is mostly applied to digital images and video files. Good compression algorithms try to maintain quality and lower the file size at the same time however it's usually a tradeoff between file size and image quality as higher quality means larger file size.

JPEG (Developed by Joint Photographic Experts Group) is a commonly used compressed image format for containing digital images which allows adjustments to the degree of compression. JPEG utilizes a lossy form of compression based on the discrete cosine transform. This mathematical operation expresses a sequence of data in terms of a sum of cosine functions oscillating at different frequencies. These functions effectively remove details from an image that typically is noticed by the human eye. The result is a balance between the need to keep file size low, while keeping the quality of the image as high as possible. JPEG compression is incorporated and utilized by many image processing programs and digital cameras as a standard compression method.

Compression will introduce artifacts in the images/frames of the video. Artifacts are elements/details that should be in the image but are lost or elements that are added to the image which do not exist originally. *Figure 12* depicts an example of an uncompressed and compressed image, that results in artifacts left on the image.

Figure 12. Uncompressed image (left) and same image compressed (right) with the artifacts

In digital videos, there are two types of compression that may take place. Compression within each frame (known as spatial compression) and compression over time (Temporal Compression). Spatial compression implements techniques within each frame of the video that represent patterns and repetitions in a simpler fashion. Temporal compression is a compression technique where it eliminates the need for encoding every single frame as a complete image and it only looks for changes (movements) in pixels over series of consecutive frames. Temporal compression takes advantage of the area of the image that remain unchanged frame one frame to the next frame.

MPEG (Developed by Moving Picture Experts Group) is a digital video and audio compression - decompression standard that was first introduced back in 1993 as MPEG-1. It was mainly designed to allow moving pictures and sound to be encoded into the bitrate of a Compact Disks (Video CDs) and continued to develop over time to be more efficient in compressing data. MPEG-4 Part 10, (Also known as H.264 or AVC) is the fourth generation of the video codecs after MPEG-2 and MPEG-4 Part 2, which is the most commonly used video encoding format on Blu-ray Discs, Broadcast Television, Streaming media and security camera systems. MPEG-4 Part 10 utilizes both spatial (intraframe) and temporal (interframe) compression algorithms. Spatial compression reduces the digital video file size by compressing the pixel data within each frame by technics similar to jpeg compression methods and temporal compression takes place over a series of frames (known as group of pictures-GOP) and takes advantage of areas of the image that remains unchanged from one frame to frame by throwing out data for repeated pixel data.

The compression of the video can also alter the way color appears when viewing the video. Since the image sensor of the camera is making a digital reproduction of the color in the real world, the color recorded can shift due to compression and digitization. Security cameras utilize data compression algorithms to minimize the video file size. These compression algorithms (also a type of lossy compression) use color approximation methods to eliminate the original color data to decrease the file size. Once an image (video frame) is compressed, original color data is lost and obtaining original color data from the compressed form is not feasible. What is left is likely an approximation of the true original color.

Frame Rate Variability:

Frame Rate Variability can exist in video hardware and software as a means to limit data size and reserve the unit for recording only when needed. An example of variable frame rate devices includes surveillance cameras that are motion or sound activated. A Digital Video Recorder (DVR) in a surveillance system may be set to record only when sufficient motion or sound is detected. Without motion, the recording of an image and the assigning of a time stamp to the image may be paused. While paused, the last recorded image and the time associates with it may simply repeat until motion is again detected, and a new set of images are recorded. When accompanied by a time stamp, the time stamp and single image associated with it may pause, repeating until both are replaced by a new event. As new images are written, the time code assigns a new time to those new images. In this respect, the frame rate is variable since any specific time segment, 1 seconds for instance may have 1 frame or multiple frames, depending on whether sufficient motion was detected. For some systems, when triggered by motion, the frame rate may require a warmup, and the number of frames per second that are recorded may start low, slowly increasing as the system maximizes its frame rate recording. In these instances, the first second may contain 2-3 frames per seconds, whereas the next second may yield 4-5, up to the maximum system capabilities. While this shows a variability of frame rate recording over several seconds because a time stamp exists, there is still the ability to know at least how many frames were recorded in one second interval, though the specific location within that second each event in the image shows may be less precise.

To know whether a video has a variable time rate, a visual analysis of the video is sometimes sufficient. If the time code jumps relative to real-time, for instance, this likely shows variability in the frame rate. Also, if the movement of a person or player visually appears unnatural when played back at real time, or the change in position or movement follow unnatural, or physically impossible patterns, this may indicate frame rate variability. Another method for determining frame rate variability is by importing the video file into a forensic image analysis software (such as iNPUT-ACE) and analyzing the available frames and associated embedded (not displayed) time codes.

Frame rate variability is not unusual and does not inherently make analysis of the video inaccurate, it just may make your measurement less precise. Take for example the following scenario. A video file shows that in one second of time, the number of frames being recorded changes form 3 frames in the first second, to 4 frames the next, and 5 frames in one second thereafter. This is a clear example of frame rate variability but is not necessarily unreliable, as long as the analysis being performed is commensurate with the information the available data can provide. In this example, for instance, while 3 frames may be recorded in one second, since the frame rate changes in the next second, without further testing or validation, this would not be known if those 3 frames are a constant interval within one second, or not constant. *Figure 13* is a visual example of how the 1 second of time with 3 frames could have a constant interval or a gradually increased interval. In both rows, the number of frames is 3 frames per second. However, the top row assumes the interval is constant while the second row assumes the rate is gradually increasing. Either could be correct without further validation. If further validation is not possible, the range of certainty should include both possibilities. In these cases, context of the series of analyzed images becomes important. Tracking of objects over longer periods of time as they approach the event in question can give clues through the trend in the data. Unaccounted for variance in timestep will lead to low precision in measured and calculated data.

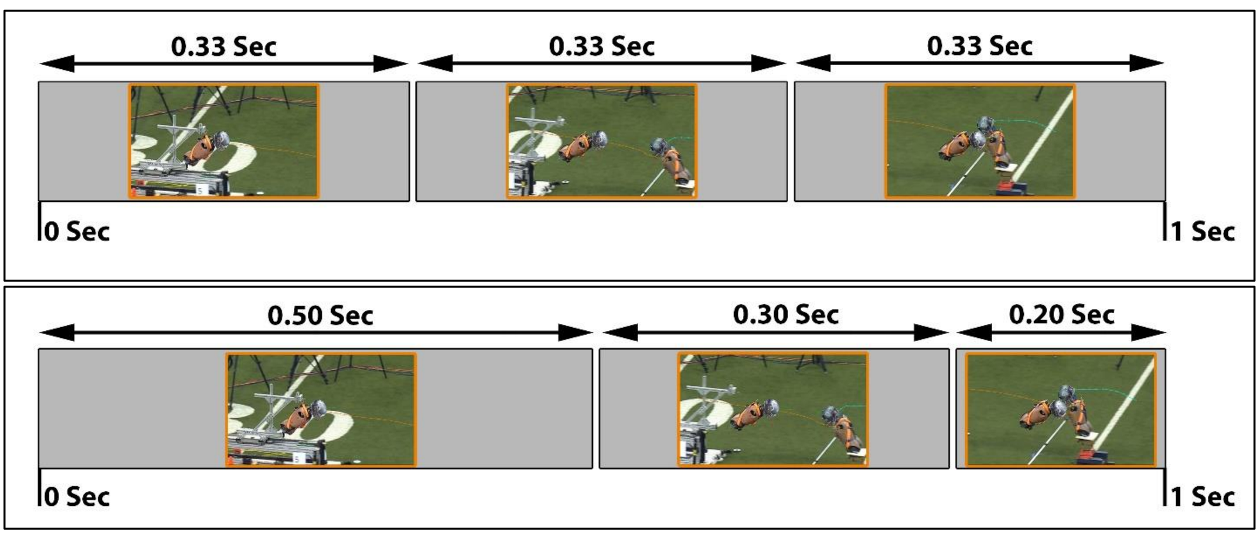

Figure 13. Example of the limits of certainty in frame rate variability

Several different methods exist for accounting for frame rate variability. These options do not apply to all conditions, and some situations exist where having more knowledge about the frame rate is simply not possible. For cameras that do not have a displayed time stamps associated with the images it has recorded, analysis of the camera itself can be performed to better understand the frame rate it records at. This includes first obtaining a substantially similar or identical digital video recorder (DVR), surveillance system, or camera set up. Research by Beauchamp, G.,[3] proposed having a digital clock or "timing light" present and in the field of view of the subject camera. In this research "The timing light is a stopwatch style clock that can be filmed so that the time of each frame (sampled duration), and time between frames (unsampled duration), can be determined. For variable frame rate camera, the deviation from an average time between frames can be determined. This deviation from the average frame rate is the uncertainty."

