Chapter 27
Privacy Protection and Face Recognition

Andrew W. Senior and Sharathchandra Pankanti

27.1 Introduction

Digital imagery—from personal cameras, cellphones, surveillance cameras and television—is now ubiquitous, being used by governments, corporations and individuals to bring undreamed-of new capabilities for entertainment and government or commercial services. This pervasiveness, together with the new technologies for analyzing and exploiting such images, lead us to ask what are the risks to privacy thus created or exacerbated, and what protections are, or could be put, in place to protect individuals' privacy. Visual privacy has been of concern since the invention of photography, but the issues are becoming critical as digital imagery is used more widely. At the same time, image processing, computer vision and cryptography techniques are, for the first time, able to deliver technological solutions to some visual privacy problems.

In this chapter, we describe the privacy issues surrounding the proliferation of digital imagery, particularly of faces, in surveillance video, online photo-sharing, medical records and online navigable street imagery. We highlight the growing capacity for computer systems to process, recognize and index face images and outline some of the techniques that have been used to protect privacy while supporting ongoing innovation and growth in the applications of digital imagery.

We first examine what is meant by privacy, and visual privacy in particular, focusing on the privacy concerns surrounding facial images. In Sect. 27.2, we examine some of the factors that determine visual privacy, and in Sect. 27.3 we summarize particular domains in which visual privacy is important. Section 27.4 describes tech-

A.W. Senior (✉)
Google Research, New York, NY 10011, USA
e-mail: andrewsenior@google.com

S. Pankanti
IBM Research, Yorktown Heights, NY 10598, USA
e-mail: sharat@us.ibm.com

S.Z. Li, A.K. Jain (eds.), *Handbook of Face Recognition*,
DOI 10.1007/978-0-85729-932-1_27,

nologies for protecting privacy in images and Sect. 27.5 presents three systems that have been developed for applying privacy enhancing technologies to face images.

27.1.1 What is Privacy?

The problem of protecting privacy is ill-posed in the sense that privacy means different things to different people [1], and attitudes to its protection vary from the belief that this is a right and obligation, to an assumption that anyone demanding privacy must have something to hide [11]. Just as it is difficult to define privacy, it is difficult to determine when privacy has been intruded upon. Equally, many people are happy to trade in their intangible privacy for only small incentives [53], whereas others guard their privacy jealously, so it is hard to determine the value of privacy and privacy intrusions. There is a continuum of privacy intrusion, and our comfort point on that continuum can easily be displaced, by a small incentive or a bout of media hype. A range of factors come into play and our personal tolerances all vary with those factors, from less-than-flattering photos on the Internet to images that form part of our medical record.

In many applications where there are privacy concerns, it is hard to point to examples where there is material effect on a person "who has done nothing wrong", yet the feeling of disquiet remains perhaps because everyone has done something "wrong", whether in a personal or legal sense. The area where the public is perhaps most concerned over privacy is video surveillance, with fears aroused by authoritarian governments, and science fiction like *1984* or *Minority Report*. Few people wish a society where all its laws (speeding, parking, jaywalking...) are enforced absolutely rigidly, never mind arbitrarily, by an omniscient state. There is always the possibility that a government to which we give such powers may begin to move towards authoritarianism and apply them towards ends that we do not endorse.

Danielson [16] views the ethics of video surveillance as "a continuously modifiable practice of social practice and agreement". What is considered acceptable or intrusive in video privacy is a result of cultural attitudes (Danielson contrasts attitudes in the UK and Canada) but also technological capability. A report of the US General Accounting Office [54] quotes the 10th Circuit Court of Appeals decision to uphold the use of surveillance cameras on a public street without a warrant on grounds that "activity a person knowingly exposes to the public is not a subject of Fourth Amendment protection, and thus, is not constitutionally protected from observation." However technology (with capabilities such as high zooms, automatic control, relentless monitoring, night vision and long term analysis) enables surveillance systems to record and analyze much more than we might naturally believe we are "exposing to the public". It has been argued that the "chilling" effect of video surveillance is an infringement of US first amendment rights.

Brin, in "The Transparent Society" [10], argues that at some level privacy cannot be preserved and suggests that in the face of inevitable ubiquitous surveillance, our only choice is whether to leave this surveillance in the hands of the authorities

or democratize access to the surveillance mechanisms and use these same tools to "watch the watchers" and so protect the populace against abuses of the tremendous power that the surveillance apparatus affords.

