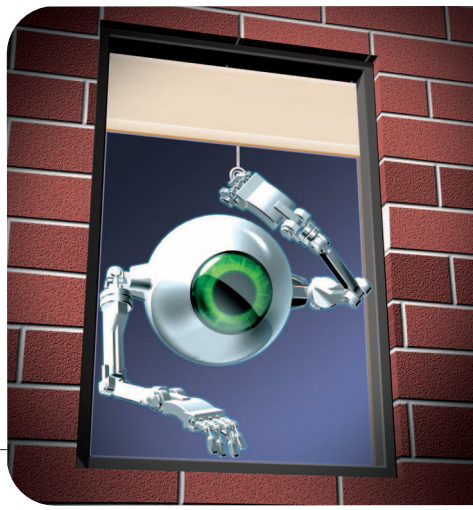
Enabling Video Privacy through Computer Vision

Closed-circuit television cameras used today for surveillance sometimes enable privacy intrusion. The authors' privacy console manages operator access to different versions of video-derived data according to access-control lists. Additionally, their PrivacyCam is a smart camera that produces a video stream with privacy-intrusive information already removed.



Andrew Senior, Sharath Pankanti, Arun Hampapur, Lisa Brown, Ying-Li Tian, Ahmet Ekin, Jonathan Connell, Chiao Fe Shu, and Max Lu
IBM T.J. Watson Research Center

In recent years, closed-circuit television (CCTV) cameras have gained widespread use worldwide. We're now seeing a corresponding rise in video-processing systems that use computer-vision algorithms to mine and interpret video data, such as movement, subject identities, and event times. Typically, CCTV surveillance conjures up images of nuclear reactor installations or secretive defense operations, but low-cost cameras are enabling a wide range of applications that put cameras on devices and in buildings and public spaces. Because human operators monitor CCTV systems, unobtrusive or deliberately hidden cameras allow spying and voyeurism, and video surveillance can be a tool for state control and oppression. When the algorithms currently under development mature and achieve human-like efficiency, their sheer scale will immeasurably increase CCTV systems' power in both benign and malignant applications.

This Pandora's box of automated video surveillance has already been opened, but some developing technologies could help control video surveillance's negative uses. We can use algorithms similar to those that extract data from raw video to filter that video, alter it, and restrict the amount of privacy-intrusive data it contains, while preserving enough useful information to complete the original task.

We've built a prototype system to record and redistribute surveillance video in a way that minimizes privacy intrusion. We've also embodied privacy-protection principles in PrivacyCam—a camera with on-board processing that produces video streams with the privacy-intrusive information already removed. We can use PrivacyCam for various automated, video-dependent applications, as well as surveillance. We hope that, along with social and legal controls preventing oppressive surveillance, we can apply these technological methods to "blinker big brother" and restrict CCTV's abilities to invade privacy.

The rise of video surveillance

Video surveillance is becoming ubiquitous in urban life. Video cameras are being installed in urban areas throughout the developed world, principally to deter crime. Theoretically, people won't commit crimes (or will commit them elsewhere) because they fear being caught by active surveillance or being identified later from video recordings. Rachel Armitage and her colleagues list 10 ways CCTV deployment reduced reported crime, although the actual effects seem limited.[1] Brandon Welsh and David Farrington,[2] after conducting a metastudy of 22 CCTV studies, found that using CCTVs contributed to an average of 4 percent reduction in crime.

The use of surveillance is spreading as the hardware becomes more affordable; video camera prices have tumbled in recent years as technology has improved and production quantities have increased. Similarly, video-storage costs have fallen as video recorders have become commodity items; digital storage is becoming less expensive, even as it provides higher quality than analog video. Finally, installation costs have fallen and should fall further as wireless networks obviate the need for cables.

Public concerns

In general, those being watched have tolerated or even welcomed the increase in video surveillance due to the

perceived public-safety and crime-fighting benefits. However, dissenting voices have always pointed out video surveillance's potential to enable abuse and privacy invasion. Recent technological developments and the threat of blanket video surveillance have heightened these fears and led to growing public concern about mass surveillance's less benign effects.

The American Civil Liberties Union (ACLU) describes five types of abuse related to CCTV:

- criminal,
- institutional,
- abuse for personal purposes,
- discriminatory targeting, and
- voyeurism.