Using an abacus style light timing device, the accuracy of the time interval between a frame can be reliable determined, since the abacus records to 1/1000 of a second, a rate substantially more frequent than the cameras being tested. From their research, error was present when measuring distance, time and velocity for certain cameras when their frame rate was not known versus correcting for the correct time interval after testing the cameras recording rate using the abacus. For distances, the difference between corrected and uncorrected time intervals yielded a measurement range of up to 2.7 feet over a 145' distance. Time difference by approximately 1/10th of a second, and the resulting speed had a difference of 8.1% between the corrected and uncorrected. These range of certainties may not matter in all situations, such as when the overall opinions account for large discrepancies in time and velocity. However, for analysis that depends on more accurate time intervals, the variability of the camera can make a bigger difference.

Analysis of the variable frame rate video may be difficult in its native format, especially with a proprietary video player. Converting the file to a more useful format may be required and care should be taken to make sure frames are not excluded or omitted, and the time stamps do not shift from their original associated frames, unless a clear reason to do so is established. One method seeks to play back the video sequence in real time, by using the time stamp stored within the video file and processing the video frames so that video plays back at an interval consistent with the real-world time elapsed during the recording. As long as the frames are associated to the time code, the events in the video frames will likewise be played back at real time. This can be performed by forensic image analysis software program if the time code is attached to each frame. These programs can then generate copies of the files that will play properly on a standard video player at the proper frame rate without affecting the quality of the recorded image. Again, the frame in post processing can be assembled and play back at real time. In this instance, where segments of the video occurred that show the recording paused, (i.e., where the recording essentially freezes on an image and time is paused), the image can be extended in the post processing software to cover the duration of paused time until the next image and time stamp appeared. In other words, each frame recorded in the video can be utilized and its time code used to determine if play back is at real time.

Resolutions:

Resolution is the number of pixels that define the image size. The number of pixels vertically and the number of pixels horizontally are its resolution. *Figure 14* depicts the common resolution types from the lowest resolution to current highest resolution of 8192 x 4320 known as True 8K.

	Resolution	Pixel Count
CGA	320 x 200	64,000
QVGA	320 x 240	76,800
VGA	640 x 480	307,200
NTSC	720 x 480	345,600
PAL	720 x 576	414,720
SVGA	800 x 600	480,000
HD	1280 x 720	921,600
SXGA	1280 x 1024	1,310,720
FHD/1080P	1920 x 1080	2,073,600
UHD-4K	3840 x 2160	8,294,400
True 4K	4096 x 2160	8,847,360
UHD-8K	7680 x 4320	33,177,600
True 8K	8192 x 4320	35,389,440

Figure 14 – Common resolution types

Display, Storage and Pixel Aspect Ratios:
The ratio of width to height (in terms of number of pixels) of an image is known as display aspect ratio (DAR). For older Televisions, monitors, cameras and other display devices, the 4:3 ratio was common and referred to as Full Screen, since the image filled the fully available width and height of the display device. This was developed to include a 16:9 aspect ratio which was referred to as Widescreen, and standard for HDTV. This format favored the wide recorded view in film, and maximized the width of the display, sometimes leaving the top and bottom of the display black. Aside from display aspect ratios, when digital images or videos are recorded as a file, they are stored at a set resolution known as Storage Aspect Ratio (SAR). A video that is recorded at 800 pixels x 600 pixels resolution, for example, has a 4:3 SAR. If this video file is then displayed at a monitor with DAR of 4:3 that means pixels are square or have a Pixel Aspect Ratio (PAR) of 1:1 (PAR=1:1). The relationship between these three parameters can be defined as: DAR=PAR x SAR.

The standard playback process of a video file is first to look for the DAR value stored in the container and according to that, it will playback the video using square pixels, if the video file does not have the DAR information, it will use Storage aspect ratio (SAR). Some playback devices have different shaped pixels than others. Computer monitors have square pixels and hence everything that is designed for display on a monitor should have a PAR of 1:1. Some televisions, however, have rectangular pixels, where they are wider than they are tall, which have a different PAR depending on the format (NTSC or PAL). This can result in a square image being displayed incorrectly by 10% and stretched vertically.

Figure 15 is an example of the effect of a square pixel aspect ratio (PAR 1:1) being represented a .9 pixel aspect ratio display (PAR 1:.9). If the image that is being stretched is not accounted for, errors in measuring distance can occur, since the size and shape of all the pixels have been distorted. Some video formats (For instance Avi) do not hold Pixel Aspect ratio (PAR) information. When these video files are encoded, the assumption of square pixels might not be accurate, and the encoded video will not have the correct resolution/aspect ratio. Error occurs not only in measuring an uncorrected PAR image, on the order of 10%, but attempts to use the uncorrected image in photogrammetric analysis such as camera matching, will compound, since the image and the three-dimensional scene to align it to will be mis-matched. Some camera matching software packages such as PFTRACK can solve for and account for errors in pixel aspect ratios as part of the camera matching or camera solving process. If utilizing point cloud data, the program will solve for a correction factor when performing the camera matching process. If tracking based solely on the imagery, then additional testing using stadia boards or other standard fields of reference can be performed to experimentally determine the correction factor. The pixel aspect ratio, and how its displayed can be accounted for by matching the pixel aspect ratio to the native display ratio of the device it is being viewed on.

Figure 15 – square PAR stretched vertically (left), .9 PAR corrected so objects appear unstretched (right)

In this example, when taking a measurement of the tire in the uncorrected image, the height is approximately 3.4 inches taller than the actual diameter of the tire. The difference, which correlates to the pixel aspect ratio, has an error of approximately 10%.

<u>Category 3: Quality of the video:</u>

Lens distortion:
Lens distortion is present in all photographic and video images due to the manufacturing process of the lenses. Lens distortion is evidenced by videos and images where lines, that should appear straight, appear bent through pin cushioning and barreling – common effects of lens distortion. Research by Neale et al.,[17] lens distortion can "shift the location of the image on the pixel matrix, and hence shift the position, size and shape of the geometry the pixel represents. As a result, when measuring a distorted image, the size, shape and position of an object of interest may be misrepresented." This paper showed that for common lens distortions, the error between the distances measured versus the actual known distances could be as high as 23%. Other research[32] reported similar error when measuring distances in an uncorrected image. *Figure 16* is an example of the effects of distortion, along with a corrected image.

Figure 16 – Comparison of an image without correction (left) and corrected (right)

Typical lens distortion types include barrel and pin cushion, named for the general appearance or characteristic of the distorted image. Distortion in images can be removed using software designed to both analyze the distortion of the image and correct the image to remove it. Adobe After Effects, Adobe

Photoshop, Nuke, and PFTrack are just some of the available software titles to do this. Published research on photogrammetric processes highlight the inherent distortion in cameras and the need to correct for distortion to avoid inaccuracies in measurement. Terpstra et al found that “Camera lenses are curved in nature and introduce varying degrees of distortion in the resulting photographs. This lens distortion has been shown to have an impact on the accuracy of photogrammetric solutions”. Specifically, the distortion associated with lenses can “cause errors to be introduced when photogrammetric techniques are used to analyze photographs” when determining position, scale, length and other characteristics of objects in the photograph or video image. Where possible, lens distortion needs to be corrected before measuring objects and distances to reduce error. In general, the center of the image contains the least amount of distortion. But as distortion is measured away from the center, the distortion increases. In general, the lower quality of the lens and the higher the field of view of the lens the larger the amount of distortion. As noted in research, care should be taken when measuring the corners of distorted images, as these areas are most prone to the effects of lens distortion.

Color Distortion:
Not only does each camera record the same color differently[12] since each manufacturer designs their own light sensitive sensors, each monitor that displays the colors will represent that color differently. Without knowing what color correction is needed to display colors correctly, a viewer would not know how different the color on their monitor is from that color in the real world. The image sensor of the camera converts an optical image into electronic signal. This conversion is different from one camera to another as it depends on many different factors including the type of sensor, lens, digitization algorithm and hardware used. These differences in the image sensor are the reason why captured colors look different across different devices. *Figure 17* depicts the difference that two cameras can have, in representing real world colors, even when taken from the same location at the same time. Note how the sky looks blue in one image but grey in the other.

Figure 17 – Two different security cameras capture varying color at the same time

The sensors of a digital camera have limited range for recording bright colors. Bright colors can appear as white in the video even when they are not white in the real world. This limitation effect is called clipping which is a result of capturing or processing an image where the intensity in certain areas falls outside the minimum and maximum intensity range which can be recorded by the sensor. The clipped area of the image will typically appear as a uniform area of the minimum or maximum brightness. *Figure 18* depicts the clipping effect in one frame of the video. In this image, the brightness of the explosion is clipped, the colors being recorded on the sensor are outside of the sensor’s range, and thus all appear uniform and white. In the real world, the colors of the flame may be distinguishable, even if the video shows them as white.

