Digital Video Quality Handbook Appendix

First Responders Group

January 2018

Homeland Security
Science and Technology

Prepared for:

Department of Homeland Security Science and Technology Directorate
Washington, DC

This document was prepared under funding provided by the U.S. Department of Homeland Security Science and Technology Directorate (Resilient Systems for Public Safety Communication). Points of view or opinions expressed in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Homeland Security.

Principal Authors: Steve Surfaro, Dan Syed (JHU/APL), Steven Babin (JHU/APL), Jay Chang (JHU/APL)

Contributing Author: John Contestabile (JHU/APL)

The authors would like to express their appreciation to the Department of Homeland Security Science and Technology Directorate and our sponsor Mr. Cuong Luu.

Please send comments to:

Mr. John Contestabile
Program Manager
JHU/APL
11100 Johns Hopkins Road
Laurel, MD 20723-6099
Phone: 443-220-8090
E-mail: John.Contestabile@jhuapl.edu

Publication Notice

Disclaimer

TABLE OF CONTENTS

INTRODUCTION AND BACKGROUND

Purpose

The following document has been generated on behalf of the Department of Homeland Security (DHS) Science and Technology Directorate (S&T). It is intended as an appendix to the May 2013 Video Quality Handbook. The objective of this supplement is to present content related to topics not covered in the original document (especially information related to video standards), and to update the material as needed to reflect innovation and changes in the video environment.

Content of this document will include: background material regarding trends and analytics, which will make up the remainder of the introduction; additional topical use-cases reflecting newsworthy events shaping the use of video; a detailed discussion of available video standards, the organizations that create them and how they might be used; and a potential implementation plan.

The emphasis of this document is on digital data captured by network or Internet Protocol (IP) cameras (network camera and IP camera are two names for the same thing) connected via a Wide Area Network (WAN). Specifically, this document will place special emphasis on the implications of digital video data being exchanged across networks with large numbers of components or participants. Much of this document will examine the following aspects of video data:

(1) Resolution aspects: Specifically, the capture and exchange of digital video data that enables presentation to observers in a form that provides sufficient similarity to that of the original object when considering the purpose of the video.
(2) Compression aspects: The efficient capture and communication of video data in order to be exchanged along with large amounts of data from other sources within bandwidth limitations.
(3) Security aspects: The ability to ensure that data are not accessed or corrupted by outside agents.
(4) Archiving and retrieval aspects: The ability to maintain large amounts of video information and retrieve it when it is needed.

Finally, this document will, when appropriate, identify opportunities to improve the utility of digital video through the use of analytics. Paragraph 1.3 below provides a more detailed discussion of content and analytics and the relationship between the two.

Trends

Emerging technology and markets are having significant impact in public safety and security. For example, advances in the consumer electronics and Information Technology (IT) industries are being incorporated into physical security systems. Critical technologies in the consumer space include:

- Ultra-High Definition (UHD, including 4K video), which enable digital video to achieve increased resolution.
- Near field communication (such as Apple Pay and Google Wallet), which enable secure, contactless data exchange.
- Cloud storage, which can enhance the archiving and retrieval of data. Cloud based solutions are very popular, but few include a detailed implementation plan or an understanding of how such solutions will affect their operations.

- Intelligent, adaptive security devices like IP cameras that automatically adjust for the extremes of ultra-low light and intense light.

Private corporations are showing more interest in entry screening technologies. With the increase in cross-border drug traffic, it has become commonplace for high-risk facilities to link mobile x-ray and backscatter technologies to live video feeds. Cyber security is becoming more important with increased awareness due to the publicity of big data breaches. Therefore, vendors and integrators, along with security departments, need to be prepared to take responsibility for securing their systems.

This Handbook will address case studies in different markets. There are a number of significant trends impacting video surveillance.

School security will continue to be important to the overall community, as will city surveillance and critical infrastructure security. Retailers will still see a return on investment (ROI) with video surveillance, even as they continue to be under pressure from online sales. Public safety will increase its use of video surveillance in event security, especially as more cities develop entertainment centers with the potential for hundreds of thousands of visitors at a single event. Part of this market will be served by the temporary surveillance and entry screening solution market.

In the Standards section of this Handbook, we will identify significant ecosystem members and describe how these groups are evolving and requiring new solutions. Regarding manufacturers and solution providers, a continued inflow of new solutions will continue as physical security continues to attract new companies, including entrants from Asia. This increasing competition means that vendors will have to continue to invest in video surveillance to stay relevant.

Dealers and systems integrators have a continued need to understand requirements from an IT perspective, and to work strategically with end users to sell not simply products, but also value in long-term relationships. Video monitoring providers are playing a greater role as high quality video becomes available even for small systems at reasonable cost, and as expectations for video verification from the end users continues to emerge. Great opportunities exist in the area of video monitoring, with additional value from increased bandwidth, mobility and advanced technology. For example, alarm systems are beginning to become integrated with video verification. New entrants such as Telco and Google will continue to make inroads in the residential security systems market.

There are initiatives in some vertical market segments, such as schools, for policies around security. The need for school safety and security standards and best practices is being met by states with the largest systems, including California, Florida, New York and Connecticut. Critical infrastructure working groups are now focusing video surveillance efforts in defending critical industries and utilities, such as petrochemicals, power, food and water.

There are, however, some pressing security industry issues that so far remain unresolved. These are areas where technology is ahead of the industry. One example is integrated systems. While most security managers would agree that security systems should be fully integrated (e.g., intrusion, access control and video), most systems today are still stand-alone systems. End users are rapidly replacing "closed" appliance-based solutions with platforms linking security devices for scalability, agility and elasticity. Video verification is an important issue because most alarms today are not verified by video, which means that guards/police are dispatched on many false alarms. Video verification could help reduce response to false alarms, and also make safety staff better prepared when they respond to a real alarm.

With more data breaches in the news, there are increasing demands for very tough cyber security requirements, which vendors and integrators need to understand and implement.

VIDEO ANALYTIC CONCEPTS

Digital Multimedia Content (DMC)

Digital multimedia content (DMC) is defined to include the video content itself, plus associated metadata (feature data) and audio. The storage location or device (i.e., virtual or cloud storage, or network video record server) where digital video, digital multimedia content or Digital Multimedia Evidence (DME) is originally stored is significant in maintaining Video Quality of Service (QoS), which is the ability to acquire, render, search, disseminate and distribute DMC.

DMC is sometimes called IP video, digital video, IP video content or DME. It may also be categorized as original, copied, local or virtual. Digital data represents audio content, video content, metadata information, location-based information, relevant IP addresses, recording time, system time and any other information attached to a digital file. DMC may either be uncompressed or compressed to save bandwidth and storage.

When compressed or transcoded from the original DMC into an industry standard file format, only a reduced amount of data are then needed to represent the original data set. For forensic readiness, the original DMC is extremely important; data recorded and retrieved to DMC media in its native file format (i.e., first usable form) must always be retained at the embedded video camera solid state media, local network attached storage, local server or virtualized cloud. For further information, see the Digital Video Handbook Volume I.

Video Content Analysis (VCA) and Video Analytics

Video analytics solutions (also referred to as Video Content Analysis or VCA) are a set of computerized vision algorithms that automatically analyze live or recorded video streams without the need for human intervention. The output of these algorithms results in actionable data.[1]

[1] Miki Schwartzburg, Agent Vi seminar on public safety, 06/2015

Adding video analytics to a surveillance network allows the system operator to be more effective in the detection, prevention, response to and investigation of incidents captured by the surveillance system. VCA can also be used to collect valuable business intelligence about the behavior of people and vehicles.

VCA can be applied to three categories of video function:

- Real-time Situational Awareness & Incident Response: VCA alerts may be defined based upon characteristics of actual events, scenarios and detections. Alerts can be generated in real-time to support immediate responses.
- Non-real-time Forensic Analysis/DMC Search: VCA can be applied to previously recorded video to detect conditions corresponding to events of interest to facilitate after-the-fact analysis.
- Business Intelligence: VCA can be applied to video data recorded over significant periods of time to compile statistical data for consumer analysis.

VCA is commonly used in support of security and perimeter protection, traffic monitoring and asset protection.

VCA can be used to detect either specific objects or behaviors. Based upon the application, detection of those objects or behaviors can be used to initiate alerts, object/feature tracking, or other responses to the detected object or behavior, or to exclude (i.e., filter) additional information about those objects or behaviors. Some common VCA functions include:

- Target Type detection – Identification within video data of objects such as people, vehicles and static objects.
- Event Type detection – Identification within video data of behaviors such as moving, remaining stationary, crossing a line, occupying a space or being part of a crowd.
- Filter by Color – Identification of objects of a specified color.
- Filter by Size – Identification of objects based upon their size.
- Filter by defined Time Ranges – Identification of objects or events based upon the time of detection.
- Search on selected cameras or group of cameras – Detection of objects or events within the field of view of a single specified camera or group of cameras.
- Search for Similar Targets – Identification within a video data set of objects having similar attributes.

These functions are most commonly used to support the following analyses:

- Accurate, wide-ranging statistical data related to people and vehicles.
- Multiple viewing options for statistical analysis of traffic volumes (people / vehicles), including numerical charts and user-friendly graphs to enable traffic comparisons, aggregates and identification of traffic trends.
- Advanced visualization options of heat map and target path to analyze movement trends and motion patterns, enabling effortless comprehension of hot / cold traffic zones and dominant traffic paths.
- Easy exporting of raw data for further analysis or integration with other systems.

Digital Multimedia Content Analysis

Digital Multimedia Content Analysis (DMCA) goes one step beyond VCA by combining information extracted from multiple modalities to derive content.[2] As shown in Figure 1 below, DMCA can range from simple "snapshot" analysis of video or audio data to processes that combine analyses of video, audio, metadata, data from multiple channels, temporal analysis and application of rules-based processes. Figure 2 provides some examples of security DMCA applications.

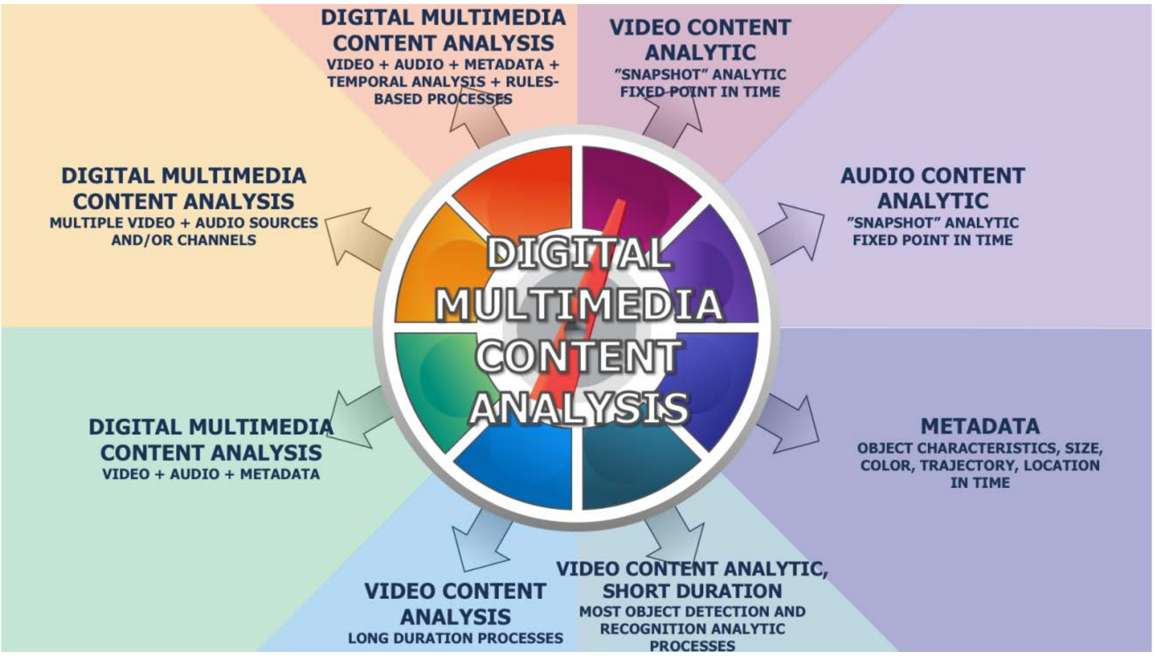

Figure 1. Security Video Analytics Application Categories.

