GOOD-BAD © CANSTOCK PHOTOS/ STUARTMILES, CHIP COURTESY OF FREEIMAGES.COM/DIMSHIK

Caught in the Act

Law-enforcement feedback on video surveillance recordings.

By Margaret H. Pinson

©ISTOCKPHOTO.COM/FILO

Video surveillance systems, designed to provide security, act like a guard in a watchtower—a security solution that has been well understood for thousands of years. Systems are designed to meet security needs of the customer who installs the system. However, such systems rarely work as useful law-enforcement tools, leading to frustration and disappointment for police, store owners, and crime victims. In many cases, a business owner selects a system because it is easy to install and manage, operates it in good faith, but when robbery or some other incident occurs, no conviction follows. The perpetrator cannot be identified because the system failed for one reason or another to produce evidence valid for a prosecution.

Video surveillance recordings are a prolific source of evidence, yet officers estimate that 95% of the video they collect from stores, residences, and other sites relevant to a crime is never viewed. This indicates vast opportunities for technology innovation.

This article looks at security video from a law-enforcement perspective. I describe the problems officers face using video surveillance recordings as evidence in court. The problems identified in this article reflect first-responder opinions, observations, and unique experiences. My goal is to encourage people to rethink best practices for video surveillance and envision new video technologies that better serve first-responder needs.

RECORDINGS AS EVIDENCE

The Public Safety Communications Research (PSCR) program acts as an objective technical advisor and laboratory for the public safety community. From 2006 to 2012, PSCR assessed the ability of first responders, using security video, to identify people, license plates, and other targets, depending on the lighting, motion, bit rate, and resolution. This effort culminated in a tool designed to help practitioners define technical specifications to meet video requirements [1].

In 2015, the PSCR program began asking first responders about quality problems they experience with images, video, and cameras. The program's goal was to identify ways to fix those problems through technology innovations, standards, and best practices. This article contains feedback on one topic: video recorded as evidence of a crime. See [2] for feedback on other topics. All opinions in this article reflect statements made by the more than 100 people interviewed. The interviews were

Digital Object Identifier 10.1109/MCE.2018.2867981
Date of publication: 10 December 2018

2162-2248/19©2019IEEE

unstructured and did not analyze these statements for accuracy, bias, or rate of occurrence. Due to privacy concerns, the people and organizations are not identified.

PROBLEMS

POORLY PLACED CAMERAS

The widespread use of video has changed how criminal cases are investigated. Twenty years ago, many investigations were stymied because of the lack of an eye witness. Today, the possibility of getting video evidence means more criminal cases are recommended for investigation. For example, through the use of video, someone recognizes a face on an image from a prior conviction, and an investigative lead is developed.

Identification is best done with a set of images showing the subject from multiple angles. For a person, this means a face shot straight on, which is suitable for facial recognition and use in database searches. The set of images would also include several full body shots at different angles. The images should accurately reproduce colors and identifying features (e.g., scars, tattoos, logos, dirt, blood, scrapes). These pictures create leads, advance the investigation, and answer the critical question—who did it help. A set of vehicle photos ideally shows the license plate plus full vehicle shots from the front, side, and back.

Surveillance cameras are typically installed at a height of eight or more feet, so the picture only shows the tops of people's heads. A typical store has two cameras. The first is over the cash register, positioned mainly to catch employees stealing from the till. The bit rate and resolution are often low [Quarter Video Graphics Array (QVGA), 320 × 240 pixels] because a higher bit rate and resolution are not needed for identifying employees. The manager knows who is on the register at any given time. The camera might be perfect for spotting dishonest employees, but provide poor images of, say, bandits robbing the store. The second camera is a door camera with the same settings and a wide field of view to show people approaching. This camera also might not produce suitable images: faces, captured at low resolution, are blurry, and, viewed from above, might be obscured.