A British study of video surveillance found that young, male, or black subjects were systematically and disproportionately targeted for "no obvious reason."[3] The study also found that some operators used video surveillance for voyeurism, and that no one used it to watch over those at risk, saying that operators don't look out for possible victims, but focus on stereotypical categories of those they think might be likely offenders. Women were also more likely to be objects of voyeuristic rather than protective surveillance. The ACLU isn't alone in concluding that video surveillance's benefits and risks are disproportionate.

Automated surveillance

Mike McCahill and Clive Norris[4] estimate that more than 4 million CCTV cameras were in operation as of 2003. Many of these are rarely monitored and of poor quality, installed as deterrents without much regard for practical use. However, automatic surveillance-video processing will result in constant monitoring, recording, and indexing of all video signals. Some CCTV systems have already publicly deployed face-recognition software, which can potentially identify, and thus track, people as effectively as CCTV systems recognize cars today.

Many groups are developing software tools that will automatically "watch" and "understand" surveillance videos. These systems can potentially gather richer information about the video subjects, judge their actions and behaviors, and aggregate this data across days, or even lifetimes. They magnify the potential for video surveillance, taking an expensive, labor-intensive operation with patchy coverage and poor recall and turning it into an efficient, automated system that observes everything in front any of its cameras and lets that data be reviewed instantly and mined in new ways. They could, for example, track a particular person throughout the day, monitor a time of day over a long period, or look for people or vehicles returning to a location or reappearing at related locations.

Algorithms exist for tracking people, understanding their interactions, determining which way they're looking, and so on. Others, such as compression algorithms, have reduced surveillance storage needs. Further algorithms could let us automatically track individuals across multiple cameras with tireless, uninterrupted monitoring across visible and nonvisible wavelengths. Such computer systems could potentially process thousands of video streams—whether from cameras installed for this purpose or from preinstalled private CCTV systems—resulting in blanket, *omnivident* (all-seeing) surveillance networks.

What is video privacy?

The current explosion in video camera deployment by governments, corporations, and individuals, together with new technologies that can exploit that video, forces us to ask what protections we might enact to protect individuals' privacy.

This problem is challenging because privacy means different things to different people. Some believe privacy protection is a right and obligation, while others assume that those wanting privacy protection must have something to hide.[5] David Brin argues that we can never completely preserve privacy, and he suggests how society can deal with this.[6] Peter Danielson views video surveillance ethics as "a continuously modifiable practice of social practice and agreement."[7] Cultural attitudes determine what's acceptable or intrusive in video privacy (Danielson contrasts attitudes in the UK and Canada), but so does technological capability. A US General Accounting Office report[8] quotes the 10th Circuit Court of Appeals' decision to uphold the use of surveillance cameras on a public street without a warrant on the grounds that any activity a person knowingly exposes to the public isn't subject to Fourth Amendment protection, and therefore isn't constitutionally protected from observation. However, technology (with capabilities such as high zooms, automatic control, relentless monitoring, night vision, and long-term analysis) lets surveillance systems record and analyze much more than we might believe we're exposing to the public. Some believe this chilling effect of video surveillance infringes on US Fourth Amendment rights (protection against unreasonable search and seizure).

In spite of such objections, surveillance technology will inevitably spread to include blanket coverage of urban areas by fixed cameras and possibly even tiny aerial vehicles that could go anywhere carrying cameras. This infrastructure is inevitable, but Brin suggests a choice:[6] we can entrust authorities with surveillance access, as we do currently (and as did Oceania in George Orwell's *1984*), or we can democratize access to surveillance mechanisms and adapt these same tools to "watch the watchers" and protect the populace against abuses of power.

A model for video privacy

To help us develop a model for protecting individual pri-

Absolute vs. relative ID

A major distinction among video surveillance systems that correlates with how likely they are to intrude on privacy is the level of anonymity they afford. We distinguish three types of systems:

- *Anonymous.* A typical CCTV system without computer augmentation is anonymous—it knows nothing about the individuals that are recorded on the tape or visible on the monitors. Although operators could abuse such a system, it doesn't facilitate that abuse.
- *Absolute ID.* These systems can identify the individuals observed (using face-recognition technology, for example, or a badge swipe correlated with the video) and associate them with a personal record in a database. Such systems require an enrollment process that registers people in the database and links their personal information (such as name or social security number) with their identifying characteristics (such as face image or badge number).
- *Relative ID.* These systems can recognize people they've seen before, but don't require enrollment. They can collect statistics about people's comings and goings, but don't know any individual information. A relative ID system might use weaker identification methods (such as clothing colors) to collect short-term statistics as people pass from one camera to another, but they can't recognize people for longer than a day.