27.1.2 Visual Privacy vs. General Data Privacy

In many legal systems, visual privacy falls under the legislation dealing with general data privacy and thence data protection. In the European Union, for instance, this is covered by EU directive 95/46/EC which is enacted by member states in their own legislation and came into force in March 2000. In the United Kingdom, with perhaps the densest video surveillance, 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.
- Not transferred to countries without adequate protection.

The act requires all CCTV systems to be registered with the Information Commissioner, extending the 1984 Data Protection act that only required registration of CCTV systems that involved "Automatic Processing" of the data. It further gives specific requirements on proper procedure in a CCTV system in order to protect privacy:

> Users of CCTV systems must prevent unauthorized access to CCTV control rooms/areas; all visitors must be authorized and recorded in the visitors log and have signed the confidentiality proforma. Operators/staff must be trained in equipment use and tape management. They should also be fully aware of the Codes of Practice and Procedures for the system. The observation of the data by a third party is to be prevented for example, no unauthorized staff must see the CCTV monitors.

It has been estimated [33] that 80% of CCTV systems in London's business district are not compliant with the DPA.

The act also guarantees the individual's right of access to information held about them, which extends to access to CCTV recordings of the individual, with protections on the privacy of other individuals who may have been recorded at the same time.[1]

[1] "The DPA supports the right of the individual to a copy of any personal data held about them. Therefore data controllers are obliged to provide a copy of the tape if the individual can prove that they are identifiable on the tape, and they provide enough detail to locate the image (e.g., 1 hour

The European Convention on Human Rights guarantees the individual's right to privacy[2] and further constrains the use of video surveillance, most explicitly constraining its use by public authorities. The Swiss Federal Data Protection Commissioner has published these guidelines: [57]

> When private individuals use video cameras, for example to protect individuals or prevent material damage, this is subject to the federal law of 19th June 1992 on data protection (DPL; SR 235.1) when the images filmed show identified or identifiable individuals. This applies irrespective of whether the images are stored or not. The processing of the images—such as acquisition, release, immediate or subsequent viewing or archiving—must comply with the general principles of data protection.

A big difference between ordinary data privacy and image privacy is the amorphous nature of images, and the difficulty in processing them automatically to extract useful information. A video clip can convey negligible amounts of information or may contain very detailed and specific information (about times, a person's appearance, actions). Privacy is hard to define, even for explicit textual information such as name, address and social security number fields in a database, knowledge of which can be used for identity theft, fraud and the mining of copious information about the individual from other databases. It becomes much harder to assess the privacy-intrusion that might result from the unstructured but potentially very rich information that could be harvested from surveillance video. A simple video of a person passing in front of a surveillance camera by itself affords little power over the individual, except in a few rare circumstances (such as proving or invalidating an alibi).

There are already strong restrictions on the use of microphones for surveillance because of the presumption of privacy of conversations, but video has been less restricted because there is an expectation of being observed when entering a public space. The UK DPA exempts from controls data where, "The information contained in the personal data has been made public as a result of steps deliberately taken by the data subject." While the act of walking along, the street could be construed as deliberate steps to make ones visual appearance public, we have seen that the DPA does provide privacy safeguards for CCTV.

Until recently, the unmanageability of images has limited their potential for abuse. Few photographs were online, and those that were only manually labeled, and mostly of celebrities, for whom privacy is handled somewhat differently. It takes time to review surveillance video to find "interesting" excerpts, and the storage requirements have added to privacy reasons to ensure that recordings are retained for only short periods of time. Long term storage, and detailed analysis have been

---

before/after the time they believe they were captured by CCTV, their location and what identifiable features to look for). They must submit an appropriate application to the Data Controller and pay a £10 fee. However, the request can be refused if there are additional data/images on the tape relating to a third party. These additional images must be blurred or pixelated out, if shown to a third party. A good example would be a car accident where one party is attempting to claim against another. The data controller is obliged to say no to a civil request to view the tape, as consideration must be given to the other party. A request by the police is a different matter though."

[2] See http://www.crimereduction.gov.uk/cctv13.htm.

reserved for situations with strong economic or forensic motivation. However, the advent of sophisticated computer algorithms to automate the extraction of data from images and video, means that imagery is becoming as easy to mine and interrelate as a queryable, machine-readable database.