NO SOUND

Audio is a critical piece of the puzzle for law enforcement. For example, backward-facing cameras with audio in a squad car fully record both images and sounds of incidents, such as when noncompliant prisoners refuse to get out or bang their head on the metal divider. Audio in other police settings might capture, for example, a hostile person's spontaneous utterance, the sound of Morse code tapped on a squad car's window, or the conversation of prisoners in separate cells yelling to each other.

Police agencies have learned from experience that the best location for the microphone is not necessarily at the camera. Consider a surveillance system in an interview room. Suspects who confess typically slump forward and mumble their confession to the floor. This soft confession is very important but impossible to record if the microphone is attached to a camera in an upper corner of the room, as in Figure 1. Depending on the application, the microphone often needs to be positioned separate from the camera.

> The perpetrator cannot be identified because the system failed for one reason or another to produce evidence valid for a prosecution.

Audio is arguably more trustworthy than video. Microphones record sounds from all directions and thus overcome the directional bias inherent to video. Video with audio is less likely to be misleading, because the sounds and conversations add context. Consider a suspect being confronted by several officers. A video-only surveillance camera behind the suspect will only show officers drawing firearms. Only if audio is recorded will the viewer be aware of someone yelling, "He's got a gun!" By contrast, privately owned video surveillance systems typically deliver poor audio or no audio at all, making any recordings less useful for criminal prosecutions.

LACK OF EXPERTISE

Law-enforcement officers multitask in a complex, time-sensitive, and hostile environment. Their technology needs to be foolproof. The officer needs to download video quickly and then move efficiently to his or her next task. Unfortunately, video surveillance systems seem designed to prevent officers from accessing and using the recorded videos. Small departments often encounter catastrophic failure: the video can only be viewed on the store's system, and the officer resorts to taking pictures of the monitor with another camera. Larger departments in time-critical situations (e.g., all-points bulletins) may, to avoid a complex or difficult exporting process, take pictures of video displayed on monitors.

FIGURE 1. Confessing people slump and mumble, making confessions difficult to record through the microphone of a camera set at an angle like this.

Today, the possibility of getting video evidence means more criminal cases are recommended for investigation.

The market offers a huge variety of surveillance systems. However, only large departments can afford a certified forensic video analyst, and most officers are not tech savvy—just like most people who own surveillance systems. Complications arise from forgotten passwords, convoluted and error-prone software, missing manuals, and the need to consult vendors. The hardware is often installed in an obscure location to prevent criminals from stealing the hard drive, and wireless connections are rarely available. In some cases, problems with the system or a belligerent owner lead officers to confiscate the video surveillance system.

INCOMPATIBLITY AND TECHNICAL PROBLEMS

The owner's belief that video was recorded does not guarantee that it can be obtained from the surveillance system. A loss of power or change of system settings can erase the video. Inexperienced users assume that video streaming live to a monitor indicates that it is also being recorded. The system's memory may be full or the recording medium may have failed, a situation unnoticed or simply ignored. Officers do not always validate the export and miss errors. Some systems erase the video before an officer can retrieve it. Older systems retain videos for only a brief period—one day, three days, or a week—as do some systems designed primarily for real-time monitoring.

Interoperability problems abound, so the exported video files may not play on the department's computers. One industry expert estimated that there are 1,000 different surveillance video file formats; another estimated that 20% of surveillance video file formats can only be read with vendor assistance. An abundance of aging systems exacerbate interoperability problems. People do not understand the need to update surveillance systems as they do other equipment (e.g., computers, vehicles).

Problems occur because no one bothered to manually record metadata or when automatically exported metadata is found to be inaccurate. The exported video's time stamp can be off by minutes, hours, days, or even years. Video feeds are often unhelpfully labeled (e.g., Cam1, Cam2) or nearly impossible to properly identify, such as in the case of videos from multiple stairwells in a building where the stairwells all look alike. These inaccuracies make it more difficult to track the movement of people and vehicles across footage from those different systems. Inconsistency between the times logged in the officer's report and the times recorded in the surveillance video cause complications in court.