Clearly, anonymity protects an individual's privacy. Someone could make an absolute ID system report on a specific individual's movements at the end of each day. With manual enrollment, someone could retrospectively convert a relative ID system with a strong identifier into an absolute ID.

vacy, we look at six aspects crucial to privacy in video surveillance systems.

What data is present?

The fundamental determinant of video privacy is the information that the surveillance captures and conveys. At a basic level, the camera system's design should limit the data captured to only those areas where surveillance is needed and not intrusive. Blinkers, blinds, or physical stops on a steerable camera's motions, as well as lens caps and on-off indicators could both restrict a camera's field of view and reassure the public that surveillance is bounded. A low-resolution camera or deliberately defocussed lens would further help protect privacy while preserving the system's usefulness. Finally, the system design can fundamentally restrict how individual identities are represented (see the "Absolute vs. relative ID" sidebar).

Has the subject given consent?

If the subject willingly consents to observation, privacy is less of an issue, but consent can vary from consciously choosing to walk in front of a camera to walking in front of one because no other option exists to being captured on hidden camera or spied on where you can reasonably expect privacy (in a hospital or at home, for example). People usually give consent with the understanding of having certain privacy protections.

Consent is hard to achieve in public places. Signs often inform the public that they're being observed (often to intentionally subdue them into good behavior), but generally those deploying CCTV don't seek consent. A future system might use facial information from the video to let subjects access portions of the data relevant to themselves. This could automate access to the video guaranteed through freedom of information provisions (as in the UK's Data Protection Act) and detect privacy clashes.

What form does the data take?

Is the data on tapes that might be removed? Is it transmitted over a network that could be tapped? Is it encrypted? Is the information accessible to a single person, or does it require multiple keys?

Encrypted data that's stored digitally will ensure that stealing surveillance tapes or eavesdropping on transmissions doesn't allow access to the video's content. Indeed, the camera itself should encrypt information to prevent eavesdropping at any stage.

Who sees the data?

Is the data restricted to the police? To security professionals? What procedures exist for enforcing the policy? How well is the data protected against hackers, burglars, or subpoena?

In addition to the physical and procedural controls that data-protection laws already require (see the "Legislating privacy protection" sidebar), we propose requiring operators to access data through a series of controls. For example, operators could be allowed to access the encrypted video only through a decoding console (see Figure 1) using an approved key; they would also need a user key to access the unencrypted data obtained using a cryptographic key and the encrypted video. Additionally, access-control rules would detail who could view what data under what circumstances, and additional restrictions would exist, such as key sharing to require multiple authorizations. Operations such as playback, searching, freeze frame, and so on could require different levels of authorization.

Legislating video privacy protection

In many legal systems, video privacy falls under legislation dealing with general data privacy and thence data protection. In the European Union, for instance, directive 95/46/EC covers this, and came into force in March 2000. In the UK, the relevant legislation is the 1998 Data Protection Act (DPA), which outlines the principles of data protection, saying that data must be

- fairly and lawfully processed;
- processed for limited purposes;
- adequate, relevant, and not excessive;
- accurate;
- not kept longer than necessary;
- processed in accordance with the data subject's rights;
- secure; and
- not transferred to countries without adequate protection.

The act requires operators to register all CCTV systems with the Information Commissioner (see www.crimereduction.gov.uk/cctv9.htm) and gives specific requirements for protecting privacy in a CCTV system, such as preventing unauthorized staff from accessing CCTV monitors. Mike Cahill and Clive Norris have estimated that 80 percent of CCTV systems in London's business district don't comply with the DPA.[1]

The act also guarantees individuals the right to access information held about them, including CCTV recordings, but protects the privacy of other subjects who might have been recorded at the same time. The European Convention on Human Rights guarantees an individual's right to privacy (www.crimereduction.gov.uk/cctv13.htm) and further constrains the use of video surveillance, most explicitly by public authorities.

References

1. R. Armitage, G. Smyth, and K. Pease, "Burnley CCTV Evaluation," *Surveillance of Public Space: CCTV, Street Lighting, and Crime Prevention*, K. Painter and N. Tilley, eds., Crime Prevention Studies, vol. 10, Criminal Justice Press, 1999.