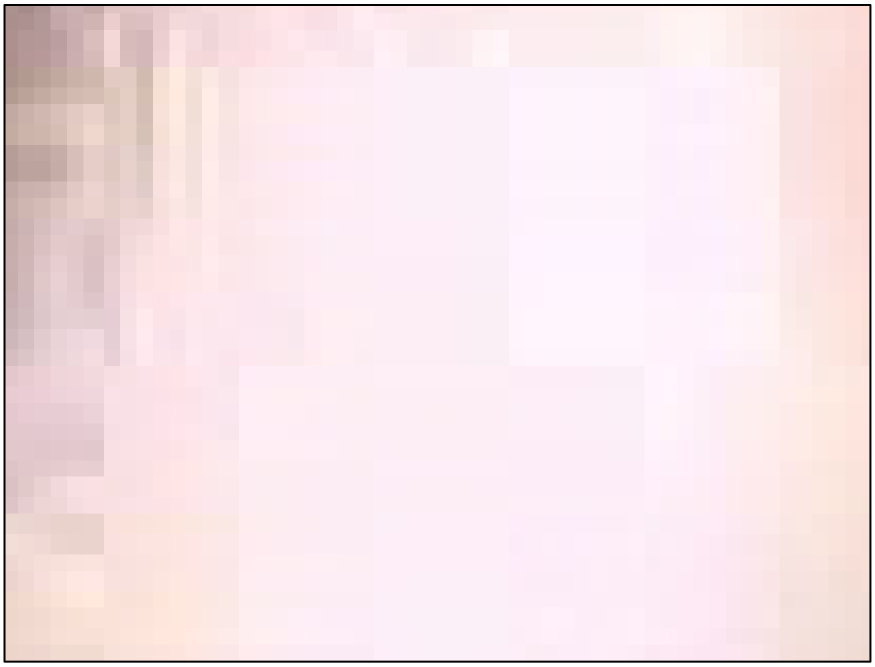

Figure 18 – Clipping effect due to the limited range of a sensor recording bright colors

Color can also be affected by the exposure of the camera settings. Exposure settings may be unknown, they may be automatic, or manually adjusted, and may adapt to the daylight conditions, which can render colors darker or lighter than they would appear in the real world. Different exposure setting affects the color intensities presented in the video frames. *Figure 19* depicts the effect of different exposure settings on a single frame and how it affects the colors. Different exposures can make yellow look orange, and vice versa. Thus, judging the color of the flame as orange or yellow may require knowledge or testing of the sensors and exposure of the cameras.

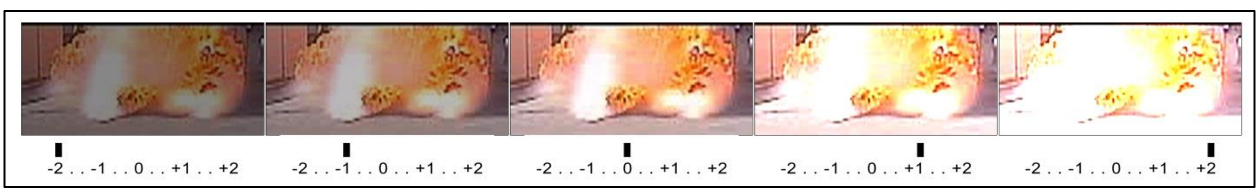

Figure 19 – Different exposure setting affect the colors

Motion Blur

Blur can be caused by the compression of the video through an averaging of pixel color which can make the recorded image differ from the real-world. Blurring is a result of loss of high spatial frequency image details due to compression where group of pixels (Picture Element) known as Macroblocks are assigned an averaged color within the macroblock. *Figure 20* depicts the Blurring effect within the macroblocks.

Figure 20 – Blurring effect in macroblocks

Blur can also be caused through motion, from either the camera or the object. The faster the object the more the blur, in general. At night, lights can streak across an image, also an example of blur. *Figure 17* is an example of blur occurring due to the object passing the camera too fast. Notice the clarity of the background but not the object. When tracking an object that has motion blur, care must be taken to make sure the point of the object being tracked is consistent between frames. In *Figure 21*, for instance, tracking the front of the vehicle in one frame and the rear of the vehicle in another will yield unreliable distances.

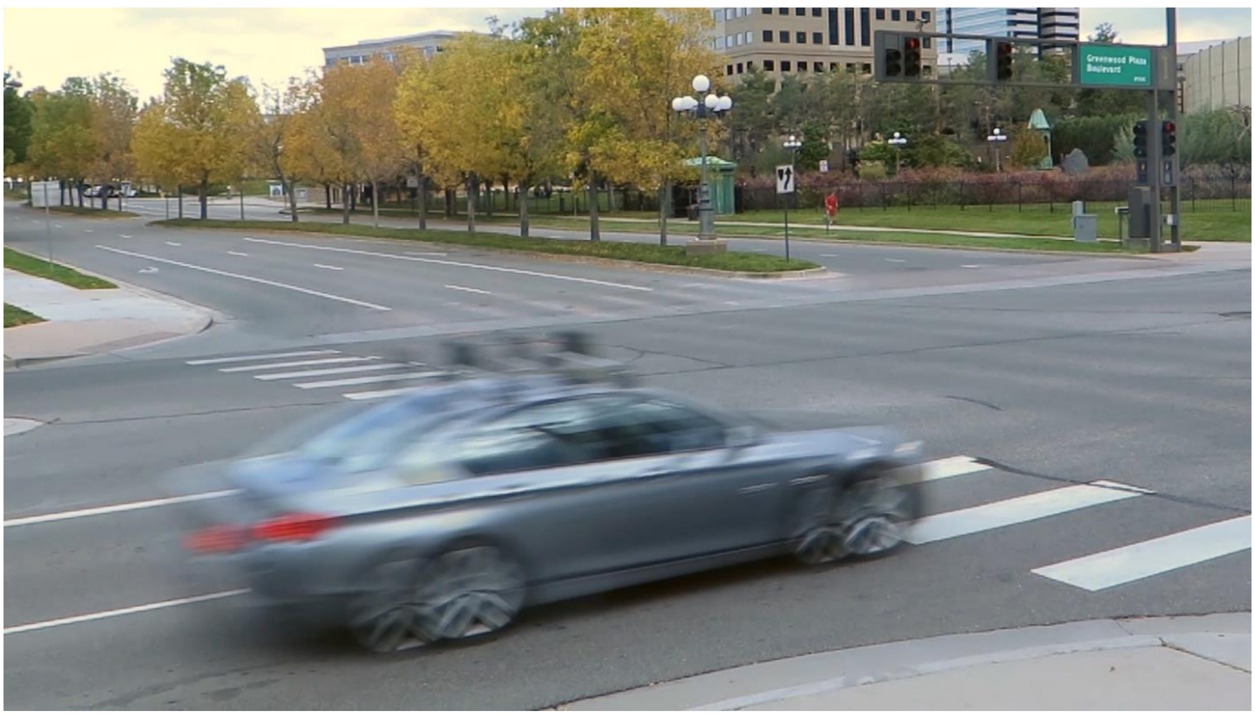

Figure 21 – Sample of motion blur

Rolling Shutter:
Rolling shutter is caused when a camera sensor does not capture the top and bottom of the image at the same time, but rather "rolls" from the top corner to the bottom corner as it stores data for a single frame. For very fast-moving cameras or objects, by the time the sensor has started recording data in one corner and arrives at the end of the image for that split second, the image may have moved and results in a distortion of the image. *Figure 22* is one example where the blade of the propeller spins fast enough that the sensor cannot pick up an accurate representation of the image in one frame, and thus it appears bent instead of straight.

Figure 22 – Sample of rolling shutter, straight objects may appear curved

Unstable Footage:

Another potential source of error when analyzing video occurs when the video is unstable. Some examples of unstable footage include footage that is hand held, with unpredictable panning, zooming and movement. Body Worn Video (BWV) is an example of unstable footage when an officer is running, moving or shifting quickly, or physically engaged with another person.

Another example of unstable footage is video recorded of another video. This occurs when someone records video that is playing on another screen. Captured video of a video can have compounding problems. The frame rate of the source video, the playback rate of the video on the external screen, the refresh rate of the screen upon which the video is being displayed, and the frame rate of the recording by the new camera can all be different. Additionally, there is a compounding error of lens distortion, since the display of the source video can be showing lens distortion, and the new camera recording can have its own lens distortion. Correcting for both of these can prove complicated. For all these reasons, care should be taken when analyzing video footage that is the recording of another screen. *Figure 23* is a sample of video recording of another video that was already recorded, being displayed on a screen.