27.2 Factors in Visual Privacy

The goal of privacy protection is to prevent access to information that intrudes on an individual's privacy, but specifying exactly what information is sensitive is difficult. For the purposes of this chapter, we limit ourselves to considering images of faces, though certainly other visual information (e.g., documents, the presence of an object, other biometric identifiers) can compromise privacy in certain circumstances. Identification of individuals is the major threat to privacy in images, and facial appearance is the most commonly captured and easily recognized biometric in images. In Sect. 27.3, we review specific domains where face images are handled. Here, we consider a number of factors (listed in Table 27.1) that play a role in the privacy-intrusiveness of automatic video surveillance, the most complex of these domains. Most systems must be designed to operate under multiple combinations of these factors, requiring multiple levels of privacy protection.

The location of the camera is certainly a principal factor. In high security surveillance environments, no privacy protection may be necessary, but to some, in the home no level of video obfuscation may be considered acceptable. The person with access to the information also determines the level of privacy-intrusiveness, as shown by [30]. A person from any of the categories of Table 27.1 may be familiar with an individual observed by the system, increasing the risk of information being viewed as sensitive, but an unfamiliar person is still subject to voyeurism and prejudiced treatment. In each category the availability of each type of data must be limited as far as possible, consistent with the person's need to access information. The person seen by the camera also plays a role, being observed with some kind of informed consent (e.g., an employee); with active consent, or denial of such, perhaps expressed through the carrying of a privacy token (Sect. 27.4.4); passively as a member of the public; or indeed as an intruder.

In preventing privacy breaches from a surveillance system, we must review the information that can be leaked, the access points to that information within the system, and the availability to different groups of people. Raw video contains much privacy-intrusive information, but much effort is required to get to that information. A key frame may convey much less information, but if well-chosen presents information succinctly. An index with powerful search facilities can easily direct a user to a particular clip of video. The power to intrude on privacy is greatly enhanced if the system has the capability to identify individuals (Sect. 27.2.1). While in principle, all the information stored in a networked digital system is vulnerable to hacking, such breaches are defended against and their effects minimized, by conventional information and physical security, for instance strict access controls should be in

Table 27.1 Factors affecting privacy protection in a video surveillance system

Scenario	Observer	Familiarity	Role of subject
High security	Law enforcement	Familiar	Member of general public
Low security e.g., workplace	System managers	Unfamiliar	Employee
	System operators		(Non-)consenting subject
Public space	Authorized accessors		Wearer of privacy tag
	Public		Intruder
Private space	Hackers		
	Person observed		

Effort	Data type	Tools
Passive	Raw video/image	Summary
Opportunistic	Redacted video/image	Video review
Deliberate	Extracted metadata	Freeze-frame
Sophisticated.	Anonymized data	Search
	Linked to an identity	Biometric ID
		Weak identifier

place to limit access to privacy-sensitive information, this information can be always encrypted when stored or transmitted, and there may be audit trails to record who accessed what data under what circumstances.

27.2.1 Absolute and Relative Identification

A major distinction that we have drawn for privacy in surveillance systems [51], that significantly correlates with how likely they are to intrude on privacy, is the level of anonymity they afford. We distinguish three types of system: *Anonymous*, *Relative ID*, and *Absolute ID*:

- **Anonymous** A traditional CCTV system without computer augmentation is anonymous—it knows nothing about the individuals who are recorded onto the tape or presented on the monitors. While open to abuse by individuals watching the video, it does not facilitate that abuse in a systematic way.
- **Absolute ID** These systems have some method of identifying the individuals observed (usually face recognition but also such identifiers as a badge swipe correlated with the video) and associating them with a personal record in a database. Such systems require some kind of enrollment process [7] to register the person in the database and link the personal information (such as name, social security number) with the identifying characteristic (face image or badge number), though the enrollment can happen without the knowledge or consent of the subject.

- **Relative ID** These systems can recognize people they have seen before, but have no enrollment step. Such systems can be used to collect statistics about people's comings and goings, but do not know any individual information. A relative ID system may use weaker methods of identification (such as clothing colors) to collect short term statistics as people pass from one camera to another, but be unable to recognize people over periods of time longer than a day, or use face recognition without any external label.

Clearly, anonymity protects the individual's privacy. An absolute ID system might, for instance be made to "Give a report on the movements of Joe Bloggs at the end of each day". A relative ID system with a "strong identifier" can easily be converted retrospectively into an Absolute ID with a manual enrollment, and as the availability of labeled data on the web increases, it is becoming easier to partially automate that enrollment. Extracting Relative or Absolute ID from an Anonymous system would require storing and reprocessing the data.