How long is data kept?

Not only does the length of storage limit how long someone can use the data, but it also limits how many people can see it. A significant difference exists between systems that present video feeds for synchronous review by guards and those that store the video or other information derived from it.

With viewing managed through the privacy console we described in the previous section, we could use keys to manage the data's lifetime independent of how long encrypted (and perhaps illicit) copies existed.

How raw is the data?

Raw video on a tape is unwieldy and difficult to use without significant resources. However, adding a time stamp or other metadata makes the video more accessible and consequently more likely to intrude on privacy. Operators might store metadata even if they discard the original video, letting users search for information without the video.

This aspect of video privacy is the most crucial, and the one we can most effectively enhance; a video-processing system could mask privacy-invasive features, to prevent it from being processed to extract metadata.

Privacy-preserving video console

Our prototype video console concentrates on the data present and the rawness of that data, and uses conventional technologies (encryption and access-control lists, for example) to deal with the other issues we outlined in the model. The system rerenders the video stream to hide privacy-intrusive details while preserving the information the operator needs for the system to be useful.

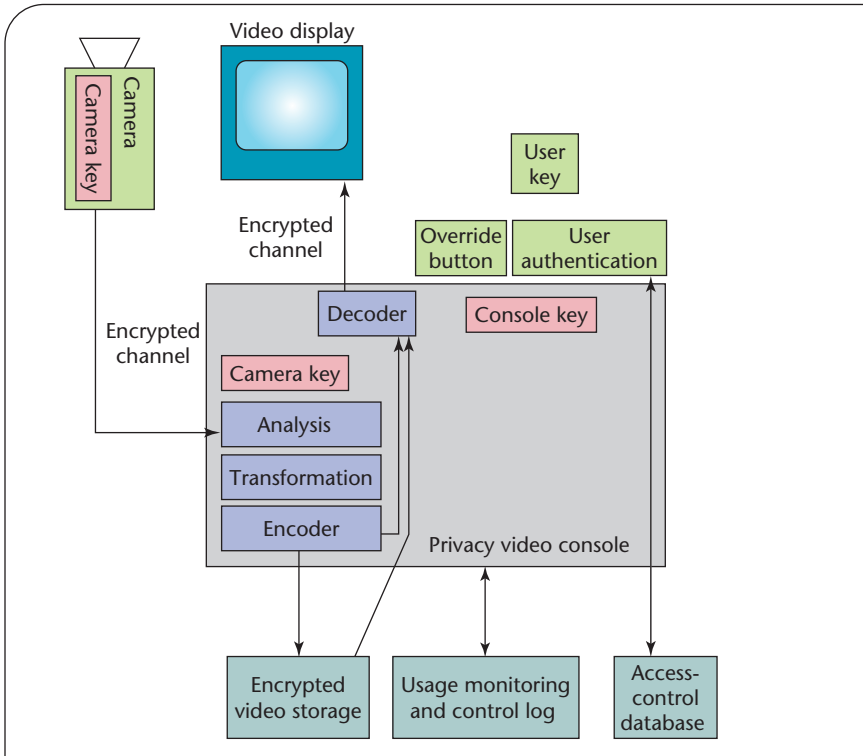


Figure 1. Privacy-preserving console. Selective dissemination and display of video information occurs through automatic video processing and analysis. The original video channel is securely analyzed for features that might reveal individual information.

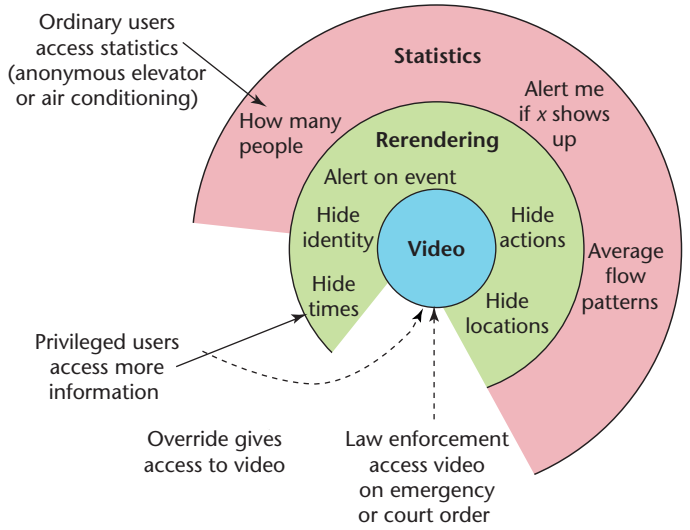


Figure 2. A layered approach to the presentation of surveillance video according to access controls. The three layers provide access to three types of users: ordinary users can access statistical information about the video; privileged users can access limited individual information; and law enforcement agencies can access raw video information and related individual identity information.