Digital Multimedia Content is more than just video data. These digital data represent audio content, video content, metadata information, location-based information, relevant IP addresses, recording time, system time and any other information attached to a digital file. This information is valuable to the security professional both in real time and for forensic use.

Video analytics can perform complex repetitive functions, such as object detection and recognition, simultaneously on many channels of video. Video analytics tools can provide quicker and more efficient searches. Searches may be based on object characteristics and behavior, including metadata-incorporating object characteristics such as color, size, trajectory, location-based information, relevant IP addresses, recording time and system time. Some network cameras include embedded video analytics where these applications run locally.

---

[2] Dr. Alan Hanjalic, "Multimedia Content Analysis", Information and communication Theory Group, Department of Mediamatics, Delft University of Technology. http://www.cs.uu.nl/docs/vakken/mir/materials/hoorcollege/college6.pdf

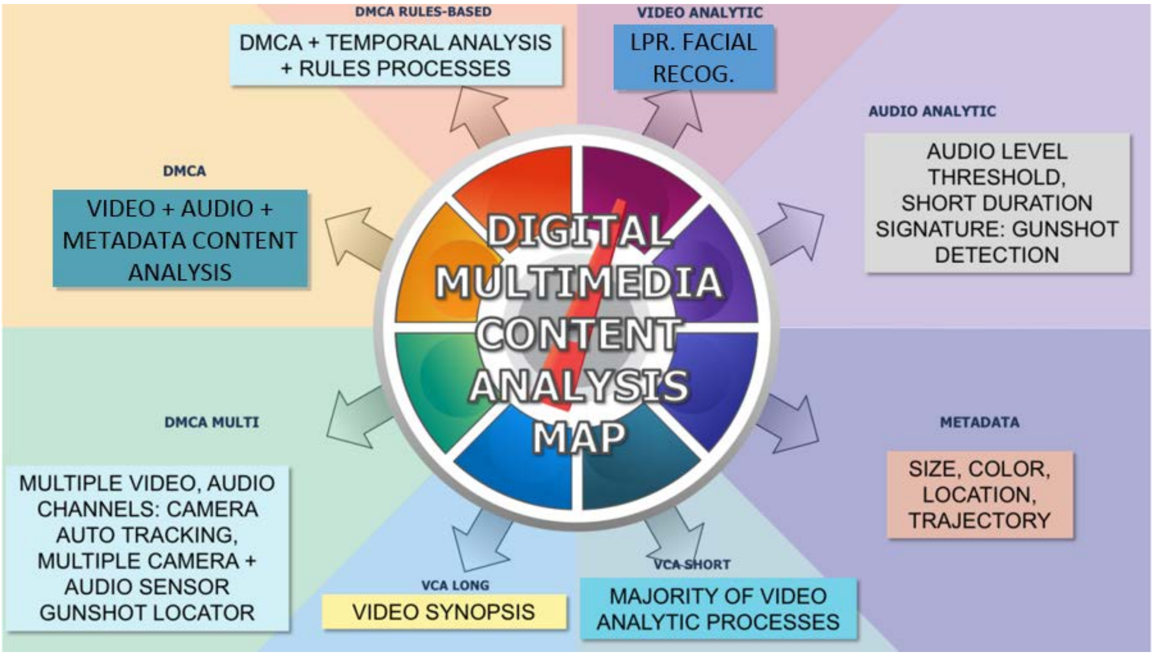

Figure 2. Security Video Analytics Application Examples.

DMCA "snapshot" or fixed point in time capabilities include thermal and color imaging. Temperature, emissivity or color signatures can be compared to support more complex processes such as license plate or facial recognition. These capabilities are currently used in applications involving access control, traffic control and toll collection.

DMCA "short duration" processing can be used to support the detection and identification of objects and behaviors. It can be used to detect specified objects (vehicles, boxes, persons), specified behaviors (motion, rest) and changes in scenes (objects introduced, objects removed, objects left behind) within a stream of video data. It can be used to support processes such as: counting of vehicles and persons; detection of persons and objects penetrating a specified boundary; detection of some complex behaviors such as tailgating, speeding or other traffic violations; or detection of objects in places where they should not be. Short duration DMCA processing can also be used to support detection of smoke and fire.

Multiple channels of video and audio detection are used in tasks such as gunshot or keyword detection.

Sophisticated DMCA techniques, incorporating information extracted via multiple modalities, can be used to achieve the following capabilities:

- Crowd detection and flow;
- Security screening queue monitoring;
- Object path analysis;
- Activity or "heat" mapping;

- Face clustering;
- Event clustering; and
- Alarm clustering.

These capabilities are frequent components of a number of business intelligence applications including:

- Customer Traffic Analysis;
- In-store Customer Behavior Analysis;
- Operational Efficiency; and
- Vehicle Traffic Analysis.

CASE STUDIES

Two topical case studies are presented in order to augment the case studies performed in support of Volume 1 of the VQIPS Handbook. The first, involving the use of analytics to support Queue Line Screening, has applicability to Airport Security. The second case study involves the use of video analytics to improve security in schools.

Example of a Video Analytic System for Queue Line Screening

Introduction

Queue line management, typically used in the retail market for "next served" or overflow redirect functions, is often used with video analytics embedded in network cameras and server-based data accumulation or analysis applications. Capabilities identified in this case study would most obviously be applicable to security screening at airports, but would also be applicable to a number of visitor, contractor, controlled substance and trusted traveler entry screening scenarios where the network camera(s) detect and alert on human throughput into a controlled opening.

Solution Requirements Guide

Entry Screening Queue Line Video Analytic Systems (ESVAS) that incorporate queue management and people counting can provide the following:

- Real-time and forensic information on the number of people entering the queue (ingress).
- Real-time and forensic information on the number of people exiting the queue (egress).
- The length of the queue, both by % of full queue and number of people in the queue.
- Wait time: the time from ingress to egress from the queue. Analytics can be used to measure the instantaneous wait time for a single person in the queue or average wait time for groups of people. By displaying continuously updated wait times, either at points located at the entrance and at various strategic points in the queue, persons in the queue can be informed regarding wait times, thus helping to manage their expectations. Alerts can be generated automatically when wait times exceed pre-defined thresholds, thereby enabling staff to make adjustments as necessary.

- Arrival rates: Video analytics can be used to determine both the average number of persons entering the queue and the average time between arrivals.
- Heat maps: identify how long persons in the queue remain in specific locations. They can be used to identify particular portions of a queue in which people are forced to “linger.” Thus, these data can be used to identify potential “choke points” in the queue and to facilitate the design of improvements.
- Identification of left behind objects: the ability to identify objects that have been left unattended in proximity to the queue.
- Facial recognition: Analytics can be used to identify persons of interest within the queue.
- Video analytics can be used to enhance the retrieval and analysis of archived data. As the quantities of archived data increase, the ability to facilitate data analysis will be increasingly valuable.
- Video analytic capabilities can be integrated with other security systems including package screening, metal detection, backscatter, hazardous substance, explosives detection, physical access control, biometric input, vehicle entry gate and fixed vehicle license plate detection systems.

In implementing an ESVAS, system designers should consider using “edge” or analytic “agents” running inside the camera. If ESVAS performance and reliability can be maintained, the application(s) may reside on a server or multiple microcomputers on the same physical network and subnet. Queue display applications should not be located on remote servers that rely on a WAN, or wireless or other connectivity subject to periodic failure.

ESVAS system designers should consider incorporation of queue management analysis so they can gather comparative metrics for different times of day at different screening locations. The user of the ESVAS should be able to configure the system to alert and capture queue line lengths based on a minimum threshold, time of day, scheduled event or an on-demand mode.

HD and Network Video: Moving Public Safety and Schools Forward in Security

One of the most significant drivers for adoption in any market is compliance. Recently, the State of Connecticut adopted recommendations as a compliance Standard based upon the “Report of the School Safety Infrastructure Council” (revised and updated to 2/4/2014), which was developed in the wake of the tragic events at the Sandy Hook school. This standard included the use of mechanical or electronic devices for observation, such as mirrors or closed circuit television (CCTV) at primary pedestrian and vehicle access and entry points.

This Standard cites the required use of guidance as provided by agencies like DHS S&T. There is a substantial amount of guidance on the DHS website (www.dhs.gov/science-and-technology/first-responders), including the guide authored by the agency’s Video Quality in Public Safety (VQiPS) Program, the Digital Video Handbook (May 2013). The focus of the VQiPS group’s efforts (and its document) is to deliver best practices on achieving video quality in public safety disciplines, including the school safety and security sector.

The Digital Video Handbook illustrates the significance of matching a required number of imaging “pixels on target” for forensic and recognition requirements. As many school surveillance systems do not use continuous observation of all cameras, the necessity for the system to be “forensic-ready” or capable of accurately reviewing critical events affecting student safety with the highest visual acuity possible can only be economically achieved through network or IP video devices augmented by the use of video content analysis.

A number of school district safety committees are developing standardized plans for implementation of controlled access rollouts. Whether they include controlling access to prevent loss, validating entries of students, faculty and staff, or locking down in the event of a crisis, the ability to integrate network cameras provide the situational awareness needed to achieve access control. The time saved by pinpointing a student’s entry or exit time using integration available with network (IP) video might just save a life.

As security measures are moving forward in many schools, there is expanded training for virtually everyone on campus. The greater simplification and availability that network video offers is making this mandatory in many cases. In one county in Alabama, every single staff member is supposed to be trained on what to do in not just an intruder situation, but in other emergency situations like a fire or severe weather, allowing for better communications.

One of several basic features of network (IP) video is the ability to distribute multiple video streams over compatible wireless or wired infrastructure. This capability gives IP video a significant advantage over “closed” analog video surveillance systems in providing situational awareness to first responders. In local schools, the focus on child protective services (e.g., protecting students from parental kidnapping) is often the driving force behind the enhanced surveillance capabilities that IP video provides.