Figure 23 – Video of another video

<u>Category 4: Post Processing of the video footage:</u>

File Conversion:
Digital video files are generally stored in a compressed format. The main purpose is to reduce file size, with a minimum loss of noticeable audio or video data. Some of the standard compression types include H.262 (MPEG-2 Part 2), MPEG-4 Part 2, H.264 (MPEG-4 Part 10), HEVC (H.265), Theora, RealVideo RV40, VP9, and AV1. Every compression type requires its own decoder (Codec) in order to decompress the file and playback the video file. Many digital video codecs are standard compression types, generally known and playable to most operating systems and video platforms. These codes are pre-loaded or easily downloadable with software such as VLC or Media Player. Proprietary video formats, such as surveillance footage or vehicle dash camera mounted systems may come with custom compression algorithms, not playable on standard media players. These type of video files may have proprietary players that allow playback by selecting the encoded file, opening it within the program itself. The video can then be viewed through the proprietary program or exported as a compressed video file, still images, or even screen captured. It should be noted that some of these proprietary players do not always display the videos correctly and may be subject to dropped frames and nonuniform temporal playback issues. The proprietary video players sometimes have additional information such as a time code, that is associated with specific frames or events, and only viewable in the proprietary player. Even though some of these players are capable of converting the video file to a format with a standard compression type, they may not necessarily export the timing information, and thus a screen capture of both the video and time code can be helpful. However, caution must be taken to make sure the screen recording is matching the playback rate of the original video file, that no frames are being excluded, and that the data is now being downsampled as described in the following sections. A better method for analyzing these proprietary file formats is to use a forensic image analysis software package, such as iNPUT-ACE, that can read the file formats natively and allow for the videos to be played back at true frame rates and will address dropped frames. These forensic software packages allow for the conversion of the native video into a standard viewable form without losing data due to down sampling or other issues related to screen capture or transcoding.

When post processing video files, the new created files may be recompressed. This can occur because a smaller file size is desired, a more compatible format is desired (as some codecs are not playable on standard play back software) or through user error. Taking an original file and converting or encoding the file with a different format can have unintended consequence. For example, recompressing the files may cause loss of data, or may create image artifacts as a result of combining pixel data in an effort to lower the data size. Another example is if the source video is a variable frame rate, and in an mpeg 4 compression format, the time between varying frames is maintained. However, if the original file is encoded to another compression format or even if the same mpeg compression is maintained but encoded to a new video file, the variability of the frame rate from the original video may be lost, and the new file will equally space the time between frames, thus producing events in the video that occurring at a time different than they did in real life. Care should be taken anytime a video format is changed or copies are made to minimize or eliminate data loss or modification.

Data Reducing and Down sampling the video:
Down sampling in post processing can remove frames from the source video. When converting from NTSC to PAL, for instance, the video sequence can drop from 30 fps to 25, losing frames in between. Europe, Australia, and other parts of the world broadcast at 25 frames per second, specified by PAL/SECAM which is their version of North America's NTSC. Whether converting an NTSC signal to PAL/SECAM standard frame the frame rate needs to match the original video source. Down-sampling from 30 fps to 25 frames per second, for instance, results in the time between frames in the video being inaccurate. *Figure 24* demonstrates the effect of down-sampling from 30 frames per seconds to 25 frames per second. As the illustration shows, the time between two frames, or two events, is 33.33 milliseconds at 30 frames per second. When this same footage is reduced to 25 frames per second, one frame of every 6 frames is randomly deleted from the video and no longer available, and the remaining video frames are re-spaced to inaccurately show two subsequence video frames, or events, occurring over 40 milliseconds of time. The new down-sampled video, at 25 frames per second, will now show the time frame between the events observed in subsequent frames occurring over 40 milliseconds instead of 33 milliseconds. Depending on how down sampling occurred, the interlaced frames may be collapsed and combined into a single frame instead of two, thus reducing available data, and removing the ability to deinterlace the frames.

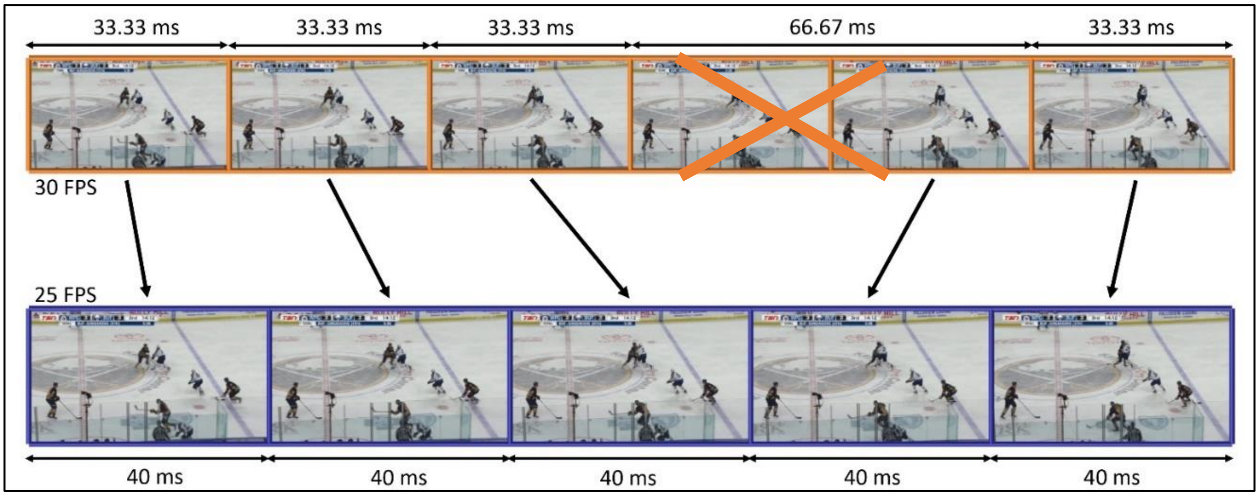

Figure 24 - Down sampling: original video has two subsequent frames at 33.33ms versus 40 ms apart when down sampled.

Video Stabilization:
Panning and zooming of the camera makes the quantification of player motion in video footage more difficult. In video footage where the camera view is stationary (e.g., not panning or zooming), player location can be tracked directly by comparison of video frames over time. However, if the camera is panning and zooming, the motion of the player in the video frame has become relative to the movement of camera itself, thus complicating measurement of player motion in absolute terms. Software programs such as PFTrack can be used to track the panning and zooming characteristics in the video, resulting in a virtual camera matching the movements of the actual camera that recorded the footage. Camera views can be 'stabilized' to remove the effects of camera movement, panning, tilting and zooming so that the background shown in each image remains stationary (PFTrack, Nuke X 12.0v2, Foundry, London, UK). This procedure simplifies

the tracking process by reducing the number of variables the tracking software needs to evaluate. To perform tracking, the video can be brought into AfterEffectsTM or similar editing software and each frame exported as still images. Nuke™, Houdini, or similar software can then be used to import the individual frames and stabilize any bouncing, shaking and rotation of the camera.

By using the technique of video stabilization, camera motion such as panning and zooming can be effectively removed from video footage. Once stabilized, visual elements which are known to be stationary, such as markings on the playing surface, don't move, while player movement is reproduced as if the original camera had been stationary. Player position can then be determined directly, as before, by comparison of video frames over time.

Research by Bailey et al.[1,2] developed techniques for stabilizing footage of a camera that was shoulder mounted on a videographer moving up and down the playing field. The videography attempted to keep the test subject in the center field of view, but the results, though good for broadcast, created streaking and blur for tracking purposes. In short, because the camera is panning to keep a single object in view, the background blurs. Through stabilization, the movement of the camera is minimized, and the object is kept as close to center as possible, while the background is also kept stationery. This provides a background that is more in focus, and a smoother position per frame for the object of interest. *Figure 25* depicts the original series of images, showing panning and un-stabilized footage. *Figure 26* depicts the same sequence stabilized.

Figure 25 - Original video of moving/panning camera

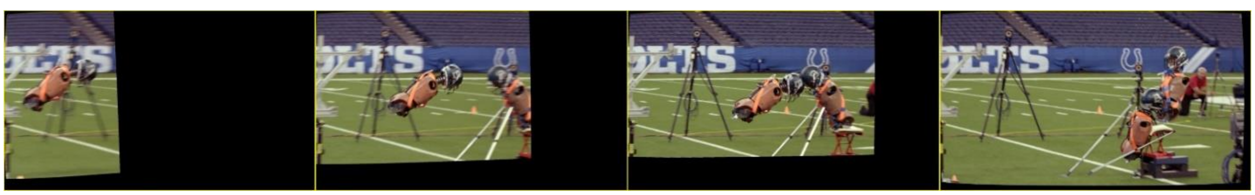

Figure 26 Video after stabilization

Color Correction:

This difference between printed colors and colors displayed on digital monitors is so well known and documented in professional industries that specific color calibration charts, techniques, methods, and devices are utilized to account for this problem[1]. Physical charts contain printed ink color whereas the digital video does not use ink but rather light for displaying colors. Printed ink comes in color combinations referred to as CMYK (Cyan, Magenta, Yellow, and Black) whereas digital displays are in RGB (Red, Green, Blue) and thus by definition cannot be the same colors, as their chemical and optical make up are formed from difference base color compositions. To determine the actual color being recorded by a digital camera, a real-world color card can be placed at the scene, and a digital recording taken of the card. *Figure 27* depicts a Macbeth ColorChecker color calibration chart that is standard and used for this process[2]. When the

[1] Terpstra, T., Neale, W., Hashemian, A., "Photogrammetry and Analysis of Digital Media", Published through SAE Technical Course Material, Troy Michigan. (2017-2021)

[2] Calibrite- Using the ColorChecker Passport in Photoshop and Adobe Camera RAW (January 2022), Available at: calibrite.com/us/wp-content/uploads/sites/2/2022/01/White_Balance_and_Colour_Calibration_in_Photoshop_ACR10.3.pdf

recorded footage is reviewed, and displayed on a monitor, the color chart can then be held in hand, and used as a live visual means to calibrate the monitor so the colors in the monitor are displayed correctly.