System architecture

Our basic premise is that the system can automatically extract various information components from the video content; different system users can access these components based on their authorization levels.

Figure 1 shows a block diagram of the complete system. At the highest level, the architecture consists of a selective video-encoding system transmitting an encoded video signal to a selective video-decoding system, either live or with intermediate storage. The decoding (and possibly encoding) system requires user authentication.

The encoding system consists of video analysis, transformation, and encryption subsystems. The video analysis subsystem takes one or more live or recorded video streams and analyzes them at successively more sophisticated levels to extract separate information streams—for example, the background's appearance or different moving objects' various attributes. The transformation subsystem selectively transforms the extracted information based on the system policy. Finally, the encryption subsystem encrypts the transformed, extracted information. Thus encoded, the video might contain multiple copies of essentially the same information, although each copy might use a different key. The encoded information stream could include an encrypted version of the raw video.

The decoding system consists of user-interface and video-synthesis subsystems. The former establishes the system operator's identity using the user authentication module (with token-, knowledge-, or biometrics-based authentication) and lets the operator apply selective operations (for example, query, view, freeze-frame, or export to analogue tape) to the synthesized video or unencrypted video information. The video synthesis module uses the keys and authorization from the user authentication system to decrypt the encoded video information received from the encoding system (or storage). The operator could then use the decrypted video information both to reconstruct the (transformed) video and to answer queries. All the operator-accessed information and operator actions are securely logged.

Video analysis

We propose a layered approach to granting access to the different kinds of data that the system extracts. Depending on the authorization level, the interface might let the user access the original raw video or even video enhanced with additional information, or it might present only reconstructed video with details deliberately obscured. It might give the user just statistical information, such as the number of people in a space. Determining what information to give which users depends on the situation and the types of users, but we provide a set of tools and basic algorithms that handle the most common cases. Figure 2 shows one way to layer access: law enforcement officials can subpoena the original video, whereas security guards can see only (identity-obscured) rerendered video, unless they use an override button, which logs the time and the video footage appearing at that time. Other registered users might be able to access other information, and anonymous users can inquire about statistics. Additionally, a device might register as a user—for instance, an elevator control computer might be able to access how many people are standing in front of the elevator doors.

The algorithms in the analysis subsystem use computer vision techniques to "understand" the video; they extract interesting objects, distinguish between background and foreground, separate people from vehicles, distinguish people walking in groups; and even distinguish between a person's different limbs. (More information on implementing video analysis algorithms is available.[9])

The system can detect objects of interest using a *generic object detection* approach or a *model-specific* approach. With the former, the system defines objects according to the video image sequence's attributes (features). For example, the system considers objects interesting when they differ from some archetypal background model. It then differentiates a detected generic object into specific object categories (or a false alarm) using a more specific model for each object.

With a model-specific approach, the system models and explicitly detects each object using model-based techniques. For example, in a surveillance application, an operator might be predominantly interested in humans

and vehicles. The system might detect these objects using specific models of the human head, the face, and vehicle types. We can also combine these two strategies.

Once the system detects an object in the video, it infers that object's type, identity, location, and pose by further processing the video and the context. The object's locus in successive frames determines the object track. Changes in the object over time can help the system infer activity; relating the activity of multiple objects defines a group interaction. Thus, the features extracted through video analysis can richly represent the video sequence's content.

To transform the video content, the transformation subsystem first selects an information component from the video, then obscures that piece of information, or its complement (subjects and their surrounding locales, for instance, are complements in the video). For example, a particular system policy might dictate that location information in the video be completely erased. Another policy might require that all faces in the video be obscured so that only gender information (but not identity, age, or expression, for example) is available. More generally, the system policies might require partially or fully obscuring or statistically perturbing certain components of the extracted information, such as a subject's location, pose, activity, track, and so on. Some simple global operations (such as noise, jitter, color desaturation, blurring, or time and space downsampling) might help prevent an unauthorized user from effectively machine-processing the video stream. The transformed information components constitute an encoded video channel. Figure 3 shows some selection and obscuration methods.