Sporting events at school gymnasiums outfitted with compatible network video systems can distribute video streams for entertainment purposes. Cameras in cafeteria check-out lanes can be used to verify student purchases. Network video combined with access control may reduce theft of high value tablets and projectors now commonly used in schools. In the State of Illinois, one city’s school system has deployed a unique “virtual gymnasium” that uses multiple projectors to promote exercise and health in a limited space and at reduced cost. The recent gamification trend has led the same school to create a fun learning experience through the use of wirelessly connected tablets at each student workstation. The continuity of these progressive strategies could easily be compromised should theft or loss occur. Network video solutions can spot check entries to the high value storage areas and even integrate with radio frequency loss-prevention tags, alerting personnel to unauthorized technology removal.

In many cases, it has become apparent that an IT manager of the school, district or other Authority Having Jurisdiction (AHJ) is heavily involved. For value engineering and simply making use of existing infrastructure investments, this can be a significant opportunity. Network video systems are *designed* to leverage existing infrastructure to deliver power, support expansion (and contraction) and offer the value of using video data for departments other than security and safety.

In a wider geographic region, densely populated districts and counties are either guided or directed by a central AHJ. These are often separated into K-12, middle schools and Pre-K in areas requiring specialized protective and safety services. State jurisdictions prevail in high population states like New York (e.g., Office of General Services) and California, supporting centralized purchasing and specification. In some cities (e.g., New York City), there is a "School Construction Authority" providing a suite of design services to the public school system. Cases like this offer a wide opportunity for economies of scale.

The following paragraphs present some of the capabilities of IP cameras networked together and integrated with video content analysis. A typical “school day” has been developed to illustrate the application of these capabilities toward making a school environment safer.

A Day in the Life of Security and Public Safety on Campus

02:00

Early in the morning at a university, a vehicle breaches a staff parking entrance, and the driver then parks near a poorly lit loading dock and forces a service door open. License Plate Recognition software provides responding security officers with a head start because when the vehicle’s license plate does not match records in the school student, faculty, staff or contractor database, software in the gate camera immediately generates an alert. The video intercom at the service door provides an indication that the door has been breached and provides video of the suspect vehicle with a waiting accomplice still in the vehicle. The alert campus command center operator dispatches law enforcement, which arrives immediately after campus security and apprehends the suspects.

05:00

The campus day begins with the arrival of a new security officer shift. This team reviews the last shift’s events with a quick overview of several video clips, providing the incoming safety and security officers a visual overview of the previous day’s issues and additional intelligence to keep their campus safe. The previous day’s main entrance monitoring and screening overviews, including students leaving class and returning to access controlled dormitories, evening deliveries, automated faculty escorts and the one overnight breach, are reviewed in minutes using the saved metadata video management system searches. The team then begins attending their posts or their tours. All officers are equipped with tablets capable of real time video view, alarm review, incident dispatching and direct messaging to the local public safety answering point (PSAP) for all first responder categories.

07:00

Staff, student and faculty arrival builds continuously in the morning, while command center operators and the safety/security director are monitoring this activity. The use of 360° field of view High-definition television (HDTV) network cameras achieves a panoramic view of school ingress and egress areas, including a useful overview of controlled access points, main student entry and reception. “My field of view has been increased tenfold,” says the safety and security director when asked about the system’s enhanced video surveillance. “If I don’t see you coming in, I’m going to see you going out.”

12:00

The campus lunch break has the students eating both indoors and outdoors, together with a number of activity tables on the common grounds. Seeing that the lunch break has just started, the command center uses the campus public address system to direct the students to the lunch break's events. The video surveillance system monitors their responses. Operators in the command center can observe crowds congregating at events, and crowd-counting analytics can confirm attendance levels at the events.

16:00

The day is just about over for most of the student population, but not for staff, faculty and an incoming evening shift that prepares for a review of the previous day's and shift's events, again made simple through intelligent searches and embedded applications inside the network cameras. These "apps," whether license plate detection, cross line detection, student activity mapping or people counting, use the network cameras to create domain awareness sensors that relay a steady stream of data available on demand.

The incoming security crew knows when to expect the student exit activity to decrease through the video surveillance "heat" activity mapping tools. They wait until after this time to conduct the shift transition to avoid any missed incidents. Heat mapping provides the security staff with business intelligence to enable them to identify times to expect heavy congestion and light traffic at various parts of the campus, and helps them coordinate their activities accordingly.

20:00

The evening's student dorm access control entries continue. Security officers on patrol are ready to receive alerts of doors ajar or propped open. The door entries are silent until a door ajar signal buzzes and automatically resets on door close. Safety and security's command responsibility is to review these door breaches on video and verify that no suspicious activity such as "piggybacking" has taken place. The simplified alarm "histogram" guides the operator to the video associated with the door alarm. Staff uses CCTV video to confirm that the alert was the result of two students carrying in a replacement microwave oven and not an actual security breach.

Safe waiting areas around the campus show pedestrian traffic as students and faculty board transportation at prearranged locations around campus. Each one of these areas includes enhanced light-emitting diode (LED) lighting, 360° field of view HDTV network cameras, area video analytics, wireless connectivity and an audio system. Should someone be walking toward the waiting area, the LED lights increase their output as a safety indication and alert a person already waiting there. They then have the option to use their smartphone or call box as an alert device if they feel unsafe.

This time of evening finds faculty and staff walking to their vehicles and using a "video escort" application on their smartphones. Should they confirm an incident, passively report or fail to check in while they walk to their cars or dorm, campus command is immediately notified and nearby cameras activated. The system has preprogrammed camera locations that are related to the alert's location from the user's smartphone or tablet. About 20 students and faculty are using this application after hours and the system is ready to process any alerts.

22:00

Our "day in the life" concludes with the security staff verifying cafeteria and facility supply deliveries. Each of the delivering vendors has checked in online prior to delivery, entering their commercial trailer license plate and approximate delivery window. The embedded license plate recognition system automatically detects the plate on entry and exit, delivering an exception alert should the plate either not be in the database or fail to exit. The safety/security officers work together with command and make a visual verification of the delivery into the transition space. The vendor does not have to enter the secured building space for delivery, thereby simplifying and shortening the process for both parties. An HDTV network camera monitors the delivery transition area and the camera's embedded video motion detector is active after hours.

The capabilities described in the above paragraphs depend on the following capabilities:

- Design of video surveillance systems for forensic video "readiness;"
- Campus video mobility, supporting enhanced situational awareness and response;
- Image quality and video analytics supporting improved response to off-normal conditions; and
- The maturity of analytics such as license plate recognition and facial recognition.

DESIGN CONSIDERATIONS

Timeframe

In the design and implementation of a video surveillance system, a number of issues must be taken into account. The primary issue is to understand how the system will be used. As stated previously, there are three primary types of use for a surveillance system:

- Real-Time Situational Awareness & Incident Response;
- Non-Real-time Forensic Analysis/DMC Search; and
- Business Intelligence.

A critical difference in these three functions is the timeframe in which they are performed. Real-time system situational awareness requires a response time of seconds to minutes; forensic analysis has a response time of hours to days; business intelligence is a long-term on-going process.

Video Quality

The second design consideration is the required video quality. Video quality refers to the degree to which the received picture from the camera captures the physical scene or object it is recording. This is a function of the camera's capabilities and how the image is captured digitally for transmission. In general, higher quality video will require higher bit rates and will thus place greater demands on the available network resources. Desire for greater video quality must be balanced against the available bandwidth and the subsequent impact of transmitting this high quality video on other tasks sharing the IT infrastructure. Greater quality results in larger file sizes, so it can also entail large storage costs.

Video quality can be affected by lighting and coverage. Applying an understanding of the effect of light on the scene can improve the image quality of the video content. Advances in camera technology that produce usable color or "find the light" in dark- or low-illumination scenes are improving forensic video content. Using standards-based, high quality image sources like HDTV IP cameras and technologies to accommodate difficult lighting will improve the recorded image quality. Other factors that can influence the required level of video quality include the target size, the desired level of discrimination (e.g., is it sufficient to identify objects or does the system have to support identification of object characteristics as well?) and motion detection. Four levels of recognition are identified: Target Positive ID, Target Characteristics, Target Class Recognition and General Elements of the Action. More details are available in the DHS Advanced Communications Video over LTE report.[3]

In addition, the placement of cameras to achieve maximum coverage and to avoid obstructed views can greatly improve the quality of the video and its usefulness in formulating responses.

[3] DHS Advanced Communications Video Over LTE: Video Design Improvement Process. Available at https://www.dhs.gov/publication/advanced-communications-video-over-lte.

Compression

Video compression is a technique used to reduce the data size by removing redundant data in order to more efficiently transmit and store video with minimal or no impact to quality. By extension, this also reduces the amount of bandwidth required from the network to transmit that information. While video compression enables large amounts of video data to be transmitted over limited bandwidth, there are limits to how much data can be compressed without loss of significant information or video quality. There are techniques for compressing data (non-"lossy" or lossless compression) in which all of the original information is retained. There are also lossy techniques in which less important data may be filtered or encoded with lower resolution.

Video can be encoded at a constant or a variable bit rate – variable bit rates generally provide higher quality video with less overall bitrate, but they take longer to encode. Compression can use intraframe or interframe compression. Intraframe compression operates on a single frame and is less computationally intensive, but it also results in larger file sizes. Interframe compression operates on adjacent video frames. It can achieve a higher degree of data compression, but is also more computationally intensive.

Video is compressed using a device or software known as a "codec" (derived from the term coder-decoder). In general, compressed data is coded and decoded using the same codec. The source video is coded, while the receiving hardware decodes and displays the video data. The challenge with compression is to reduce the size of the data, while maintaining an acceptable level of quality.

Video transcoding is the process of coding video files using multiple codecs with various parameters in order to support end user hardware that may not be natively compatible with the original format. Transcoded data can suffer some level of degradation. However, having the ability to transcode video into multiple formats increases video compatibility with a larger number of devices, especially for systems that are required to support older devices.[4]

Compressed DMC is the most common video data available, having been transcoded from the original DMC in an industry-standard file format, so that there is a reduction in file size and thus the network bandwidth required to represent the original data set. Advances in H.264/AVC video compression, the ability to store DMC within the camera or video-encoding device itself, and virtualized or cloud computing have dramatically improved the management of video data available to investigations.

Uncompressed DMC or a copy of the original DMC in an industry-standard file format with no further compression or loss of information — although desirable by professional video-evidence examiners — is often unavailable, so this may be an unreasonable expectation due to the far larger storage requirements of uncompressed data. Bandwidth limitations can often be addressed and mitigated through the use of cameras that incorporate analytics capable of changing the compression and bandwidth based upon activity or lighting.

---

[4] DHS Advanced Communications Video Over LTE: Video Design Improvement Process. Available at https://www.dhs.gov/publication/advanced-communications-video-over-lte.

Security

A third consideration in design of a networked Video Surveillance System (VSS) is security. It is critical that the VSS cannot be accessed or compromised by unauthorized users. As the VSS will be sharing network infrastructure with other organizational Enterprise services, it is important that the VSS does not provide a conduit for malicious actors to access the broader IT infrastructure. Access control, including credentialing and other intrusion protections, needs to be designed into the system. DMC authentication and tamper detection are examples of maintaining the chain of custody for DMC evidence as specified under Law Enforcement and Emergency Services Video Association "Guidelines for the Best Practice in the Forensic Analysis of Video Evidence."