UNREALISTIC EXPECTATIONS

There is a fundamental mismatch between what consumers expect, what surveillance systems deliver, and what investigators need. Consumers assume incorrectly that the video streamed live to the surveillance system monitor will be available in the future without loss of quality. Instead, the recorded video is typically of much lower quality. Exporting the video usually requires a format conversion, which lowers the video quality a second time. The quality typically drops a third time when the surveillance video is paused to create photographs for investigators.

The human visual system fuses sequential video frames to create an illusion of motion. Forensic video software uses super-resolution algorithms to simulate this phenomenon, but the popular video pause shows defects that are normally imperceptible. Thus, a suspect's face might be recognizable when playing the video on the system but not in the final photo. Conversely, the entire video may be indistinct. Meanwhile, images can be obscured from, say, restaurant grease or spider webs on lenses. At dawn or dusk, the exterior video may be washed out because the camera points toward the sun. At night, the interior video may be severely underexposed.

Rapid advances in video technology have changed our expectations for picture quality. Many stores have a QVGA system purchased a decade ago for US$5,000. The business owner is loath to replace it with a new US$500 system, despite the many advantages, such as 720p video (1280 × 720 pixels), four more cameras, and dramatically improved video quality.

SHORTAGE OF FORENSIC SKILLS

Forensic video analysis is the scientific examination, comparison, and/or evaluation of video in legal matters [3]. Forensic video analysts are trained to differentiate between real events and artifacts created by the video technology. Analysts help the courts interpret video. Many departments interviewed did not have a forensic video analyst.

From an analyst's perspective, video is not an accurate medium. Cameras alter colors, distort perspective, discard peripheral vision, and impair distance perception. As the quality drops, video encoders can warp shapes, smooth contours, invent texture, accelerate some gestures while decelerating others, and contrive movement from stillness. Video typically conveys only a tiny fraction of the information delivered to our brains by our natural senses, yet paradoxically it may also record details imperceptible to the human eye.

Nonexperts often assume each video frame has photographic accuracy and thus rely on video to reach unwarranted conclusions. When comparing two sequential frames in a 30 frames/s video, nonexperts conclude that the changes occurred over 1/30 of a second. Sometimes this is true; other times it is not. A coding artifact on a suspect's neck may look like a tattoo. What looked like a gun to the officer may be clearly visible in the video as a garden hose. A moving person may be briefly depicted as motionless. As quality drops,

the video becomes more inconclusive, but people sometimes see what they want to see—like an officer swinging a club to hit someone. Surveillance videos often have color consistency problems (i.e., perceived colors change in response to the illumination and camera) [4].

EMPHASIS ON SECURITY

Basic investigations consist of identifying suspects and vehicles. A man enters a store, grabs razors from a display near the entrance, and runs out. Thieves break into cars parked at an isolated trailhead that is only reachable by vehicle. More complicated investigations involve tracking images of several people and vehicles captured on multiple cameras and potentially over many hours of video. The goal is to determine who was where and at what time. Just as no man is an island, no system is an island. Investigators build a cohesive understanding of events from an array of privately owned and independently designed video surveillance systems. Investigators use surveillance video for purposes beyond their intended use. For example, cameras set up to spot break-ins taking place might be used to help identify someone walking across the street.

Investigation needs differ from security needs. Most video surveillance systems are designed to observe events that affect site security, not to help the investigator identify and track suspects and vehicles. Investigators have mixed feelings about surveillance video that seems to stem from this discrepancy. A few departments even said that surveillance videos rarely helped them solve cases. However, a video of the crime is very compelling in court. Defendants will typically plea-bargain when presented with high-quality video depicting their crime. When the quality is bad, the defendant can argue "that's not me" and avoid conviction. An increasing problem in court is the attitude that, "If there is no video, it did not happen."