To protect and restrict the information in our system, we follow a three-pronged approach, using transformation, summarization, and encryption to deliver only authorized information to users. The transformation subsystem implements an obscuration policy for one channel; users can access the channel, but they can't access the obscured information, which is irrevocably lost and can't be recovered from that channel. However, a user might have access to some information (albeit at different levels of detail) through multiple channels, so we must design our system to ensure that users can't use information from multiple streams (perhaps from different colluding users) to reconstruct data. The statistical query processor summarizes and delivers information of even less sensitivity (for example, it will let the processor know that there are typically 20 people in the parking lot at 8 a.m. based, statistically, on how many people the video captures over a particular time period). The encryption and decryption processes protect all information channels from tampering and interception. We can combine user authentication with transformation to let subjects access video of themselves, while still using obscuration to protect other subjects' privacy.

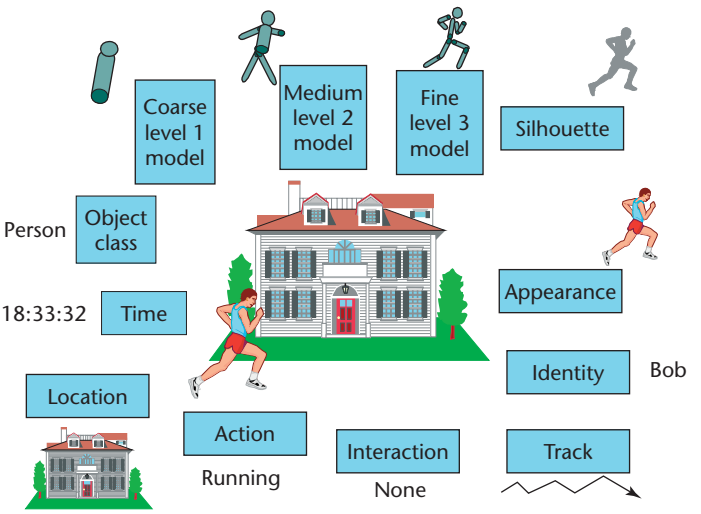


Figure 3. Video rerendering options. The central image shows a video frame surrounded by different methods of rerendering video, which the operator, if permitted, can select individually or use together. If such an operation isn't permitted, the operator can only access statistical information (for example, when is the street most crowded).

PrivacyCam

The PrivacyCam is a standalone implementation of some concepts that we've described in the privacy-preserving video console section. In our version, the camera outputs a rerendered National Television Systems Committee (NTSC) video stream, so the PrivacyCam can replace a standard CCTV camera, but with privacy-preserving features built in. In this case, the camera has on-board processing power, so the video transformation, summarization, and encryption occur on the camera before transmission. The on-board processor implements any processing algorithms available in the privacy console. In addition to the rerendered video stream, PrivacyCam can also transmit encrypted information streams via other output ports, such as over a wireless network. We can limit the camera to produce only one type of video stream, or we can integrate it with a privacy console to let authenticated users see the original video or some other information stream.

Guaranteeing video privacy

Video information processing systems, including the system we've outlined (see the related work sidebar for other examples), are error prone. Perfect performance isn't guaranteed, even under fairly benign operating conditions. The system makes two types of errors when separating video into streams: missed detection (of an event or object) and false alarms. These errors counterbalance each other: an operating point with high sensitivity has

Related work on privacy-protecting technologies

Little work has been done to protect video privacy beyond creating legislation (principally in Europe) that describes how to run a CCTV system and what to do with the data. In recent years, a few companies and researchers have attempted to use technology to protect surveillance subjects' privacy.