Figure 27 – Color Correction Chart used to calibrate digital images

5. Limitations in technology, and in analysis of video

Video analysis and tracking is an essential tool to perform prospective and retrospective analyses of recorded events. The accuracy of video tracking is dependent on many variables. In sports, several studies have been conducted to assess the effects of frame rate[1,33], resolution or number of pixels on the object.[1,10] A full exploration of the factors presented previously has not been undertaken to assess compounding factors resulting in degradation or improvement in accuracy. Several factors such as number of cameras used, image calibration, camera frame rate, resolution of images and quantification/recognition of error or uncertainty are some of the additional limitations that must be recognized in video tracking.

Validation studies are helpful in establishing techniques that result in high quality analysis, and assessing the software used, steps taken, range of certainty considered, since the accuracy of the video tracking can be variable depending on the video data, camera calibration, software used as well as the analyst performing the video tracking. An analyst or research paper simply quoting a stated accuracy from an existing research study is not sufficient unless that research is being applied correctly to the subject analysis. Validation of the tracking methods that incorporates the range of cameras used for the analysis, frame rates and object tracking methods, inclusive of software and analysts is essential to understand the uncertainty of an analysis. Some video tracking software use ray casting methods (ProAnalyst3D) or pixel error (PFTrack) to understand the accuracy of the camera calibrations and also the positional uncertainty when tracking linear movement of an object in three-dimensional space. This provides real-time feedback on the accuracy of the video tracking and is a tool to understand uncertainty.

When applied to helmeted or un-helmeted events it must be recognized that gross kinematics such as linear or rotational position and angles are the outputs from a video tracking analysis. Pre-impact velocities and post-impact velocities are calculated through differentiation. What is critical, is that pre- and post-impact velocities do not represent the impact severity. Change in velocity (ΔV) and change in rotational velocity (Δω) must be calculated through vector analysis to assess a measure of the severity of a collision. In calculation of these impact severities, the duration of the impact must be considered for the vector analysis. When selecting a time window to calculate dV, a time window that is too short and does not capture the entire impact event can underpredict the Delta-V (ΔV, Δω), whereas, a time window that is too long and incorporates velocity changes not associated with the impact can overpredict Delta-V (ΔV, Δω). As it relates to tracking accuracy linear and angular velocities and change in velocities of athlete heads and helmets in sports have been reported to be tracked with linear velocity errors of 10%-11%[15,20] or better.[1,10] The calculation of impact severity in a helmeted event relates to helmet kinematics and not head kinematics. It has been shown that although using helmet impact severities (ΔV, Δω) to estimate head impact severities may be appropriate[11] the same cannot be said for accelerations. Helmet accelerations are not representative of head accelerations.[11,10,26] Furthermore, typical in-game video is of frame rates of 240 images per second or less. Head accelerations in an impact event typically occur over a time window of 15 – 20 ms.[23] Therefore, reporting head accelerations from video tracking of these lower frame-rate videos is inaccurate and tends to underpredict head acceleration. Frame rates of approximately 1000 fps would provide the fidelity necessary to accurately calculate head accelerations.

6. Limitations in interpretation of results

All measured data has limitations, and these limitations are based on the data from which they are measured, and the methodology used to perform the measurement, along with the intentions or goals of the analysis. The higher the quality of the data and the more robust the analysis is, the higher the precision and accuracy of the results of that analysis will be. A robust analysis can be hampered by lower quality data and vice versa. When developing conclusions, the analyst must consider the quality of the data being analyzed as well as the methodology being employed. In some cases, data reduction and filtration techniques can be utilized to account for some of the limitations in the methodology of the analysis by averaging the data and looking at trends. Most videos analyzed do not have the fidelity of a high enough frame rate and resolution to measure the accelerations sustained in an impact directly. However, many videos do have a high enough fidelity to measure average accelerations over longer periods of time. We as analysts must rely upon other methods of estimating/calculating the peak accelerations in impacts based on time, position, velocity, and orientation measurements resulting from the video tracking analysis. These estimation/calculation techniques may fall outside the scope of this paper.

All video tracking data, both measured and calculated, are driven from a spatial/displacement tracking basis. Since all the data in video analysis is displacement based, calculated values require the use of differentiation. As previously described, there are many factors that can affect the precision and accuracy of the measured data that results from a photogrammetric analysis. In many cases, precision limitations lead to “noise” in the data. These limitations will be amplified when calculating the first order derivative (speed/velocity) and even more so when calculating the second order derivative (acceleration). For this reason, varying filtering and regression methods can utilized to account for high accuracy data that has low precision. However, these filtering and regression techniques have limitations and must be applied properly.

Precision limitations typically come from four sources: image resolution/quality, pixel tracking, temporal measurements, and differentiation. The first two factors relate to the tracking of a pixel, and therefore also relate to variances in the measured spatial displacement of a camera or tracked object. This typically results from a cyclical precision error where the displacement of an object is first overrepresented in one measurement while being underrepresented in the next measurement. This is caused by the tracked positions being imprecisely located. This imprecision typically comes in two forms. The first of these is that the resolution, contrast, or clarity of an image is insufficient to properly define the exact location of the object being tracked. *Figure 28* consists of two images generated from the same raw file. The image on the left is the raw high-resolution image and on the right is a reduced resolution copy of the same image. Each of these images allows accurate tracking of the vehicle’s average speed/velocity and average acceleration over multiple frames. However, the image on the right includes much more noise in the results as the reduced resolution induces uncertainty into each individual measurement. A similar comparison could be made based on the precision at which the user and/or the analysis software tracks a pixel or a set of pixels. The more time and effort applied in verifying the exact position of each tracked position, can lead to a higher precision in measurement.

Figure 28 – Precision limitations of resolution and pixel tracking

The next precision limitation in raw data can come from the temporal side of the measurement equation. As previously described, there can be a variance in the time between frames. If not properly accounted for, this can lead to errors in calculated speeds/velocities and accelerations. A high variance in calculated data can be an indication of a variable timestep. However, it will not be an indication of a biased offset of the data as biased offset will not produce noise in the data. An example of a biased offset is calculating the speed of a vehicle using 15 frames per second (fps) for a video captured at 30 fps. This error in frame rate will result in a calculated speed that is half the speed at which the actual vehicle was traveling. In the case of a variable timestep, filtering, averaging, or regression can be applied to isolate affected regions of the data for further analysis.

The fourth precision limitation in raw data can come from the numerical differentiation of the displacement data. The act of calculating the derivative of discrete numerical data will inherently exaggerate any noise in the data. The only way to reduce the noise in the data is to increase the precision of the raw data or to average, filter, or regress the resulting calculated data. As an example of this behavior, Figures 29 and 30 show an example of what occurs when noisy/uncertain data is differentiated. Figure 24 is a plot of a linearly increasing data set with a slope ($\Delta Y/\Delta X$) equal to 10. This baseline data is represented by a grey line. To simulate noise/uncertainty in the data a random number between –5 and +5 was added to each datapoint. This simulated data is represented by an orange line.