Video Content Analysis

One of the challenges of large VSS systems is that they produce prodigious amounts of data. In general, operators of these systems will not have sufficient human operators available to monitor real-time events, and retrieval of archived data will often be a time-consuming, inefficient task. The problem is exacerbated by the fact that the vast majority of data is of little or no use. This means that humans may continuously monitor hours of uninteresting data for real-time situational awareness, or forensic applications could involve sifting through hours or days of video to find a single event or a small number of events of interest. Applications that use video analytics can perform complex repetitive functions, such as object detection and recognition simultaneously on many channels of video. These tools provide improved searches by sifting through large amounts of data from a large number of cameras, searching for readily observable attributes, objects or behaviors.

Video analytics embedded in the network camera represents a growing market segment where applications run so that values or decisions based on recognition are available using the "edge" network camera and minimal software.

One popular example that can report behavior at main entry/egress points uses a "people counter," where the network camera and built-in app return the number of people passing into a zone, through a boundary or into the field of view. This information can then provide criteria on which to increase camera frame rate and stored resolution, such as during the time of highest traffic.

Another popular video-recognition solution that runs either as an embedded network camera application or in the Video Management System is fixed License Plate Recognition and Capture (LPR/LPC). This specialized app captures license plate information for immediate processing by LPR software. The software may run in a rapid-acquisition mode to collect the information and then compare plates later against an approved list, or it can perform the recognition sequentially as the vehicles pass within the camera field of view. In either case, LPR is a mature application embraced by campus safety for entry and exit locations. The trend to embed this function within the camera reduces cost and allows greater flexibility.

"Heat" activity mapping provides a visual color-coded summary showing how students, faculty and staff move about a campus. This type of video content analysis can improve safety by analyzing the flow of pedestrian and vehicular traffic on campus. Understanding personnel traffic flow will often help camera placement and ultimately the video forensic-review process.

Finally, another form of video analytics that can be useful in forensic data analysis is video summarization. With video synopsis or summarization, a condensed clip of all motion for selected criteria is continuously generated and stored, thereby allowing an “instant review” of a readily available “video synopsis.” It is possible to summarize a 24-hour period of event entries in as little as 15 minutes, reducing incident-review time by at least 50 percent. Video analytics capable of detecting abnormal scenes allows the user to set specific object criteria and direction. The scene is analyzed continuously, so that abnormal behavior or differences from the majority of the scene content are detected and annunciated or marked for later review.

Legal Constraints

A final consideration for systems that perform forensic analyses in support of law enforcement is meeting requirements for use as evidence in a trial. DMC evidence to be used in a legal setting may carry additional requirements for traceability, quality control and tamperproof assurance. The following checklists have been applied for the use of video in support of forensic analysis.

Checklist: Preparing for Digital Forensics – Issues/opportunities include the following:

- Digital multimedia content (DMC), incorporating forensic video data requirements.
- Cost assessment for preparation: balance cost-effectiveness with the technically feasible features.
- Target collection capability on the risks to the business/event/assets.
- Consider implementing ALICE: alert, lockdown, inform, counter and evacuate.
- Collect admissible evidence within the legal compliance requirements; review the legality of any monitoring (note: not a technical issue of what can be obtained through forensic video review).
- All forms of potential evidence should be considered, such as IP cameras or legacy CCTV cameras, personnel records, access control systems, still images.
- Understand the functional differences between systems: Observation, Forensic Review and Recognition/Content Analysis.
- Understand the difference between pixel density and visual acuity/image processing and how the highest quality video content is produced.
- Differentiate between video content analysis/recognition systems optimized for “pro-active” vs. “reactive;” understanding that many “reactive” tools are best for forensic video review.

Checklist: Implementing a Forensic Video Readiness Program

- Define the business/industrial/first responder scenarios that require digital multimedia content (DMC), incorporating forensic video.
- Identify available sources and different types of potential evidence.
- How will you retrieve the DMC?
- Establish the capability for securely gathering legally admissible evidence to meet the requirement.

- Establish a policy for secure storage and handling of potential evidence. Ensure monitoring is targeted to detect and deter major incidents (consider ALICE as defined above, and proactive vs. reactive technologies).
- Specify circumstances when escalation to a formal investigation (which may use the digital evidence) should be launched.
- Train staff in incident awareness so that all those involved understand their role in the digital evidence process and the legal sensitivities of evidence.
- Document a sample case with forensic video evidence; use as a model to describe incident and its impact.
- Ensure legal review to facilitate action in response to the incident.

Top Technology Considerations in Forensic Video

- Simplified high quality redundancy of recording: “edge” camera recording.
- High quality, High Definition (HD), low light tech (full color, thermal).
- “Proactive,” “Ahead of the threat” advance warning video tech (abnormality detection).
- “Reactive” video technologies help investigations (video summarization, synopsis, LPR, face recognition).
- Video + Mobility and/or Alarm Automation.

STANDARDS

Normative vs. Informative Standards – Requirements vs. Recommendations

There are two kinds of standard. Standards with a capital "S" – also known as normative standards – contain specific requirements that must be followed. Standards with a small "s" – also known as informative standards – are best practices' guidelines for achieving a specific security goal.

One example of a normative Standard would be the specifications published by the Society of Motion Picture and Television Engineers (SMPTE), which outline what qualifies as an HDTV camera. The image must conform to a 16:9 widescreen format, contain 720 or 1080 scan lines, be able to stream at 30 or 60 frames per second, and generate video in high color fidelity. An example of an informative standard would be the recommendation to install redundant local archiving as a fail-safe in case network connectivity to the remote server is disrupted.

How Standards are Created

Normative Standards are created by accredited Standards Developing Organizations (SDOs). There are a host of SDOs that play an important role in shaping electronic security. Here are just a few of them:

ASIS International

In the physical security and security applied sciences end-user community, the largest global accredited SDO is ASIS International. Their comprehensive educational programs, such as study for the Certified Protection Professional (CPP) and Physical Security Professional (PSP) credentials, are based on industry standards and guidelines. ASIS has organized its membership into regional chapters, as well as vertical markets and councils to apply each domain's standards. In the retail market, both the National Retail Federation and ASIS International work together on interpreting the significance of the Payment Card Industry Data Security Standard (PCI-DSS), which governs the payment card data security process and includes prevention, detection and appropriate reaction to security incidents.

Physical Security, Facility Security and Advanced Security Solutions like explosives detection have resulted in ASIS focusing members into the Physical Security, Security Architecture & Engineering, and the Security Applied Sciences Council (SAS). SAS and the International Information Systems Security Certification Consortium ($(ISC)^2$) are working together to deliver advanced guidance on trending solutions, such as Mobile Device Forensics, Explosives and Contraband Detection, and Active Shooter Response.

SIA

The Security Industry Association (SIA) has evolved into a significant provider of focused collaborations for industry manufacturers and solution providers. If security and safety devices are interoperable, they are more easily deployed and solutions can be scalable, agile and elastic, thereby meeting end user requirements. SIA also provides a lobbying point by bringing policy makers and stakeholders together to address federal and state initiatives affecting the security industry. SIA Education is using trends in industry standards to deliver classes on UltraHD video and Near Field Communications (NFC) at industry events. NFC-based Apple Pay and Google Wallet services allow consumers to “Tap and Pay,” while the same NFC technology is turning smartphones into electronic access control credentials.

BICSI

In the building industry, Building Industry Consulting Service International (BICSI) supports the advancement of the information and communication technology (ICT) community, which covers voice, data, electronic safety and security, project management, and audio/video technologies. BICSI recently published a Data Center design Standard to specify how to properly engineer a data center. BICSI also recently published an Electronic Security and Safety Standard (ESS), becoming the first SDO to unify physical security, physical infrastructure and safety in a single document.

ESA

Another important SDO is the Electronic Security Association (ESA) whose membership includes independent national and global systems integrators. One of its charters is to provide extensive vertical industry education to its members. Recently, they have taken a leadership role in developing Electronic Security Guidelines for Schools to ensure the safety of children, teachers and school personnel.

CSAA

The Central Station Alarm Association (CSAA) represents protection service providers, users and bureaus certified by nationally recognized testing laboratories such as Underwriters Laboratories (UL). CSAA activities and standards are encouraging the industry practices that lead to life-saving false alarm reduction and improved central station performance. Through CSAA’s Automated Secure Alarm Protocol (ASAP) to the Public Safety Answering Point (PSAP) program, the second largest Next Generation 911 center in the City of Houston can often process alarms in 15 seconds, where they previously took several minutes.

LEVA

In the law enforcement sector, the Law Enforcement Video Association (LEVA) not only publishes best practices guidelines for conducting forensic video investigation, but also offers a rigorous certification program for Forensic Video Analysts and Forensic Video Technicians. Nowhere was the value of that training more evident than during the Vancouver Stanley Cup riots in 2011, when more than 18,000 arrests were made.

SISC

In the electronic security industry, there is a unique working group known as the Security Industry Council (SISC). It reviews and coordinates the standards activities of accredited member SDOs, identifies related organization with relevant expertise for SDO assistance, and coordinates with their individual standards projects.

How do Standards Extend into the User Community?

Standards come into play on multiple levels in a security scenario. Let's use an after-hours jewelry store robbery as an example. The robbery is detected by the collaborative processing of three alarm sensors. A glass-breaking detector is activated by a breach in the glass panel of the store's front door. A remote central station monitors the audio level at the store and is able to recognize the sound of multiple people within the store. The central station also monitors the video cameras on the premises as a third verification of the burglars in action. The operator can now call law enforcement with full details of the situation and let officers know how many suspects are on the premises so that all participants can be apprehended.

Ensuring that all the critical systems function properly and in concert with one another requires adherence to multiple standards. For instance, there are standards that govern the alarm transmission and the video verification. There are protection services standards adopted by the CSAA, which are certified by a CSAA-approved Nationally Recognized Testing Laboratory (NRTL) such as UL. There are also image quality standards that make it possible to identify the suspects, such as HDTV and Ultra High Definition (also known in the industry as 2K and 4K/8K, respectively), whose specifications are governed by the Consumer Electronics Association (CEA). In addition, there are video compression standards such as H.264 (also known as Advanced Video Codec or AVC) and H.265 (also known as High Efficiency Video Codec or HEVC), which govern how the video is coded and decoded so as to reduce bandwidth consumption without degrading image quality.

Both types of standards can be found in just about every major industry. For instance, security system manufacturers rely on technology standards like H.264/AVC codecs and display resolution when developing products to ensure component interoperability.

Resolution Standards

Video data must first be digitized[5] in order to be viewed or stored. Within North America, real-time video consists of a stream containing 30 images every second (referred to as frames per second). The National Television System Committee (NSTC) set this standard, and televisions in North America receive and display video at this rate. Televisions operating in Europe receive and display 25 frames per second – a rate consistent with Phase Alternating Line (PAL) standard.[5]

Images or frames contain a set number of pixels depending on the standard. Standard definition television (referred to as 720p) consists of rectangular frames of 720 x 480 pixels. Full High Definition (HD) televisions present images consisting of 1920 x 1080 pixels.[5] The CEA defines UHD as having an aspect ratio of at least 16:9 and at least one digital input capable of carrying and presenting native video at a minimum resolution of 3840×2160 pixels. UHD video content, in 4K, has approximately four times

[5] http://www.embedded.com/design/real-time-and-performance/4013520/5/IP-video-surveillance-standards

the resolution (4 times the number of actual pixels) of 1080p full HD. 1080p content has more than twice the 720p resolution.