A Sony patent[1] describes a system that detects skin tone and replaces it with another color to hide surveillance subjects' race to avoid discrimination. Matsushita has patented a system that obscures a "privacy region"—that is, camera views or portions of camera views that depict possibly sensitive information such as private property—being observed by a pan-tilt-zoom camera.[2] Elaine Newton and her colleagues recently described a system for de-identifying faces by transforming them in shared surveillance video so that they can't be picked up by a face recognition system.[3]

References

1. A.M. Berger, *Privacy Mode for Acquisition Cameras and Camcorders*, US patent 6,067,399 to Sony Corp., Patent and Trademark Office, 2000.
2. J. Wada et al., *Monitor Camera System and Method of Displaying Pictures from Monitor Camera Thereof*, European patent EP 1 081 955 A2 to Matsushita Electric Industrial, European Patent Office, 2001.
3. E. Newton, L. Sweeney, and B. Malin, *Preserving Privacy by De-Identifying Facial Images*, tech. report CMU-CS-03-119, Carnegie Mellon Univ., 2003.

few missed detections but many false alarms, whereas one with low sensitivity has few false alarms but can fail to detect events when they occur.

To alleviate these problems, we must select the appropriate system operating point. The costs of missed detection and false alarms are significantly different, and differ for privacy protection from those for a surveillance system. Given the information's sensitive nature, a single missed detection would likely reveal personal information over extended time periods. For example, failing to detect a face in a single video frame could reveal unobscured identity information, compromising days of aggregated track information associated with the supposedly anonymous individual. On the other hand, an occasional false alarm (for example, obscuring something that isn't a face) has a limited impact on the system's effectiveness. The operating point can be part of the access-control structure—a higher authority might suggest reducing the false alarm rate, increasing the risk of compromising privacy.

Even perfect detection can't guarantee anonymity, however. Contextual information might be enough for an operator to uniquely identify a subject even when all identifying characteristics are obscured in the video. Obscuring biometrics (face or gait, for example) and weak identifiers (such as, height, pace length, or clothing color) will nevertheless reduce the potential for privacy intrusion. In general, these privacy-protection algorithms, even when operating imperfectly, will make it harder, if not impossible, to run automatic algorithms to extract privacy-intrusive information, making abuses by human operators more difficult or costly.

Increasing public acceptance

Obviously, the techniques we've described are optional for CCTV systems, and will both cost more and risk impinging on the offered surveillance's effectiveness. So, why would anyone accept this extra burden? The main reason is likely to be legislation. In the future, regulating authorities might require CCTV systems to impose privacy protection similar to what we've described. Indeed, we could argue that existing legislation would require deploying these techniques as they become commercially available.

Without legislation, companies and institutions deploying CCTV might still choose (possibly due to pressure from the public, shareholders, or customers) to include privacy-protecting technology in their surveillance systems. Liability for privacy infringement could encourage such a movement.

We must also ask, however, what guarantee a citizen has that a claimed privacy protection is actually in force? Legislating public access to surveillance systems as Brin proposes[6] is one solution, but that still begs the question—does some data exist that's not open to the public? A potential solution is system certification and registration, perhaps similar to the system that's evolved for Internet privacy (see www.TRUSTe.org, for example). Video system vendors might invite some independent body to certify their privacy-protection systems. For purpose-built devices with dedicated camera sensors (such as PrivacyCam), this would suffice. Individual surveillance installations could also be certified to comply with installation and operating procedures, and operators could display certification of the privacy protection offered on the equipment and CCTV advisory notices. Such notices might include a site (or even camera) identification number and the surveillance privacy registrar's URL, where people could confirm the system's certification.

Video privacy is an exciting and burgeoning research area. Our team is further developing the core video analysis issues and video privacy prototypes we presented in this article (www.research.ibm.com/people

vision/videoprivacy.html).[10] More specifically, our recent research focuses on determining sophisticated video analysis technology that will detect objects in increasingly complex environments, such as cluttered scenes, and difficult imaging situations, such as snow. We plan to harness the increased sophistication in automatic video processing to provide more effective video privacy in realistic deployment situations. We strongly believe that it is merely a matter of time before video privacy technologies will be pervasively used to both prevent abuse of sensitive video data and to share useful impersonal information acquired using video surveillance networks. □