27.3 Explosion of Digital Imagery

In this section, we review some of the domains in which the privacy of face images is important.

27.3.1 Video Surveillance

CCTV deployment is undoubtedly expanding rapidly. In 2003, McCahill and Norris [33] estimated that there were more than 4 million CCTV cameras in operation in the UK. At the time, most such CCTV systems were rarely monitored and of poor quality, installed largely as a deterrent. Automatic processing of surveillance video, however, is bringing a new era of CCTV with constant monitoring, recording and indexing of all video signals.

Many groups around the world [6, 8, 26, 29, 34, 55] are developing software tools to automate and facilitate the task of "watching" and understanding surveillance videos. These systems also have the potential for gathering much richer information about the people being observed, as well as beginning to make judgments about their actions and behaviors, as well as aggregating this data across days, or even lifetimes. It is these systems that magnify the potential for video surveillance, taking it from an expensive, labor-intensive operation with patchy coverage and poor recall, to an efficient, automated system that observes everything in front of any of its cameras, and allows all that data to be reviewed instantly and mined in new ways: tracking a particular person throughout the day; showing what happens at a particular time of day over a long period; looking for people or vehicles who return to a location, or reappear at related locations. This brings great power in the prevention and solution of crimes.

Some CCTV systems have already publicly deployed face recognition software which has the potential for identifying, and thus tracking, people as effectively as cars are recognized today (for instance, the London Congestion Charging scheme [13]). Currently, face recognition technology is limited to operating on relatively small databases or under good conditions with compliant subjects [42]. Further algorithms bring the potential to automatically track individuals across multiple cameras, with tireless uninterrupted monitoring, across visible and non-visible wavelengths. Such computer systems may in future be able to process many thousands of video streams—whether from cameras installed for this purpose by a single body, public webcams [56] or preinstalled private CCTV systems [38]—resulting in blanket, *omnivident* surveillance networks.

Yu et al. [63], in work supported by the US Department of Justice, describe one potential future direction for higher-level learning based on face recognition. They show how automatically captured location tracks and face images from fixed and steerable cameras can be used to learn graphs of social networks in groups of people, particularly targeted at identifying gangs and their leaders in prisons.

27.3.1.1 Camera-Based Sensors

While surveillance has driven the widespread deployment of cameras, low cost sensors and more sophisticated algorithms are enabling many other applications that involve the installation of cameras that will see people, but in which it is not the images themselves that are of interest, but rather the data extracted from them. These range from today's traffic cameras and cameras that anticipate drownings in swimming pools [44] to "human aware" buildings that adjust heating, lighting [32], elevators and telephones according to the locations and activities of people, as well as controlling physical access and assisting with speech recognition by lip-reading [45]. Many future devices and systems will have cameras installed because they are a low-cost sensor that "sees the world as humans see it". While the purpose of these sensors is often merely to detect a single piece of information, such as the number of people at a check-out line [60], the same hardware could equally be used for surveillance and face recognition. It is impossible for the subjects of the observation to know what is happening to the data once it has left the sensor, so without suitable oversight these devices are a potential and perceived privacy intrusion.

27.3.1.2 Ambient Video Connections

Some of the earliest work on image-based privacy relates to the use of video for ambient awareness in media spaces, particularly video for awareness of co-workers in a remote location. Here, a worker may choose to be shown in a constant video feed to provide a sense of copresence. However, in times when there is no explicit face-to-face conversation the worker may wish to reveal only general information, such as presence or general location without revealing specific details that would be visible in a full-resolution video. Such a privacy protection system that uses model-based face obscuration is described in Sect. 27.5.2.

27.3.2 Medical Images

Medical images are also proliferating, with the advances in medical science and the lowering cost of imaging devices. Much attention has been paid to the electronic patient record and its privacy implications. The ability to copy and transmit sensitive patient records in electronic form as well as access them remotely, together with the increasing richness of the records has led to stricter controls on medical record privacy, such as the HIPAA [58] regulations in the USA. These regulate medical records as a whole, but photographs of patients that show their faces are of specific concern here. Face images may be an important component of a patient record for such areas as oral and maxillofacial surgery, dentistry, dermatology and plastic and reconstructive surgery. It is important to protect the patient from exposure of the data both through unauthorized access and use for teaching or research material. It is essential in the latter case to remove identifying information while preserving the usefulness and accuracy for the intended purpose. De-identifying faces (Sect. 27.5.2) is an important technique here.