Table 1 provides a list of currently available video display resolutions as they have evolved from HD to UHD. Figure 3 illustrates the various display resolutions used historically, with the color of the rectangle indicating the display aspect ratio.

Name	Known As	Vertical Resolution	Horizontal Resolution	Pixel Count
HD High Definition	720	1280	720	922K pixels
FHD Full HD	1080	1920	1080	2M pixels
UHD Ultra HD	4K	3840	2160	8.3M pixels
FUHD Full UHD	8K	7680	4320	33M pixels
QUHD Quad UHD	16K	15360	8640	132M pixels

T able 1. List of currently available high definition resolutions.

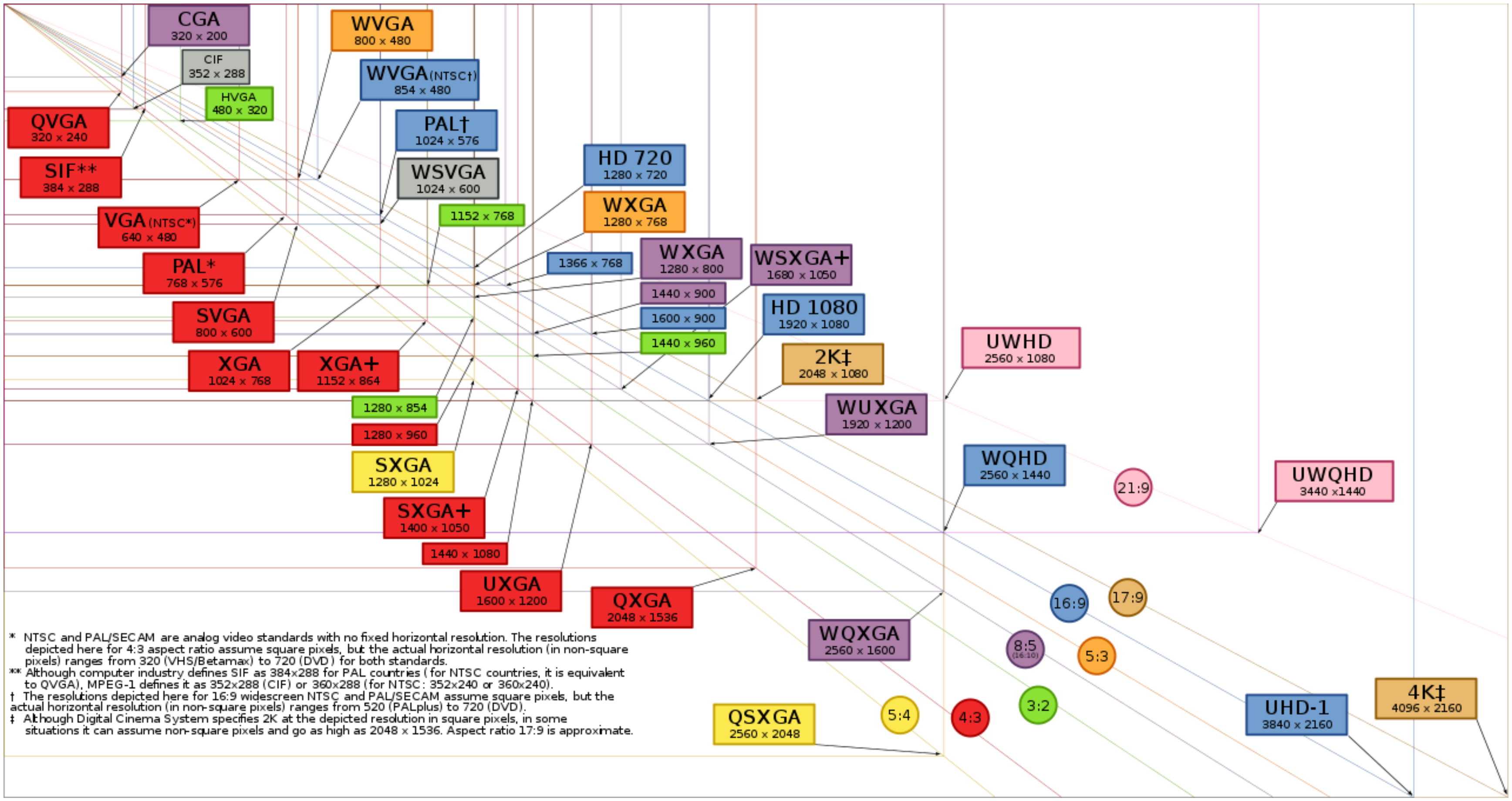

Figure 3. Currently available (as of 2016) video display resolution (from Wikipedia). The display aspect ratio is indicated by the colors, as noted in the key circles near the lower right-hand side of this figure.

Modern IP cameras and higher image quality have successfully contributed to investigations that would not be possible without today's high definition standards like HD and UHD.

UHD video formats include both 4K and 8K formats, which have advantages for surveillance. The widespread deployment of 4K video surveillance cameras not only provides increased opportunities for improved investigations, but also business intelligence. For example, personnel and customer paths are more accurately traced, and the resolution is high enough to often support multiple content analysis opportunities like people counting and vehicle license plate recognition. Additionally, due to the increased resolution, Pan-Tilt-Zoom (PTZ) can be performed directly from the image instead of mechanically.

4K resolution is especially important if the images will be viewed on larger displays. For example, given a viewing distance of 10 ft, the approximate display resolution required will increase with display size. At the 10-ft viewing distance, a 36-inch display will require at least 720p HDTV resolution, a 60-inch display will require at least 1080p, and a 100-inch display will require 4K.[6]

VSS using 4K video formats are also useful if captured video is to be used in court proceedings. Higher resolution provides improved opportunities to identify an individual, vehicle or object in real time observation and forensic review. A video source capable of supporting the identification of a person or vehicle of interest depends not only on resolution, but also the imaging hardware, image processing, lens, illumination and compression efficiency. With all parameters being equal, 4K provides four times the resolution of 1080p HDTV video sources.

Implementation of a 4K VSS requires that certain components be incorporated into the system. The most important consideration in 4K video surveillance is the source, usually an IP video 4K UHD camera. A 4K IP video camera equipped with solid-state internal storage like an SD Card can retain several hours of recorded video, depending on the storage size, even if recording server and display should fail. The second most important device is usually located near a group of IP cameras and is called a Network-Attached Storage (NAS) unit, which provides redundant recording and the opportunity to reduce bandwidth on the network. The recording and application server must be powerful enough to store, index and conduct searches of the 4K video content. If the network architect or IT designer decides not to use NAS Devices, the network infrastructure will be required to bear the 4K streams to the recording server. Internal and distributed storage provides redundancy and efficient video stream management.

Standards for Performance Testing of Camera Image Quality

UL 2802

UL 2802 is a video-imaging standard developed by Underwriters Laboratory (UL) and published in September 2013. The development included input from various stakeholders such as producers, supply chain and distributers, authorities having jurisdiction, practitioners, commercial / industrial users, government, etc. It provides progressive and objective test methods to assess the image quality of

[6] Digital Trends, 3/8/2013, “720p vs. 1080p: Can You Tell The Difference Between HDTV Resolutions?”

digital camera equipment. UL 2802 defines testing procedures and quantifies image quality based on an objective set of performance tests that are conducted on production camera samples and that measure the following nine image quality attributes:

- **Image resolution/sharpness** – measures how closely the digital image captured by the camera matches the actual image.
- **TV distortion** – quantifies the extent to which the two-dimensional image captured deviates from the actual image.
- **Relative illumination** – measures the ability of a camera to effectively capture related light intensity across an object.
- **Maximum frame rate** – measures how effectively a camera can capture a subject in motion at full resolution.
- **Sensitivity** – determines the amount of light required to digitally re-create the image as realistically as possible.
- **Veiling glare** – quantifies the impact of stray light on the camera.
- **Dynamic range** – assesses the ratio of the minimum and maximum light intensities captured by the camera.
- **Grey level** – quantifies how well a camera can differentiate areas of interest under different illumination, reflectance or luminance levels.
- **Bad pixel** – measure the level of pixel defects.

Because no single attribute or criterion can provide a reasonably accurate and objective evaluation of a given camera, lens, software, image processor, camera lens housing, electronic components or any combination of critical elements, UL 2802 uses several different quantifiable metrics to assess a video camera's performance. These metrics, along with consistent and documented test methods, eliminate any potential variations in the evaluation of a camera.

It is anticipated that UL 2802 will become one of a series of video imaging standards developed to address the complete video ecosystem.

During the test program, the manufacture and testing organization (UL) collaborates to ensure that the video camera settings are optimized and the test results reflect the camera's best abilities. The process typically involves fine tuning resolution properties, optimizing exposure time and gain during each test, enabling or disabling default features, etc., to optimize camera performance and achieve the most favorable test results based on the camera's abilities. The process of optimizing settings is no different from what is typically done during deployment of a video system.

Below are a few unique features of the new UL test program that offers a different approach from traditional test methods. All test apparatus and procedures are comprehensively detailed in the published standard.

Electronic Test Targets. Traditional methods for testing video camera image quality involve the use of standardized, published test charts for resolution and grey level tests. UL's program implements a different approach that incorporates calibrated light sources of known luminous levels, frequency, spatial orientation and grey levels. Each of these factors contributes to a video camera's ability to capture, record and store an image relative to the actual object. Using these particular electronic test targets for the test procedures eliminates some of the known potential inconsistencies associated with the use of printed test charts. Printed charts must be used under very specific lighting conditions and lighting angles, which can be difficult to control. The electronic test targets are calibrated and measured during each test program with all of the data and settings recorded for each test (Figure 4).