References

1. R. Armitage, G. Smyth, and K. Pease, "Burnley CCTV Evaluation," *Surveillance of Public Space: CCTV, Street Lighting, and Crime Prevention*, K. Painter and N. Tilley, eds., Crime Prevention Studies, vol. 10, Criminal Justice Press, 1999.
2. B.C. Welsh and D.P. Farrington, *Crime Prevention Effect of Closed Circuit Television*, research study 252, UK Home Office, 2002.
3. C. Norris and G. Armstrong, *The Maximum Surveillance Society*, Berg, 1999.
4. M. McCahill and C. Norris, *CCTV*, Perpetuity Press, 2003.
5. M. Caloyannides, "Society Cannot Function without Privacy," *IEEE Security & Privacy*, vol. 1, no. 3, 2003, pp. 84–86.
6. D. Brin, *The Transparent Society: Will Technology Force Us to Choose Between Privacy and Freedom?*, Perseus Publishing, 1999.
7. P. Danielson, "Video Surveillance for the Rest of Us," *Proc. IEEE Int'l Symp. Technology and Society*, IEEE Press, pp. 162–167.
8. R. Stana, *Video Surveillance*, tech. report GAO-03-748, US General Accounting Office, 2003.
9. A. Hampapur et al., "Face Cataloger," *Proc. IEEE Conf. Advanced Video and Signal Based Surveillance*, IEEE Press, 2003, pp. 13–20.
10. A. Hampapur et al., "Multiscale Tracking for Smart Video Surveillance," *Signal Processing*, vol. 22, no. 2, 2005.

Andrew Senior *is a research staff member at the IBM T.J. Watson Research Center, where he has worked in the areas of speech, handwriting, audio-visual speech, face and fingerprint recognition, and, most recently, video privacy and visual tracking. His research interests include pattern recognition, computer vision, and visual art. Senior received a PhD in engineering from the University of Cambridge. He is a senior member of the IEEE. Contact him at aws@us.ibm.com.*

Arun Hampapur *manages the Exploratory Computer Vision Group at the IBM T.J. Watson Research Center. His research interests include smart video surveillance, visual-tracking systems, and visual-data management. Hampapur obtained a PhD in electronic engineering and computer science from the University of Michigan. He serves on the program committees of several IEEE and ACM conferences and is a senior member of the IEEE. Contact him at arunh@us.ibm.com.*

Sharath Pankanti *is a research staff member at the IBM T.J. Watson Research Center. His research interests include computer vision and pattern recognition. Pankanti received a PhD in computer science from Michigan State University. He is a senior member of the IEEE. Contact him at sharat@us.ibm.com.*

Yingli Tian *is a research staff member at the IBM T.J. Watson Research Center. Her current research interests include computer vision problems, from motion detection and analysis to human identification, 3D reconstruction, facial expression analysis, and video surveillance. Tian received a BS and an MS in electronic engineering from TianJin University, China, and a PhD in electronic engineering from the Chinese University of Hong Kong. She is a senior member of the IEEE. Contact her at yltian@us.ibm.com.*

Lisa Brown *is a research staff member in the IBM T.J Watson Research Center's Exploratory Computer Vision Group. Her primary research interests include head tracking, pose estimation, and, more recently, object classification. Brown received a PhD in computer science from Columbia University. She is a senior member of the IEEE. Contact her at lisabr@us.ibm.com.*

Ahmet Ekin *is a research staff member in the Video Processing group at Philips Research, Eindhoven, the Netherlands. His research interests include statistical object detection, data fusion, and video enhancement. Ekin received a PhD in electrical and computer engineering from the University of Rochester. He is a member of the IEEE Signal Processing and Computer societies. Contact him at ahmet.ekin@philips.com.*

Jonathan Connell *is a research staff member at the IBM T.J. Watson Research Center, where he has worked on mobile robot navigation, reinforcement learning, vision-based object segmentation and recognition, audio-visual speech processing, and biometrics. His research interests include robotics, computer vision, and artificial intelligence. Connell received a PhD in computer science from MIT. He is a senior member of the IEEE. Contact him at jconnell@us.ibm.com.*

Chiao-Fe Shu *is an architect, programmer, and researcher and cofounder of Virage, where he develops commercial applications based on content-based retrieval technology. His research covers oriented texture-pattern analysis, classification, and segmentation, in-situ wafer inspection systems based on Fourier imaging, and multimedia indexing and retrieval. Shu received a PhD in electronic engineering and computer science from the University of Michigan. Contact him at cfshu@us.ibm.com.*

Max Lu *is a contracting senior software engineer at the IBM T.J. Watson Research Center. His research includes building complex systems, such as video surveillance systems, machine tool computer numeric control systems, image-based 3D model capture systems, and wafer defect detecting systems. Lu received a BE in electronic engineering from HuaZhong University of Science and Technology and an MS in pattern recognition from the National Laboratory of Pattern Recognition (NLPR), Chinese Academy of Sciences. Contact him at maxlu@us.ibm.com.*