27.3.3 Online Photograph Sharing

Photo-sharing has recently become one of the most popular activities on the Internet, featuring in social networking sites like Facebook, and special photo storage and sharing sites like Flickr or photobucket.com. Billions of photographs are stored by such services.[3] As traffic has grown, the affordances for labeling have become more sophisticated. Text tagging has evolved to labeled bounding boxes and to the automatic face recognition found in Picasa and Windows Live Photo Gallery. Now the task of labeling a photo album has been made much easier by software which allows the user to name a person in one picture and then propagate that label to other similar photos of the same person. These new labels can be confirmed or corrected and the face model is improved accordingly, so a large photo collection can be iteratively labeled with relatively little manual intervention.

Companies such as PolarRose are seeking to apply these techniques to social network sites, and companies such as Google have developed face recognition technologies to label photographs and videos [48] of celebrities on the web. As recognition technology improves and the quantity of labeled data increases, it seems that it is only a matter of time before all photos of you on the Internet can be tagged as such and searchable.

Google Goggles which allows visual search using images captured with a smart phone has the potential to carry out face recognition, but privacy concerns have prevented it from being made available [39], according to a spokesman. “We do have the relevant facial recognition technology at our disposal.... But we haven’t implemented this on Google Goggles because we want to consider the privacy implications and how this feature might be added responsibly.” [46].

[3] More than 3 billion photos a day are uploaded to Facebook [20].

27.3.4 Street View

Online services such as Google's Street View, Bing Streetside, Mapjack, and Everyscape present systematically captured street-level imagery on an unprecedented scale, allowing users to explore distant places through intuitive user interfaces in their computer browser. The extent of their coverage, the high image quality and the easy access have aroused concern over the effect of the imagery on privacy. Individuals are concerned about the possibility of their presence in a particular location being publicly visible and about their property being easily examinable without their knowledge and scouted by burglars. In Japan, privacy concerns led to Street View imagery being recaptured with the car-mounted cameras lowered by 40 cm so that the service would not present imagery taken over people's garden walls, and there has been considerable opposition to the service on privacy grounds in Switzerland and Germany [2]. Mechanisms are provided for individuals to request that particular images are not made public, but the ubiquity of faces and license plates in the imagery, and the general unease that these elicit, required an automated solution to attempt to automatically obscure all the faces and license plates. The automatic system that Google deployed to blur faces and license plates is described in Sect. 27.5.1. Flores and Belongie [21] have shown preliminary work using multiple views and inpainting to remove isolated pedestrians images from Street View images.

27.3.5 Institutional Databases

Increasingly in recent years, governments and corporations have sought to harness Information Technology to improve efficiency in their provision of services, to prevent fraud and to ensure the security of citizens. Such developments have involved collecting more information and making that information more readily available to searching and through links between databases. Silos of information, collected for an authorized process are readily accepted for the benefits they bring, but the public becomes more uneasy as such databases succumb to "function creep", being used for purposes not originally intended, especially when several such databases are linked together to enable searches across multiple domains. Plans for Australian identity cards were rejected because of just such fears [17] and there was a significant backlash when retired Admiral John Poindexter conceived the "Total Information Awareness" (TIA) project [43] which aimed to gather and mine large quantities of data, of all kinds, and use these to detect and track criminals and terrorists. The Orwellian potential for such a project raised an outcry that resulted in the project being renamed the *Terrorist Information Awareness* project, an epithet calculated to stifle objection in post-September 11th America.

Naturally, faces are an important part of such electronic databases allowing the verification of identity for such purposes as border control and driver licensing, but registered faces provide a link between definitive, exploitable identification information such as name, address, social security number, bank accounts, immigration

status, criminal record and medical history and the mass of images of individuals that is building up from other channels like surveillance and photo-sharing.[4] Many authors, from Bentham [5] to the present have expressed concern about the potential for state oppression by the exercise of extensive monitoring and the projection that such monitoring is pervasive if unknowable.

The widespread use of electronic records and their portability has led to numerous cases of records being leaked or lost, and their potential value for identity theft has made them a target for theft and hacking, from within as well as outside the controlling institution. This inadvertent exposure is a major reason for strong automatic privacy protection controls such as encryption, tight access control and image redaction even in databases where normal use would not lead to privacy intrusion.