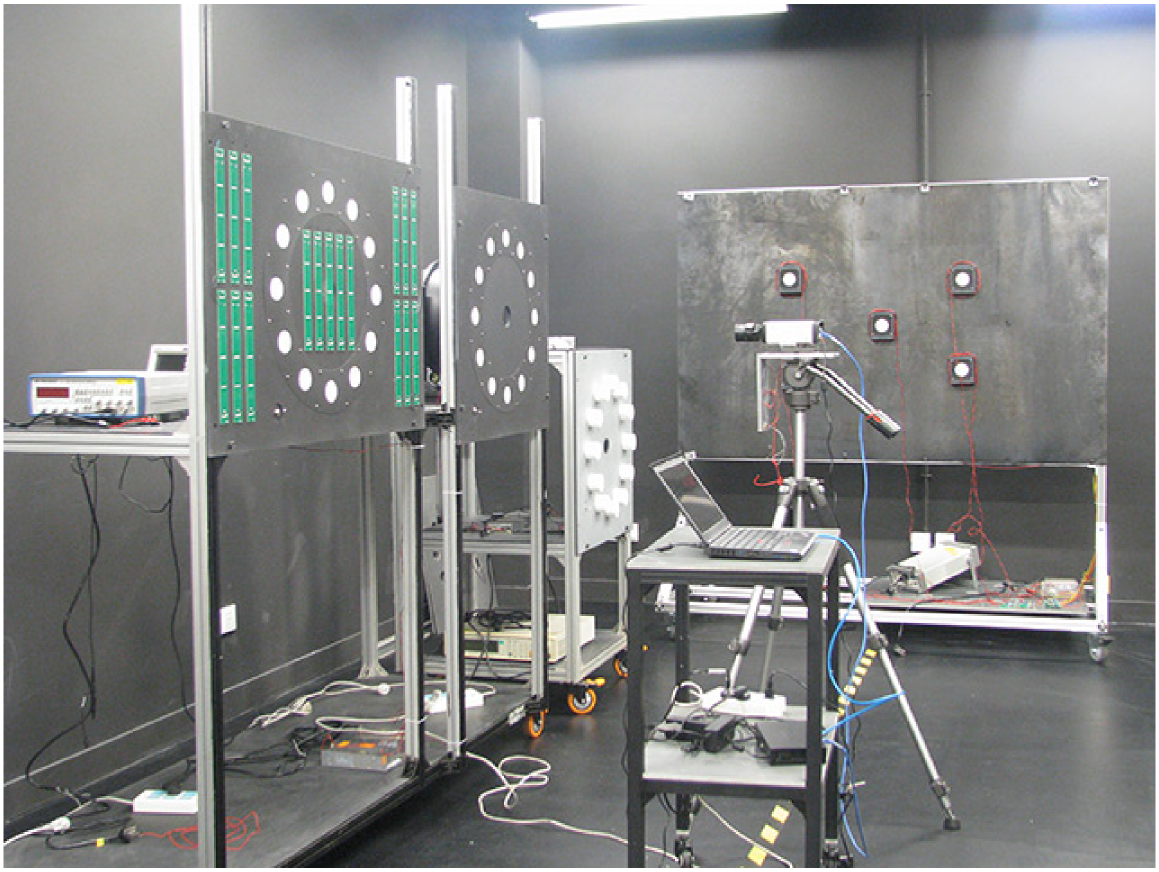

Figure 4. Test chamber with electronic test targets.

Circular Edge Analysis. Image resolution is a critical test because it is a measure of how closely the digital image captured by the camera matches the actual subject. Digital images consist of pixels, which are stored in the relative dimension of the actual image. An accurate resolution measurement requires an evaluation of a camera's lens and sensors, as well as its imaging software, and is often presented as line pairs per picture height or LP/PH. The metric is a measure of how many distinguishable alternating colors can be represented in an image.

Modulation transfer function (MTF) is a technique used to quantify image resolution in more complex images. MTF corresponds to the spatial frequency of image LP/PH; that is, the ability to represent the "real" object by taking the light intensity and plotting it along imaginary lines traversing the representation of the object.

Other mechanisms for measuring distortion are detailed in ISO 12233 (photography — Electronic still picture cameras — Resolution measurements) and in the standard mobile imaging architecture (SMIA) forum specification. UL 2802 uses the same spatial frequency response (SFR) method for resolution. The primary difference is that UL 2802 uses a circular edge analysis method (see “The Circular-edge Spatial Frequency Response Test,” by Richard Baer, Agilent Laboratories, 2002) versus a more traditional slanted-edge method to detect the transition edge of a known image (Figure 5). Benefits of the circular edge methods include an accurate method of averaging SFR from all circular directions versus multiple (horizontal and vertical) measurements using slanted-edge techniques. An added benefit is that consistency of measurements is improved using circular-edge techniques. This is because the slanted-edge SFR assumes that the transition edge of the printed resolution chart is a straight line and factors such as lens distortion can deviate the edge from a straight line, resulting in measurement inaccuracies. The benefits of using the electronic test targets and circular edge techniques for measuring SFR are cumulative in producing accurate test data.

Figure 5. Circular edge vs. slanted edge analysis.

Performance Score. Each test results in an interpolated performance score that is calculated from traditional photographic units. UL reports the scores in two usable units of measure. For simple comparisons, the UL unit scores of 0-100 make comparing one camera’s test score to another’s relatively simple without requiring an expert for interpretation. For those that are more technically involved, the traditional photographic unit scores may be more meaningful in comparing test results. Either way, the performance scores reflect the image quality achieved for each test parameter.

Performance scores will help a video system integrator make an accurate decision of what camera would be best for their particular use case based on known conditions. For example, if I need a camera that will perform well under low lighting conditions and with fast moving objects, I may want to compare camera parameters that specifically relate to those conditions, such as sensitivity, dynamic range and frame rate.

No single number or criterion can provide a reasonably accurate and objective evaluation of a given camera. Therefore, it is necessary for an integrator to consider multiple parameters based on the camera’s use case application. UL 2802 is a standard that provides objective test results based on the camera’s tested image quality.

Other UL 2802 Considerations. Unlike other UL Standards that generally focus on addressing fire and shock safety concerns, UL 2802 provides guidance for both safety and performance characteristics of cameras. Video cameras evaluated according to the performance criteria of UL 2802 must also comply with the safety requirements found in one or more other applicable product standards, including:

- UL 60950-1, the Standard for Safety of Information Technology Equipment, Safety – Part 1: General Requirements;
- UL 60065, the Standard for Safety of Audio, Video, and Similar Electronic Apparatus – Safety Requirements;
- UL 62368-1, the Standard for Safety of Audio/Video, Information and Communication Technology Equipment – Part 1: Safety Requirements; and
- UL 2044, the Standard for Safety of Commercial Closed Circuit Television Equipment.

Video cameras used in outdoor settings must also comply with the safety requirements found in UL 62368-1 60950-22, the Standard for Safety of Information Technology Equipment – Safety – Part 22: Equipment to be installed Outdoors.

This progressive camera image quality standard can assist the end user in determining the most appropriate video camera(s) for specific use cases.

Tactical Video Emerging Standards: UL 3802

UL 3802 represents a new set of standards being developed by UL and the National Fire Protection Association (NFPA) for tactical video cameras and equipment. This includes cameras that are hand-deployed, robot-mounted, body-worn, covertly placed, etc. The NFPA and UL formed the UL Standards Technical Panel 3802 in 2015 to develop performance standards for these systems. The standard will cover video cameras used by law enforcement and military in tactical operations for surveillance and situational awareness. The proposed standard will define a method for evaluating tactical system features, including image quality, audio quality, ruggedness of both the camera and associated monitoring device, length of battery operation, and remote control capabilities. As of this writing, UL 3802 was still under development.

Compression Standards

Raw video data requires a large amount of transmission bandwidth and significant storage space. In order to reduce these requirements, compression is used to decrease the video data size with the goal of minimizing the impact to video quality. Many coding/decoding (codec) standards have been developed over time. These are listed below. While not all codecs are described, the list highlights a few of the popular codecs that were previously used and currently used to support transmission of video over data networks.

+MPEG-2 Part 2[7]

The Motion Pictures Experts Group (MPEG) was formed in 1988 by the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) to set video and audio compression standards.

H.262 or MPEG-2 Part 2 (also known as the MPEG-2 Video) is the second part of the ISO/IEC MPEG-2 standard. MPEG-2 specifies that the raw frames be compressed into three kinds of frames: intra-coded frames (I-frames), predictive-coded frames (P-frames) and bi-directionally-predictive-coded frames (B-frames).

An I-frame is a compressed version of a single uncompressed (raw) frame. It takes advantage of spatial redundancy and of the inability of the eye to detect certain changes in the image. Unlike P-frames and B-frames, I-frames do not depend on data in the preceding or the following frames. Briefly, the raw frame is divided into 8 pixel by 8 pixel blocks. The data in each block are transformed by the discrete cosine transform (DCT). The result is an 8 by 8 matrix of coefficients. The transform converts spatial variations into frequency variations, but it does not change the information in the block; the original block can be recreated exactly by applying the inverse cosine transform. The advantage of doing this is that the image can now be simplified by quantizing the coefficients. Many of the coefficients, usually the higher frequency components, will then be zero. The penalty of this step is the loss of some subtle distinctions in brightness and color. Next, the quantized coefficient matrix is itself compressed via coefficients combination and Huffman run-length encoding. This encoding effectively reduces the image data amount significantly.[8]

MPEG-4 Part 2[9]

MPEG-4 Part 2, MPEG-4 Visual is a video compression format developed by MPEG. It is a discrete cosine transform compression standard, similar to previous standards such as MPEG-1 Part 2 and H.262/MPEG-2 Part 2.

MPEG-4 Part 2 is H.263 compatible in the sense that a basic H.263 bit stream is correctly decoded by an MPEG-4 Video decoder. That is, an MPEG-4 Video decoder is natively capable of decoding a basic form of H.263. In MPEG-4 Visual, there are two types of video object layers: the video object layer that provides full MPEG-4 functionality, and a reduced functionality video object layer, which is the video object layer with short headers that provides bit stream compatibility with base-line H.263.[8]

---

[7] The Motion Picture Experts Group, H.262/MPEG-2 Part 2, http://mpeg.chiariglione.org/standards/mpeg-2/video (accessed 26 October 2015).

[8] DHS Advanced Communications Video Over LTE: Video Design Improvement Process. Available at https://www.dhs.gov/publication/advanced-communications-video-over-lte.

[9] The Motion Picture Experts Group, MPEG-4 Part 2, http://mpeg.chiariglione.org/standards/mpeg-4/video (accessed 26 October 2015).

M-JPEG[10]

Motion JPEG (M-JPEG) is an intraframe-only compression scheme. Intraframe compression is less computationally intensive than interframe compression because each video frame contains all the necessary information like a single photo. In contrast, interframe compression groups adjacent video frames that then reference each other and uses the temporal redundancy to obtain increased compression. While interframe compression uses smaller file sizes than intraframe compression, it requires more processing power because the entire frame group has to be examined instead of a single frame. M-JPEG's lack of interframe prediction limits its efficiency to 1:20 or lower, depending on the tolerance to spatial artifacts in the compressed output. Because frames are compressed independently of one another, M-JPEG imposes lower processing and memory requirements on hardware devices. As a purely intraframe compression scheme, the image-quality of M-JPEG is directly a function of each video frame's static (spatial) complexity. Frames with large smooth-transitions or monotone surfaces compress well and are more likely to hold their original detail with few visible compression artifacts. Frames exhibiting complex textures, fine curves and lines (such as writing on a newspaper) are prone to exhibit DCT-artifacts, such as ringing, smudging and macro blocking. M-JPEG compressed-video is also insensitive to motion-complexity (i.e., variation over time).

This type of compression handles rapidly changing motion in the video stream well, whereas compression schemes using interframe compression can often experience unacceptable quality loss when the video content changes significantly between each frame.

M-JPEG and H.264/AVC are the two most widely adopted codec standards in IP-based surveillance cameras.[11]

RealVideo[12]

RealVideo is a proprietary video format created by RealNetworks. The early version of RealVideo was based on H.263 technology. However, after RealVideo 8 was deployed, the company switched the video codec to its proprietary video format.

RealVideo can be played from a RealMedia file or streamed over the network using the Real Time Streaming Protocol (RTSP), a standard protocol for streaming media developed by the Internet Engineering Task Force (IETF). However, RealNetworks uses RTSP only to set up and manage the

---

[10] Codec Central, Motion JPEG, https://www.siggraph.org/education/materials/HyperGraph/video/codecs/MJPEG.html (accessed 26 October 2015).