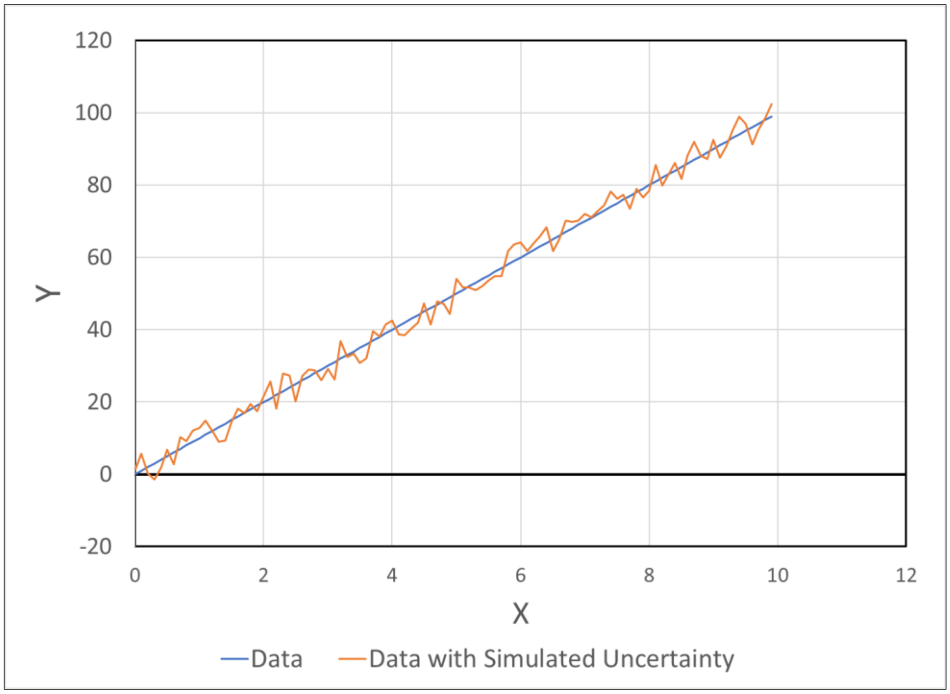

Figure 29 – Linearly increasing data with simulated noise/uncertain data

The data shown in Figure 24 was then differentiated. The results of the differentiation are shown in Figure 25. The green line represents the differentiation of the base data without noise. This green line equates to a constant value of 10 as expected. Next, the red line represents the results of the differentiation of the simulated noisy/uncertain data. Notice how the data oscillates about the true value of 10 but the noise and uncertainty in the data are amplified. One way to reduce noise amplification is to calculate the differential over a larger window. This helps to average out the noise and can result in a more precise measurement of the true differential. Notice in Figure 25 how the data becomes more precise as a larger window for differentiation is utilized (grey and blue lines). This is a way of calculating an average differential over a larger time period. While this method provides more precise measurements with less uncertainty of the calculated differential, there are limitations to averaging techniques that will be discussed later.

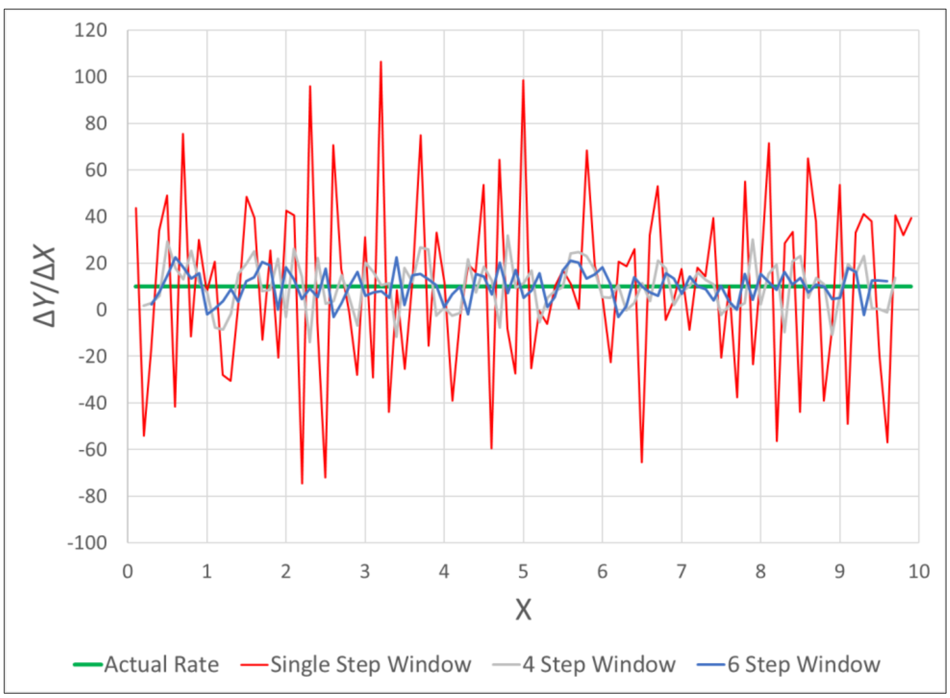

Figure 30 – Example Plot of noise amplification due to differentiation

In any video tracking analysis, one or all these factors may combine to produce noisy or less precise results. When presented with noisy results, an analyst should first review the raw data and the analysis to determine whether the precision of the raw measurements can be increased. This will result in a higher precision of the outcome. If the precision of the raw measurements cannot be increased, the calculated values can more precisely be determined by using filtering, averaging, or regression.

Let’s first consider filtering. Various types of filters exist and several of them used in the reconstruction field are a form of a low-pass, a high-pass, or a band-pass filter. These filtering methods are traditionally used in physical testing when analyzing measured results. A low pass-filter dampens out high frequency data while leaving low frequency data unchanged. Conversely, a high-pass filter dampens out low frequency data while leaving high frequency data unchanged. A band-pass filter is generally a combination of a high and low frequency filter where all frequencies outside of a specified range are dampened out. Again, these filtering methods have been historically used in varying forms of physical testing. The problem with using these filtering methods within a video analysis is they require a very high sampling rate.[25] Standard frequently references and recommends the use of a Class 180 filter prior to integrating acceleration data when calculating velocity or acceleration. Section 8.2 of SAE J211 requires that the sample rate of the raw data to be filtered should be sampled at a minimum of 10 times the filter frequency. Therefore, a minimum sample rate for a Class 180 filter is 1,800 samples per second. This does not account for the fact that the data is usually prefiltered with a Class 1000 filter which requires 10,000 samples per second before it is filtered by the Class 180 filter. These methods are utilized for filtering acceleration data before integrating to calculate velocity and displacement. The calculations of video analysis are performed in reverse order of what is described in SAE J211. In video analysis, the differentiation of displacement data to calculate speed/velocity and acceleration will likely require a much lower frequency filtering method. Even so, most video clips being analyzed are going to be anywhere from 1 to 60 fps, so using a filter is likely not going to be possible or appropriate due to the low sample rates (frame rates).

In the absence of using industry standard filters, the analyst must resort to other methods to increase the precision of calculated data. The simplest option is to use a moving average. By default, the average speed is determined by calculating the distance traveled between two adjacent frames. This distance is then divided by the time between the two frames to determine the average speed over that time. Low precision in raw data can occur when the uncertainty of a measurement becomes large compared to the distance being measured. In simpler terms, when the error in the measurement becomes significant compared to the value being measured, it results in uncertainty in the measured data. When this occurs, the precision of the measurement decreases. A similar comparison can be made using the time between frames. Often, the limitations of the analysis and/or the raw video prevent the reduction in the uncertainty of the raw measurement. When this occurs, an analyst must increase the distance and time over which the measurement is performed. By increasing the distance and time in which the value is measured, the uncertainty becomes increasingly less significant. Subsequently, this leads to more precise results.

A downside to averaging is that there is a temporal shift in the data that must be accounted for in the analysis. Additionally, a moving average assumes the subject object is traveling at a constant average speed over the timeframe being measured. If there is a sudden change or a localized maximum or minimum in the value being measured or calculated, the change could be averaged out. Averaging works well for tracking objects where the value being measured does not undergo a sudden change over the time in which the value is calculated.

To illustrate the limitations of a moving average analysis, *Table 4* is a plot of example speed data from a passenger vehicle. The data begins at a constant travel speed of of 70 mph. At 4 seconds before impact, the vehicle began to brake at a rate of 0.5g. The braking continued until impact (0.0 seconds) where the braking stopped for 1.25 seconds. The braking then resumed at 1.25 seconds until the vehicle came to rest. For simplification, noise has not been added to the raw speed data as the effect on the base data is what is being illustrated. In *Table 4* there are 4 lines: the black line is the raw data, the red line is a 1 second moving average, the green line is a 2 second moving average, and the yellow line is a 3 second moving average. While this method efficiently removes noise by calculating the average over a longer time, should the true value abruptly change, the moving average will serve to flatten the data. This is evident in *Table 4* as there is almost no remanent of the 1.25 second constant speed shown in the yellow line. It should also be noted that this analysis truncates the values due to the time shift in the data. In other words, there is an absence of data on the left and right side of the plot. For this reason, data points must be on either side of the desired measurement point to utilize this method.