27.4 Technology for Enabling Privacy

In recent years, a number of technological solutions have been proposed for the general problem of privacy protection in images and video, and for face privacy protection in particular. In this section, we review the principal methods being developed: intervention, redaction, and provably secret processing, together with a discussion of privacy policies and tokens for claiming or relinquishing privacy protection.

27.4.1 Intervention

Patel et al. [41] have proposed a system that prevents unauthorized photography by detecting cameras using their retro-reflective properties. In their detection system, a bright infra-red source is located near a camera. If the lens of another camera is pointed toward the detector, a strong retro-reflection is seen in the image, which can easily be detected automatically. When a camera is detected, a light is flashed towards it using a digital projector, spoiling any images that it may record. This unusual approach, dubbed an "anti-paparazzi" device, exploits computer vision to create a privacy-protection solution where no control can be exerted over the use of the images once recorded. As well as privacy protection, the system is envisaged for copyright protection, for instance to prevent recording of new release films in cinemas.

27.4.2 Visual Privacy by Redaction

Most recent work on visual privacy protection has focused on systems that modify, or redact, visual images to mask out privacy sensitive information. Such systems

[4] Consider the case of the British fraudster John Darwin who faked his own death but was identified in a photograph on a real estate web site after subsequently buying property [61].

typically use computer vision technology to determine privacy sensitive regions of the image, for instance tracking moving people in a video [51], or detecting faces [22] in still or moving images. Such regions of interest are then changed in some way to prevent subsequent viewers or algorithms from extracting the privacy sensitive information. Obscuration methods that are commonly used include blurring, masking, pixellating [27], scrambling [18], or permuting pixels [12]. Recent work has investigated the limitations of some of these, for instance Gross et al. [25] show that simple pixellation and blurring may not be strong enough to defeat face recognition systems. They train a *parrot* [37] recognizer on gallery images with the same distortion as the probe and obtain markedly higher recognition rates than using a system trained on clean images. Neustaedter et al. [35] have also found global blurring and other obscuration techniques to be unable to supply simultaneously both sufficient privacy and adequate information for always-on home video conferencing. Koshimizu et al. [31] have explored the acceptability of different obscuration and rerendering techniques for video surveillance.

Stronger masking with greater changes to the image may have the limitation of reducing the usability of the video for its intended purpose, but rerendering [51] may alleviate this by showing computer generated images to convey important information hidden by the redaction process. One example of this would be to obscure a person's face in an image with a computer generated face—hiding the identity yet preserving the gaze direction and expression. Two extensions of this using face modeling are described in Sect. 27.5.2.

One important aspect of redaction systems is reversibility. It may be desirable for some purposes to permanently destroy the privacy-intrusive information, but for others it may desirable or necessary, perhaps for evidential reasons, to be able to reconstruct the original video.

When redacted information is to be presented to the user, one fundamental question is what redaction is necessary and when the redaction is to occur. In some scenarios, the system may need to guarantee that redaction happens at the earliest stage, and that unredacted data is never accessible. For such a scenario, we proposed the PrivacyCam [52], a camera with on-board redaction that behaves as a normal video camera but outputs video with privacy-sensitive areas redacted. Only one level of redaction at a time is possible when such a system is a drop-in replacement for an analogue camera. However for a general system, it may be necessary to present both redacted and unredacted data to end users according to the task and their rights, and to allow different types and extents of redaction according to the circumstances.

In a distributed surveillance system, there are three principal locations through which the data must pass: the video processor, database and browser (or end-user application), at each of which the redaction could take place:

Browser: Here the unredacted data is delivered to the client and client software carries out the redaction and presents the redacted information to the user. This scenario means that redacted data does not need to be stored and transmitted but metadata for redaction does need to be transferred with the raw data. Since the browser is the part of a system most exposed to attack, transmitting the unredacted data there is not secure.