[11] DHS Advanced Communications Video Over LTE: Video Design Improvement Process. Available at https://www.dhs.gov/publication/advanced-communications-video-over-lte.

[12] Codec Central, Real Video, http://www.siggraph.org/education/materials/HyperGraph/video/codecs/ClearVideo.html (accessed 26 October 2015).

connection. The actual video data are sent with RealNetwork's proprietary Real Data Transport (RDT) protocol.[13]

VP8/VP9/VP10[14]

VP9 is an open and royalty free video streaming coding format being developed by Google. VP9 had earlier development names of Next Gen Open Video (NGOV) and VP-Next. VP9 is a successor to VP8, which is an earlier version video compression standard from Google. Chromium, Chrome, Firefox and Opera support playing VP9 video format in the Hypertext Markup Language 5 (HTML 5) video tag. VP10 is the next generation of codec that is under development by Google. VP10 is designed to have better compression efficiency than VP9 by targeting a reduction of the bandwidth requirement by half.[15]

VP9 supports compression of HDTV, 4K and above, and is used extensively in YouTube. In Figure 6, we see how a video stream can be first downloaded and stored (buffered) prior to decoding and live display. While this works very well for the video content streaming in the entertainment industry, it places less priority on high quality, real time tactical video within a limited bit rate. Google maintains that its codec is significantly more efficient than the Advanced Video Codec (AVC) or H.264. However, while it does not have the licensing issues associated with H.265/HEVC, it is currently less mature and does not achieve as high compression rates.[13]

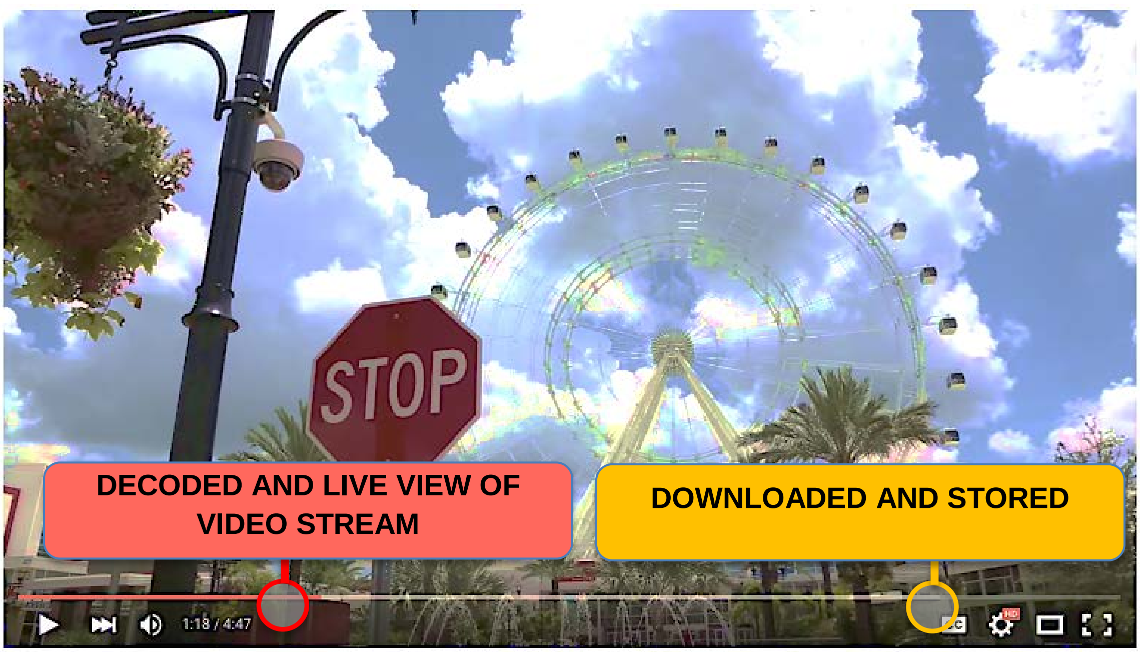

Figure 6. Video stream decoding, live view and storage with YouTube.

---

[13] DHS Advanced Communications Video Over LTE: Video Design Improvement Process. Available at https://www.dhs.gov/publication/advanced-communications-video-over-lte.
[14] VP9 Video Codec: http://www.webmproject.org/vp9/ (accessed 26 October 2015).
[15] DHS Advanced Communications Video over LTE: Efficient Network Utilization Research. Available at https://www.dhs.gov/publication/vqips-advanced-communications-video-research.

H.264/AVC

H.264, also referred to as MPEG-4 AVC, is a block-oriented motion-compensation-based video compression standard developed by the ITU-T Video Coding Experts Group (VCEG) together with the ISO/IEC JTC1 MPEG.

The H.264/AVC standard can be viewed as a "family of standards" composed of the profiles. A specific decoder decodes one or more profiles. The decoder specification describes which profiles can be decoded. H.264/AVC is normally used for lossy compression, in which the amount of loss is imperceptible by general viewers.

Recent extensions of the standard included adding five new profiles intended primarily for professional applications. One major feature added to the standard was Scalable Video Coding (SVC). SVC allows the construction of bit streams that contain sub-bit streams that also conform to the standard. This includes one bit stream known as the "base layer" that can be decoded by a H.264/AVC codec that does not support SVC.

Another major feature added to the standard was Multiview Video Coding (MVC). MVC enables the construction of bit streams that represent more than one view of a video scene. An example of this functionality is stereoscopic 3D video coding. Two profiles were developed in the MVC work: Multiview High Profile supports an arbitrary number of views, and Stereo High Profile is designed specifically for two-view stereoscopic video.[16]

HEVC (H.265)[17]

The High Efficiency Video Coding (HEVC) Standard, also known as H.265, was developed by the Joint Collaborative Team on Video Coding (JCT-VC) to increase the Advanced Video Codec (AVC, H.264) compression and efficiency, as well as to endorse the development of UHD systems. H.265/HEVC supports increased use of parallel processing architectures and effective motion vector data prediction techniques adopted to reduce bandwidth.

The goal of H.265/HEVC is to develop an encoding standard capable of supporting 4K up to 8K resolution by providing twice the compression efficiency of H.264/AVC. In H.265/HEVC, images in adjacent frames are analyzed, and divided into blocks of varying size based upon complexity. Using larger block sizes on less complex aspects of an image enable those portions of the image to be encoded more efficiently. Motion vectors can be encoded with higher precision in H.265/HEVC than H.264/AVC.[18]

Mobile providers need to conserve bandwidth to effectively deliver a quality mobile video experience. Efficiencies achieved from the H.265/HEVC codec could provide substantial improvements for

---

[16] DHS Advanced Communications Video Over LTE: Video Design Improvement Process. Available at https://www.dhs.gov/publication/advanced-communications-video-over-lte.

[17] PC Magazine Encyclopedia – definition of High Efficiency Video Coding (HEVC), http://www.pcmag.com/encyclopedia/term/63887/hevc (accessed 26 October 2015).

[18] DHS Advanced Communications Video Over LTE: Video Design Improvement Process. Available at https://www.dhs.gov/publication/advanced-communications-video-over-lte.

transmission of video on networks with limited bandwidth.[18] The replacement rate on phones is much faster than other consumer electronics, so native H.265/HEVC support is primarily focused on the smartphone industry. As of this date, Apple's iPhones 6/6S and iPhone 6/6S Plus natively support H.265/HEVC for FaceTime; Google's operating systems (Android Intel Core i7 4790KQualcomm) also include support.

H.265/HEVC encoding is highly processor intensive, which could create performance issues for network cameras that do not use efficient architectures. For example, an Intel Core i7 4790K 4 GHz processor has an AVC/x264 benchmark of 52 frames per second,[19] while H.265/HEVC[20] has only 15 fps. This basically means more dedicated central processing unit (CPU) cores to accomplish the encoding. With processor manufacturers focusing on lowering power consumption as a priority, H.265/HEVC's intensive encoding requirements may be better suited for servers or high performance network cameras dedicated to video streaming only, without additional processes like video analytics.

H.265/HEVC made its introduction at the 2016 National Association of Broadcasters (NAB) event. H.265/HEVC's significant (often 40%) efficiency over existing H.264/AVC solutions made it a compelling solution for the substantially increased bandwidth required for 4K. Adoption breeds competition.

There are, however, significant licensing issues associated with H.265/HEVC. The standard is based upon proprietary technology owned by multiple entities, and unlike for the H.264/AVC standard, no single licensing organization has yet emerged. The cost for this codec is expected to be substantially higher than for H.264/AVC, and the use of H.265/HEVC in freely distributed software is restricted.[18]

Figure 7 shows a comparison between different encoding methods versus the resulting bitrate, while still maintaining the same image quality. As can be seen from the graph, the scene encoded using H.264/AVC (baseline profile) was at least six times more efficient than a scene encoded using Motion JPEG, with a comparable level of image quality as shown in Figure 7.