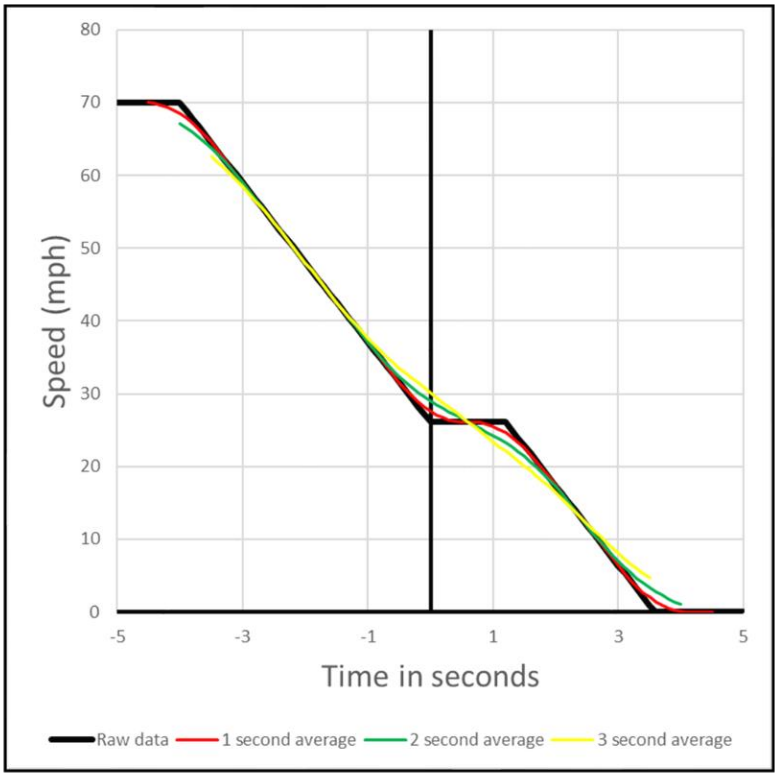

Table 3 – Example of moving average analysis limitations

Some general rules when applying the moving average noise reduction method include:

1. The longer the period over which the data is averaged, the more adequately the subject data will be smoothed.
2. The longer the period over which the data is averaged, the less the resulting data's precision will be subject to the uncertainty in each individual raw measurement. This can be used to account for poor data such as poor resolution, low frame rate, nonuniform timecode, and low precision tracking.
3. The longer the period over which the data is averaged, the more real-world sudden changes in the data will be averaged out.
4. The longer the period over which the data is averaged, the more the data will be truncated requiring more and more samples on either side of the data being measured.
5. The less constant, or linear, the data being measured, the shorter the averaging time window needs to be.

To address some of the limitations of a simple moving average filtering technique, a regression analysis may be performed. A regression analysis can better represent a more precise measurement for a value being calculated. The regression method excels over other methods when the data being analyzed does not represent a constant or a simple linear increase or decrease. The downside to using a regression method over a simple average is that more samples are needed. Therefore, the regression method cannot represent sudden and drastic changes in data. Additionally, the general expected shape of the data being fitted needs to be known. In this analysis, the analyst is essentially fitting the data to a theoretical shape. If the wrong shape is chosen, the regression will not properly represent the value being measured. In the case of sudden changes in data, the data set can be broken into subsets of data and a regression can be performed on the subsets individually to properly represent the measured data.

To illustrate an example of the use of data regression, *Figure 31* is a plot of two different regression analyses performed on a control data set. The orange line is the control data that was generated to have noise. The orange line is equal to $Y = X^2 + R$ where R is a random number between -1 and 1. This illustrates a base function of $Y = X^2$ with some simulated random noise added. To determine the base equation, two regressions are shown. The first is a regression where the equation is assumed to be a first order polynomial equation (linear). This analysis does not represent the data set well which is reflected by the R^2 value. R^2 is a statistical measure of how well the regression represents the raw data. A perfect match to the data would result in an R^2 value equal to 1. The further R^2 deviates from 1, the more poorly the regression equation represents the raw data. With $R^2 = 0.0001$, the linear regression does not represent the raw data, which is expected. On the other hand, the second order polynomial regression represents the data well as expected. Again, this is depicted by the resulting R^2 value.

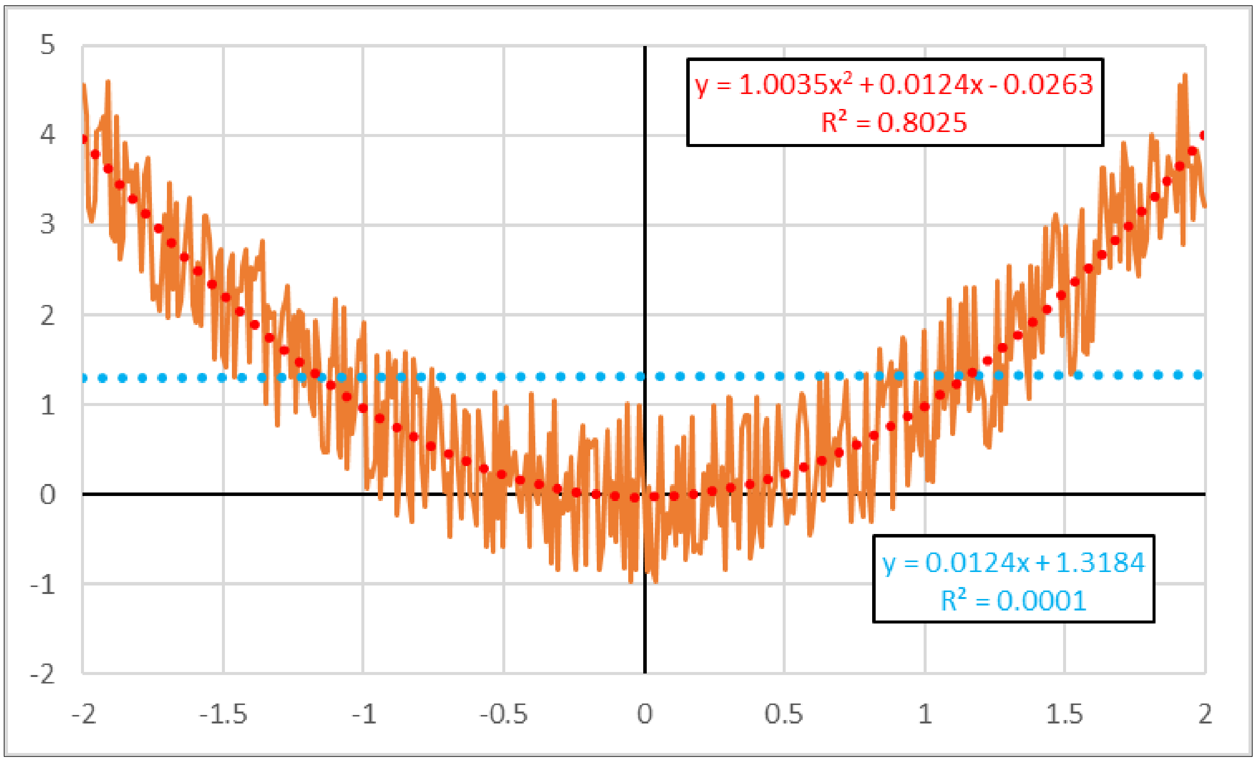

Figure 31 – Example of data regression analysis

To summarize, there are many factors that can induce low precision into a measurement in a video tracking analysis. When a data set shows indications of low precision, it does not directly indicate that the data is inaccurate. There are many data filtering methods that can be used to increase the precision of the measurements, in other words reduce the noise in data. These methods, however, must be selected and implemented properly as to minimize the effects on the accuracy of the values that are being measured. When proper methodology is utilized, the precision of noisy data can be increased significantly without affecting the accuracy of the data.

7. Conclusions and Summary

The term “objects in the video” used in this paper should be interpreted in liberal terms. An object in the video would include not only typical three-dimensional traditional objects, but anything that has shape, and can be measured. Something is still an “object” even when two dimensional in the video, such as paint, liquid or fluid, or even the shadow of a person. Even though these items may not have thickness, the can be analyzed and tracked like any object in video, since video reduces all real world geometry to a 2 dimensional surface, i.e the film or picture plane. “Objects” in video are simply shapes, but through techniques developed over the last century, the location, dimensions, shape, scale, position over time, and other characteristics can be known for these shapes. For the purposes of video analysis and photogrammetry and tracking, objects in the video cover pretty much anything you see in the video, regardless of its three dimensionality.

The parameters described in this paper, and their potential limitations are not exclusionary or absolute. Simply because a video has been down sampled, for instance, does not mean the video is no longer usable. Accurate and useful data can still be obtained from grainy, blurry or recompressed video imagery. The important point being made, is that when analyzing the video, these parameters and their limitations need to be considered, and accounted for. Where possible, the best video versions, of highest resolution are optimal. If not available, care should be taken when analyzing the video to make sure the opinions and results that are being obtained from the analysis are supported by the video and the potential information in it.

With respect to the use of video in determining impact-induced head accelerations, the previously described examples, methodologies, and limitations provide a framework for determining an object’s speed and/or velocity. Video analysis, even when performed using highly specialized marker-based tracking systems, typically do not have the necessary spatial or temporal resolution to accurately measure impact-induced accelerations, which may be occurring and peaking over a period of milliseconds. Consequently, the use of video in determining head accelerations provides a starting point for a more detailed analysis. Specifically, video analysis can often provide a reasonable range of an object’s velocity just prior to impact. This velocity can be used in calculations, physical experiments, or computational simulations to determine accelerations.