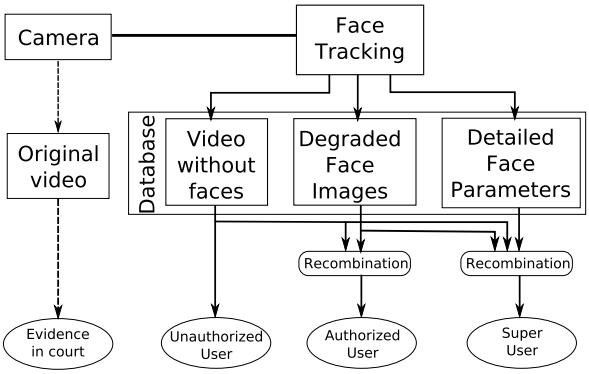

Fig. 27.1 Double redaction: Video is decomposed into information streams which protect privacy and are recombined when needed by a sufficiently authorized user. Information is not duplicated and sensitive information is only transmitted when authorized. Optionally an unmodified (but encrypted) copy of the video may need to be securely stored to meet requirements for evidence

Content management: The content management system can redact the information when requested for viewing, which will minimize storage requirements and allow complete flexibility, but involve additional processing (with the same keyframe perhaps being redacted multiple times), latency and imposes image modification requirements on the database system. If the unredacted video is stored, unauthorized access can reveal the privacy-intrusive data.

Video analytics: The video analytics system has access to the richest information about the video activity and content, and thus can have the finest control over the redaction, but committing at encoding time allows for no post-hoc flexibility. In the other two scenarios, for instance, a set of people and objects could be chosen and obscured on-the-fly. Sending redacted and raw frames to the database imposes bandwidth and storage requirements.

Double redaction: Perhaps the most flexible and secure method is *double redaction* [50], in which privacy protection is applied at the earliest possible stage (ideally at the camera), and privacy-protected data flows through the system by default. Separate encrypted streams containing the private data can be transmitted in parallel to the content management system and to authorized end users, allowing the inversion of the privacy protection in controlled circumstances. The operating point of the detection system can even be changed continuously at display time according to a user's rights, to obscure all possible detections, or only those above a certain confidence. Figure 27.1 shows an example of such a double redaction scheme, with two levels of redaction.

Several authors [28, 40] have adopted such a double redaction system and have explored options for embedding or hiding additional data streams in the redacted video data, for instance Zhang et al. [64] store the information as a watermark, Carillo et al. [12] use a reversible cryptographic pixel permutation process to obscure information in a manner that can be reversed, given the right key, and that is robust to compression and transcoding of video. Li et al. transform sensitive data using the

Discrete Wavelet Transform, preserving only low frequency information and hide the encrypted high-frequency information in the JPEG image.

27.4.3 Cryptographically Secure Processing

A recent development in visual privacy protection is in the development of cryptographically secure methods for processing images. These methods establish protocols by which two parties can collaborate to process images without risk of privacy intrusion. In particular, if one party owns images and another party has an image processing algorithm, algorithms such as "Blind Vision" [3] allow certain algorithmic operations to be carried out by the second party on the first party's data without the data itself or the algorithm being made available to the other party. Such systems have been applied to the problems of face detection and recognition, as will be discussed in Sect. 27.5.3.

27.4.4 Privacy Policies and Tokens

An important aspect of privacy protection systems is the policies for determining what data needs to be obscured for which users. As we have seen in Sect. 27.2, privacy systems may need to operate in different modes according to different factors including the roles, authorization and relationship of the observer and the observed. Determining privacy policies is a complex area, made more so when detection of the privacy sensitive information is not reliable.

Brassil [9] and Wickramasuriya et al. [62] explore the use of devices (detected through separate sensors) that can be used to claim privacy protection in public cameras, and Schiff et al. [49] use visual cues (hats or jackets) to designate individuals whose faces are to be redacted, or preserved from redaction.

27.5 Systems for Face Privacy Protection

In this section, we describe three approaches that have been specifically designed for privacy protection of face images.

27.5.1 Google Street View

As mentioned in Sect. 27.3.4, Google's Street View and similar sites present particular privacy challenges with their vast coverage of street-level public imagery. Frome et al. [22] describe the system that they developed to address this problem.

They highlight both the scale of the problem and the challenging nature of their "in the wild" data. They use a standard sliding-window face detector that classifies each square region of each image as face or non-face. The detector is applied with two operating points, and the results are combined with a number of other features (including face color and a 3D position estimate) using a neural network to determine a final classification of the region as face or non-face. All face detections are blurred in the final imagery that is served in Google Maps. A similar system is used to detect and blur license plates.

They describe the criteria used for choosing the redaction method, that it should be: (1) irreversible; (2) visually acceptable for both true faces and false positives; (3) makes it clear to the public that redaction has taken place, a requirement that precludes the use of rerendering techniques from the next section. To meet these requirements, the authors choose to redact the faces with Gaussian blurring and the addition of noise.