[19] X264 Benchmark
[20] X265 Benchmark

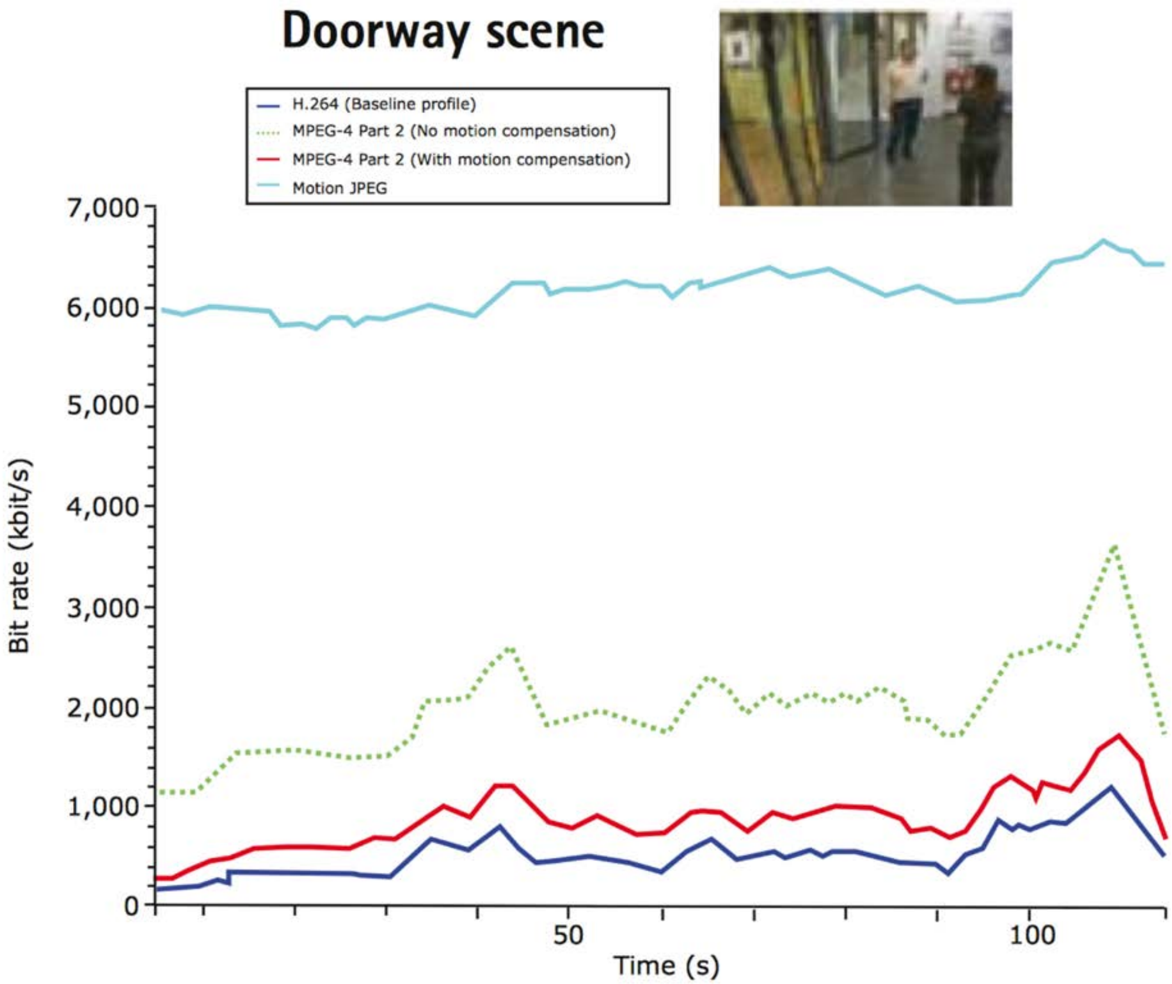

Figure 7. Comparison graph of bit rate across various encoders[21]

Security Standards

Public safety professionals are now closely examining surveillance video retention policies, best practices and agency standards. Are we keeping enough video content to perform a comprehensive investigation? Do we have a policy in place to discard video storage after a period of time? These were topics of discussion at a Public Safety Summit held by the Video Quality in Public Safety (VQIPS) working group of the DHS S&T Directorate. It was discovered that some agencies in the Los Angeles area require at least 90 days of retention for Digital Multimedia Content (video plus metadata), while an agency in

[21] Taken directly from http://www.axis.com/files/whitepaper/wp_h264_31669_en_0803_lo.pdf.

Little Rock, Arkansas, must dispose of video kept longer than 120 days. A VQiPS public safety policy team published a guide[22] to assist public safety professionals.

The following diagram (Figure 8) illustrates the participation of security and public safety ecosystem entities in various industry associations and Standards Development organizations.

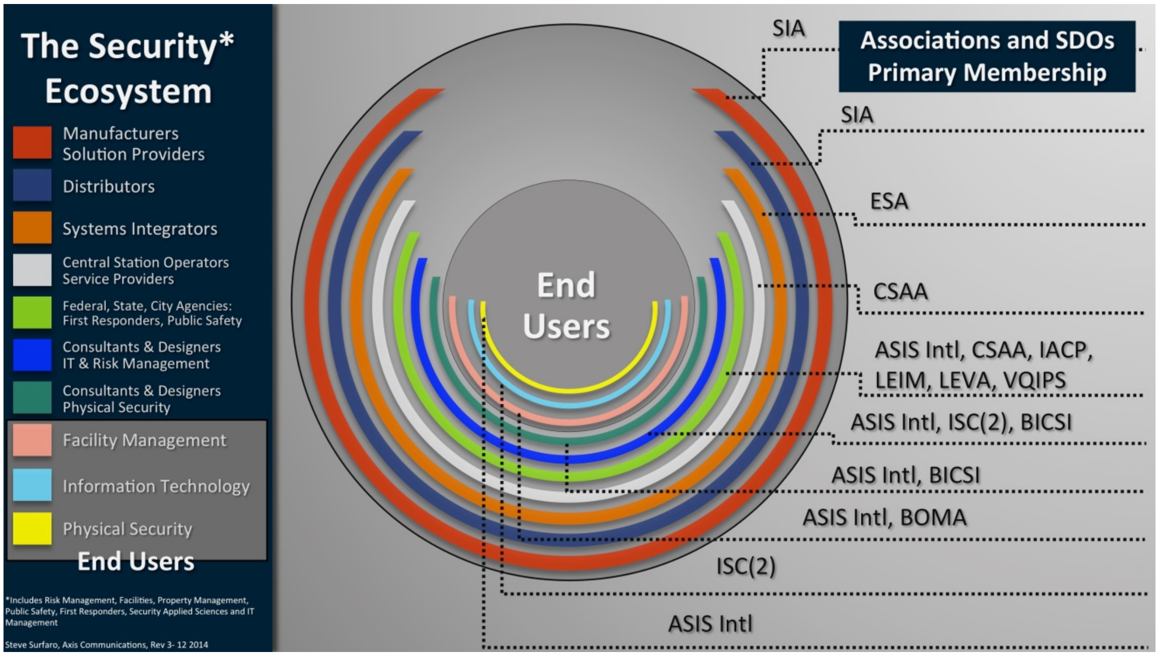

Figure 8. Participation of security and public safety entities in different standards development organizations.

[22] VQiPS – Policy Considerations for the Use of Video in Public Safety. Available at https://www.dhs.gov/publication/vqips-policy-considerations-use-video-public-safety.

IMPLEMENTATION CONSIDERATIONS

The following paragraphs present a sample project implementation plan for a network video surveillance solution if an agency chooses to seek commercial assistance.

1. Planning

During the planning phase, it is important that the contractor who will implement the system become familiarized with the operational environment and needs of the system, and that the end-user of the system understands both the capabilities and limitations of the system to be implemented. Some critical issues that will need to be resolved include the following:

a. The contractor will need a clear understanding of what the client requires and what concerns the system is intended to address. Ideally, this would include an understanding of the operational environment and existing security procedures and concepts of operation. Systems that are inconsistent with the existing environment or practices will result in less efficient operation and could entail future costs.
b. The contractor and the client will need to identify any compliance issues related to prevailing laws, safety standards or organizational practices.
c. Contractors will need to familiarize themselves with the IT infrastructure and capabilities, and understand the demands already being placed upon it. The client, and more specifically the client IT department, will need to work with the contractor to understand how the addition of the surveillance system might affect the existing IT network.
d. Contractors will need to familiarize themselves with the location infrastructure. Specifically, the contractor will need to perform a walk-through of work areas, and points of access and entry to assess how best to implement the surveillance system and the infrastructure’s ability to support the system.
e. The client and contractor should work together to identify clearly defined requirements for the system. These requirements should reflect the system functional and performance capabilities, and will serve as acceptance criteria for implementation and test of the system.
f. The contractor should provide a schedule, which identifies critical documentation deliverables.
g. The contractor and the client should agree on a schedule of status meetings and reviews to ensure project transparency.

2. Infrastructure impact

Surveillance systems employing IP cameras connected on a WAN for all their advantages will place burdens on the underlying IT infrastructure. Contractors will need to evaluate the demands that their systems will place on that infrastructure, the other demands placed on it, and develop a plan to minimize the impact of the surveillance system on other organizational functions. To achieve this, the contractor will need to work closely with the organization’s IT Department. Some critical tasks include the following:

a. The contractor will need to develop and implement an infrastructure plan.
b. The contractor will need to develop network, routing and bandwidth maps for the infrastructure.

c. The contractor will need to work with the client to develop scenarios to support measurement of bandwidth under both nominal and stressing conditions. The following steps can be followed to measure bandwidth:
    i. Note individual security device (e.g., camera) bandwidth values.
    ii. Use camera and recorder bandwidth calculators as required.
    iii. Accumulate to nearest network switch—apply totals.
    iv. Accumulate multiple network switch bandwidth—apply totals.
    v. Accumulate total security device (e.g., camera) bandwidth for each aggregation point (e.g., network video server or physical access control panel).
    vi. Apply totals.
    vii. Note individual user monitoring station bandwidth values.
    viii. Note command center bandwidth values.
    ix. Apply scenarios given in Command Center Display Scenarios chart.
    x. Accumulate to nearest network switch—apply totals.
    xi. Accumulate typical multiple user monitoring station bandwidth values for each aggregation point—apply totals.

d. The contractor will need to work with the client IT Department to define protocols, power and QoS for the network.
    i. Verify infrastructure compatibility and protocol support.
    ii. Verify QoS and desired performance (e.g., with Personal Access Communication System (PACS), credential acceptance delay; with IP Video, refresh rate, control delay, stream quality, individual image capture acuity).
    iii. Consider that cable infrastructure needs to support bandwidth capacity requirement.
    iv. Consider that cable installation and cable quality significantly impact data rate.
    v. Verify design of wiring plant topology and network switches needed to support placement of security and surveillance system devices.

e. The contractor will need to identify methods for provision of system power. This should include evaluating the potential use of Power over Ethernet devices and designing resiliency into the system.

3. Analytic Capabilities

For larger systems, the contractor should consider implementation of a modeling and simulation capability for the network. Ideally, contractors would have a capability to model and evaluate network conditions and loads under the broadest range of scenarios. Models should capture recording system utilization and network switching functions. They should also be able to facilitate assessment of performance during failure and recovery scenarios, and be scalable to support analysis of additional components, edge devices and infrastructure.

4. Deliverables

a. The contractor will need to develop a commissioning plan to control deployment of devices on the network.

b. The contractor will need to ensure that a full suite of system diagrams is available and accessible. Required diagrams will include the following:
    i. Security block diagram.
    ii. Data closet design.
    iii. Security device type schedule and bill of materials.
    iv. Security device type detail.
    v. Riser diagrams.
    vi. Point-to-point diagrams.
    vii. Command center elevations and stretch-out.
    viii. Command center sequence of operations by scenario
c. The contractor should provide a master Bill of Materials (BOM) including the following:
    i. Security device BOM.
    ii. Telecom Room BOM.
    iii. Command center BOM.
    iv. Workstation BOM.

5. Acceptance Testing

In collaboration with the client, the contractor shall develop and execute a plan for acceptance testing of the system. The test plan should provide a framework to enable the client to adequately review plans for verification (and validation) of system requirements, identify long-lead test items and facilities, identify roles and responsibilities, and define opportunities for the client to observe and formally accept the system. Testing of the system should be done incrementally, but should include a formal acceptance test to be performed following deployment and check-out of the system by the contractor (acceptance may be preceded by a brief period of operational use, in which problems can be detected and mitigated). Additional aspects of the acceptance test include the following:

a. Acceptance testing should include testing individual devices for operation, scenarios and system responses.
b. All access control points, intrusion detection points, video cameras and intercom systems require testing and observation to ensure that they operate as required in the construction documents.
c. In order to account for all lighting conditions, video cameras must be examined during the day and at night.
d. Acceptance testing may also include a “defeat the system” test to demonstrate that there are no potential shortcomings within the hardware and software system that would compromise the integrity of the system under normal operating conditions. Acceptance testing is to be completed and test documentation approved by the client prior to the project completion.
e. Testing should include provisions to verify that performance meets compliance requirements.
f. Testing should also include tests to validate the video data meets legal forensic and evidentiary requirements.

6. Training

The contractor should plan on providing training to ensure successful operation of the system.