Preventing video tampering and maintaining the evidence are major issues for law enforcement that video surveillance systems do not address. A major roadblock is interoperability. Video evidence moves between diverse computer systems operated by many departments (e.g., law enforcement, state attorney's office, defense, prosecution, and the court).

INADEQUATE VIDEO MANAGEMENT

In the United States, each local jurisdiction may have a unique policy for video retention, distribution, and redaction. Video retention requirements depend upon the type of case and might be 30 days, 90 days, one year, or indefinitely. These policies balance competing needs, such as accountability, privacy, expense, and exploitation [5]. Surveillance videos can be misused for criminal abuse, voyeurism, or other inappropriate purposes. Such concerns drive policies for video distribution and, in some jurisdictions, redaction.

There is no clear best solution for data storage. Read-only digital versatile discs and compact discs provide a surprisingly popular solution; they lay flat and thus store efficiently in a filing cabinet. Local computers are more expensive due to the need for triple fault protection, full separation from the in-house system, and a high level of computer expertise that many departments lack. Online options can be prohibitively expensive.

FIGURE 2. As more bodycams are deployed, law-enforcement agencies will contend with increasingly complicated redaction policies.

Outside of the law-enforcement community, the conventional wisdom is that data storage costs are declining. Law-enforcement officers disagree. Their digital storage costs are rising rapidly due to the proliferation of video recordings and the demands to extend retention durations. Logistically, these costs cover surveillance video, photos of crime scenes (≈500 photos for a typical homicide or suicide), in-car camera footage, and bodycam footage. Unusable videos and images cannot be deleted ahead of the retention schedule because such deletions would raise concerns that the department is trying to hide something. Moreover, the trend toward higher resolutions and faster frame rates, which require more data capacity, is swelling everyone's video data storage needs [6].

Redaction digitally removes portions of a video, such as faces, tattoos, profiles, license plate numbers, patient names, and addresses. Reflected faces provide an unusual challenge, e.g., an officer's face reflected in a car mirror or on a windshield. Automatic redaction software is not yet reliable, so each video frame must be manually checked [7]. The swelling numbers of first-responder bodycams are raising concerns about personal and medical privacy [5]. Figure 2 is an example of an image recorded in an emergency that touches on medical privacy.

FIRST RESPONDERS' WISH LIST

First responders have proposed innovative surveillance systems designed as law-enforcement tools that would do more than simply record video. Such a surveillance system must be designed to produce video, images, and reports that could be checked for tampering and submitted to the courts unchanged. The system would enable courts to view the original video files using a generic computer and without installing new software. Forensic video analysts and the courts need to differentiate between real events and artifacts created by the video technology. Law enforcement and the courts need a reliable mechanism, such as a digital watermark, to verify the integrity of the videos and images [9]. Without such a

(a) (b)

FIGURE 3. (a) Practitioners want high-quality head shots for identification and facial recognition. (b) Surveillance systems provide images at low resolution and high angles, which make the identification of suspects difficult.

system, it is difficult to accommodate redaction and image enhancement while preserving evidence.

Law-enforcement officers want surveillance systems that streamline their investigations. The basic video surveillance system would be designed for a reasonably sized store with one or two entrances and exits. This system would photograph each person who enters the building and each vehicle that enters or leaves the store's lot. The door camera would be at head height to show faces. One might think that a camera positioned in such an obvious way would be obstructed or ripped down by the thief. But in real life, officers say that thieves don't care that their image is being recorded. Multiple cameras to capture images at multiple angles would increase the likelihood of fast identification and better evidence. The system would intelligently craft photos for identification purposes, store them at high resolution in JPEG files, and timestamp each photo with date and time synced to universal time. Figure 3 shows how resolution can affect facial recognition. This basic system would not record video. Basic investigations involve identifying suspects; video showing the crime is not essential. The photos would help judges and jurors reach decisions about the guilt or innocence of suspects.