Funding and/ or Conflicts of Interests/ Competing Interests:
This research and resulting paper were self funded. No outside funding was involved in the production of this research. Further, there are no competing or conflicting interests in the production of this research

Appendix A:

1. Bailey, A., J. Funk, D. Lessley, C. Sherwood, J. Crandall, W. Neale, N. Rose. Validation of a Videogrammetry Technique for Analysing American Football Helmet Kinematics. Sports Biomech 19. 678-700 (2018). doi: 10.1080/14763141.2018.1513059.
2. Bailey, A.M., C.P Sherwood, J.R. Funk, W. Neale, N. Carter, D. Hessel, S. Beier. Characterization of Concussive Events in Professional American Football Using Videogrammetry. Ann Biomed Eng 48, 2678–2690 (2020). doi: 10.1007/s10439-020-02637-3
3. Beauchamp, G., D. Pentecost, D. Koch, A. Hashemian, et al., "Speed Analysis from Video: A Method for Determining a Range in the Calculations," SAE Technical Paper 2021-01-0887, 2021
4. Campolettano, E. T.,R. A. Gellner, & S. Rowson. (2017). *High-magnitude head impact exposure in youth football*. *Journal of Neurosurgery: Pediatrics, 20(6), 604–612. https://doi.org/10.3171/2017.5.PEDS17185*
5. Cole MH, W. Van den Hoorn, J. K. Kavanagh, S. Morrison, P.W. Hodges, et al. (2014) *Concurrent Validity of Accelerations Measured Using a Tri-Axial Inertial Measurement Unit while Walking on Firm, Compliant and Uneven Surfaces.* PLOS ONE 9(5): e98395. https://doi.org/10.1371/journal.pone.0098395
6. Gyemi, D. L., D.M. Andrews, & R. Jadischke. (2021). Three-dimensional video analysis of helmet-to-ground impacts in North American youth football. *Journal of Biomechanics*, *125*, 110587. https://doi.org/10.1016/j.jbiomech.2021.110587
7. Hutchison M. G. *Concussions in the National Hockey League (NHL): The Video Analysis Project. PhD Dissertation. University of Toronto. 2011*
8. Instrumentation for Impact Test – Part 1 – Electronic Instrumentation. Society of Automotive Engineers. March 1995. SAE J211-1.
9. Jadischke, R., J. Zendler, E. Lovis, A. Elliott, A & G. Goulet, (2019). *Development of a methodology and preliminary analysis of head impacts in American 7-v-7 non-tackle football. IRC*-19-87, 615–616. Florence, Italy.

10. Jadischke, R., J. Zendler, E. Lovis, A. Elliott, & G.C.Goulet, (2020). *Quantitative and qualitative analysis of head and body impacts in American 7v7 non-tackle football. BMJ Open Sport and Exercise Medicine,* 6(1), 1–7. https://doi.org/10.1136/bmjsem-2019-000638

11. Joodaki H, A. Bailey, D. Lessley, J. Funk, C. Sherwood, J. Crandall. *Relative Motion between the Helmet and Head in Football Impact Test. J Biomech Eng. 2019 Mar 5. doi: 10.1115/1.4043038. Epub ahead of print. PMID: 30835289.*

12. Kamalkeet K. and S. Baljit, *Analysis of CCD and CMOS Sensor Based Images from Technical and Photographic Aspects (January 19, 2020). International Conference of Advance Research & Innovation (ICARI) 2020,* Available at SSRN: http://dx.doi.org/10.2139/ssrn.3559236

13. Lall et al., "*Correlation of Accelerated Tests with Human Body Measurements for Flexible Electronics in Wearable Applications," 2020 19th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), 2020,* pp. 877-887, doi: 10.1109/ITherm45881.2020.9190514.

14. Lessley, D., J. Kent, R. W. Funk, J. R. Sherwood, C. P. Cormier, J. M. Crandall, J. R., Myers, (2018). *Video analysis of reported concussion events in the National Football League during the 2015-2016 and 2016-2017 seasons.* The American Journal of Sports Medicine, 46(14), 3502–3510. https://doi.org/10.1177/0363546518804498

15. McIntosh A S, D.A. Patton, B. Frechede, et al. *The biomechanics of concussion in unhelmeted football players in Australia: a case-control study*. BMJ Open 2014; 4:e005078.doi:10.1136/bmjopen-2014-005078

16. Neale, W., S. Fenton, S. McFadden, N. Rose, “A Video Tracking Photogrammetry Technique to Survey Roadways for Accident Reconstruction.” SAE Technical Paper, 2004-01-1221, 2004, doi:10.4271/2004-01-1221

17. Neale, W., D. Hessel, T. and Terpstra, “Photogrammetric Measurement Error Associated with Lens Distortion,” SAE Technical Paper 2011-01-0286, 2011, doi:10.4271/2011-01-0286.

18. Neale, W., J. Marr, and D. Hessel, "Video Projection Mapping Photogrammetry through Video Tracking," SAE Technical Paper 2013-01-0788, 2013, doi:10.4271/2013-01-0788.

19. Newman, J., M. Beusenberg, E. Fournier, N. Shewchenko, A. King, K. Yang, G. McGinnis, (1999*). A new biomechanical assessmet of mild traumatic brain injury. Part 1*: Methodology. IRCOBI Conference, (September), 17–36. Sitges (Spain)

20. Newman, J., A, M.C. Beusenberg, N. Shewchenko, C. Ã. Withnall, & E. Fournier, (2005). *Verification of biomechanical methods employed in a comprehensive study of mild traumatic brain injury and the effectiveness of American football helmets. Journal of Biomechanics, 38*, 1469–1481. https://doi.org/10.1016/j.jbiomech.2004.06.025

21. Niederer, P., B. Birchler, F. Mesqui, and Y. Lehareinger, "Computer Assisted Single-View Photogrammetry for Accident Scene Documentation," SAE Technical Paper 850067, 1985, doi:10.4271/850067

22. Patton D. A., C. M. Huber, C. C. McDonald, S. S. Margulies, C. L. Master, K. B. Arbogast. Video *Confirmation Of Impact Sensor Data From High School Soccer Players. The American Journal Of Sports Medicine.1-8.* DOI: 10.1177/03635465209D6406. 2020.

23. Pellman E. J., D. C. Viano, A. M. Tucker, and I. R. Casson. *Concussion in Professional Football: Location and Direction of Helment Impacts- Part 2. Neurosurgery 53:6. 2003*

24. Press JN, S. Rowson. *Quantifying Head Impact Exposure in Collegiate Women's Soccer. Clin J Sport Med. 2017 Mar;27(2):104-110. doi: 10.1097/JSM.0000000000000313. PMID: 26978008.*

25. SAE International, "Instrumentation for Impact Test - Part 1 - Electronic Instrumentation," SAE Standard J211-1, Rev. Mar. 2014.

26. Siegmund, G. P., K. M. Guskiewicz, S. W. Marshall, A. L. Marco, & S. J. Bonin, (2016). *Laboratory validation of two wearable sensor systems for measuring head impact severity in football players.* Annals of Biomedical Engineering, 44(4), 1257–1274. https://doi.org/10.1007/s10439-015-1420-6

27. Smith, G., and D. Allsop, "A Case Comparison of Single-Image Photogrammetry Methods," SAE Technical Paper 890737, 1989, doi:10.4271/890737.

28. Terpstra, T., J. Dickinson, A. Hashemian, and S. Fenton, "Reconstruction of 3D Accident Sites Using USGS LiDAR, Aerial Images, and Photogrammetry," SAE Technical Paper 2019-01-0423, 2019, doi:10.4271/2019-01-0423.

29. Terpstra, T., J. Dickinson, and A. Hashemian, “Using Multiple Photographs and USGS LiDAR to Improve Photogrammetric Accuracy,” SAE Technical Paper 2018-01-0516, 2018, doi:10.4271/2018-01-0516.
30. Terpstra, T., A. Hashemian, R. Gillihan, E. King, S. Miller, W. Neale, “Accuracies in Single Image Camera Matching Photogrammetry,” SAE Technical Paper 2021-01-0888, 2021, doi:10.4271/2021-01-0888.
31. Terpstra, T., A. Hashemian, T. Voitel, J. Priest. A Comparison of Metrology Used in Documenting Shooting Incident Trajectories. J Assoc Crime Scene Reconstruction. 2020;24:23-42.
32. Terpstra, T., S. Miller, A. Hashemian, “An Evaluation of Two Methodologies for Lens Distortion Removal when EXIF Data is Unavailable.” SAE, Paper 2017-01-1422.
33. Tierney G. J., H. Joodaki, T. Krosshaug, J. L. Forman, J. R. Crandall, and C. K. Simms. *Assessment of model-based image matching for future reconstruction of unhelmeted sport head impact kinematics.* Sports Biomechanics. 17:1, 33-47. 2018.