Advanced systems would act as law-enforcement tools. Smart security sensors [10] would provide real-time situational awareness. Officers say that if a school video surveillance system could detect a fight involving weapons, the police response time would improve by approximately four minutes. Advanced systems would create written reports to aid postevent evaluations. The system design goal would be to produce analyses and evidence, not simply video.

To be a usable law-enforcement tool, the surveillance system interface must be easily accessed and operated by someone with little or no technical expertise. The system would allow first responders to establish a wireless connection using industry-wide standard protocols that provide limited functionality for remote control (e.g., data export). Metadata would be automatically created and exported with the video. The metadata would identify any video analytics used and how they work—so that this information can be explained in court.

Useful video evidence must meet the legal needs of the court and the investigatory needs of detectives. These needs are not being met because surveillance systems are designed for security instead of law enforcement. A surveillance system installed in a school shows a student getting beaten but not the faces of the perpetrators. A street surveillance system shows a shooting but the shooter's license plate cannot be read. New systems designed as law-enforcement tools would transform surveillance system video from an overwhelming and troublesome data stream into targeted and effective evidence.

ACKNOWLEDGMENT

Photography credits are available at the Consumer Digital Video Library (www.cdvl.org). Release approvals were obtained where required.

ABOUT THE AUTHOR

Margaret H. Pinson (mpinson@ntia.doc.gov) has been at the Institute for Telecommunication Sciences, National Telecommunication and Information Administration, Boulder, Colorado, where she is currently a cochair of the Video Quality Experts Group.

REFERENCES

[1] U.S. Department for Homeland Security. (2010, July 26). Defining video quality requirements: A guide for public safety. [Online]. Available: https://www.its.bldrdoc.gov/outreach/video/vqips/vqips_guide/

[2] M. H. Pinson. (2017, May). Technology gaps in first responder cameras. Institute for Telecommunication Sciences. Boulder, CO. [Online]. Available: https://www.its.bldrdoc.gov/publications/3171.aspx

[3] Scientific Working Group on Imaging Technology. (2009, Dec.). Best practices for forensic video analysis. [Online]. Available: https://www.swgit.org/

[4] R. R. Varior, G. Wang, J. Lu, and T. Liu, "Learning invariant color features for person reidentification," *IEEE Trans. Image Process.*, vol. 25, no. 7, pp. 3395–3410, July 2016.

[5] K. Kampfe, "Police-worn body cameras: Balancing privacy and accountability through state and police department action," *Ohio State Law J.*, vol. 76, no. 5, pp. 1153–1200, 2015.

[6] T. Coughlin, "How big are your dreams? Gauging the size of future content," *IEEE Consum. Electron. Mag.*, vol. 6, no. 2, pp. 108–124, Apr. 2017.

[7] S. Sah, A. Shringi, R. Ptucha, A. M. Burry, and R. P. Loce, "Video redaction: A survey and comparison of enabling technologies," *J. Electron. Imaging*, vol. 26, no. 5, July 2017.

[8] NPSTC. (2016, Feb.). EMS telemedicine report: prehospital use of video technologies final report. Nat. Public Safety Telecommunications Council. Littleton, CO. [Online]. Available: http://www.npstc.org/download.jsp?tableId=37&column=217&id=3612&file=EMS_Telemedicine_Report_Final_20160303.pdf

[9] S. P. Mohanty, A. Sengupta, P. Guturu, and E. Kougianos, "Everything you want to know about watermarking: From paper marks to hardware protection," *IEEE Consum. Electron. Mag.*, vol. 6, no. 3, pp. 83–91, July 2017.

[10] M. L. Gavrilova, Y. Wang, F. Ahmed, and P. Polash Paul, "Kinect sensor gesture and activity recognition: New applications for consumer cognitive systems," *IEEE Consum. Electron. Mag.*, vol. 7, no. 1, pp. 88–94, Jan. 2018.