27.5.2 De-identifying Face Images

Coutaz et al. [15], described a system for preserving privacy in the CoMedi media space which gives remote participants a sense of co-presence. The system offered shadowing and resolution lowering redaction methods [27] for privacy protection but also used eigenspace filtering for face redaction. In this technique an eigenface [59] representation is constructed using a training set of face images, and faces detected in the mediaspace video are projected into that eigenspace before rerendering. This effectively constrains the rendered face to conform to variations seen in the training set, and obscures other kinds of appearance differences. While this can protect privacy in some ways, such as hiding socially incorrect gestures or expressions, it is also shown to have limitations. The choice of the correct model and the corresponding training set is crucial. Using a mismatched model may unintentionally change the identity, pose, or expression of the face.

In several papers, Sweeney and collaborators [23–25, 36] have described a series of algorithms for de-identifying faces that extend this eigenface approach, tackling the problem of identity hiding. They use the Active Appearance Model [14], face representation which normalizes for pose and facial expression. Their algorithms create a weighted average of faces from different people, such that the resulting face is not identifiable. This deidentification is termed k-same in that it results in a face whose chance of being correctly identified is no more than $\frac{1}{k}$. In their more recent work [25], they use a multifactor decomposition in their face representations that reduces blending artifacts and allows the facial expression to be preserved while hiding the face identity. They also consider the application of this in a medical video database, showing patients' responses to pain, in which facial expression, not identity, is important.

27.5.3 Blind Face Recognition

As described in Sect. 27.4.3, a new field of research is cryptographically provable privacy preserving signal processing, or "Blind vision". Recent work has applied this to face detection [4] and recognition algorithms. Erkin et al. [19] describe a secure implementation of an eigenface face recognition algorithm [59]. Their system performs the operations of projecting a face image onto the eigenvectors of "face subspace" and calculating the distances to each of the enrolled faces, without the querying party, Alice, having to reveal the query image, nor the owner of the face recognizer, Bob, having to reveal the enrolled faces. Such a secure multiparty computation can be very laborious and time consuming, with a single recognition taking 10–20 s, though speed-ups have been proposed [47].

27.6 Delivering Visual Privacy

The technological tools of the previous section can help to prevent privacy intrusion from image-based applications, but as we have seen they form only part of a privacy solution, along with information security and privacy policies. To ensure that the privacy benefits are delivered effectively, two further factors must be considered—ensuring that systems are used where appropriate and operate effectively when installed.

27.6.1 Operating Point

Video information processing systems are error prone. Perfect performance can not be guaranteed, even under fairly benign operating conditions, and systems make two types of errors when determining image regions for redaction: missed detection (of an event or object) and false alarm (triggering when the event or object is not present). We can trade these errors off against one another, choosing an *operating point* with high sensitivity that has few missed detections, but many false alarms, or one with low sensitivity that has few false alarms, but more often fails to detect real events when they occur.

The problems of imperfect image processing can be minimized by selecting the appropriate system operating point. The *costs* of missed detection and false alarm can be quite different, as seen in Sect. 27.5.1, where not blurring a face reveals private information and blurring a non-face degrades the quality of the information provided. In a surveillance system, the operating point for privacy protection may be chosen differently than for general object detection for indexing. Given the sensitive nature of the information, it is likely that a single missed detection may reveal personal information over extended periods of time. For example, failing to detect, and thus obscure, a face in a single frame of video could allow identity information

to be displayed and thus compromise the anonymity of days of aggregated track information associated with the supposedly anonymous individual. On the other hand, an occasional false alarm and unnecessary redaction may have a limited impact on the effectiveness of the installation. The operating point can be part of the access-control structure—greater authorization allows the reduction of the false alarm rate at a higher risk of compromising privacy. Additional measures such as limiting access to freeze-frame or data export functions can also reduce the risks associated with occasional failures in the system. For some applications it will be some time before algorithms are accurate enough to deliver an operating point that gives useful privacy benefits without degrading the usefulness of the data provided.

Even with perfect detection, anonymity cannot be guaranteed. While face recognition is the most salient identifier in video, a number of other biometrics such as face, gait or ear shape; and weak identifiers (height, pace length, skin color, clothing color) can still be preserved after face redaction. Contextual information alone may be enough to uniquely identify a person even when all identifying characteristics are obscured in the video. Obscuring biometrics and weak identifiers will nevertheless reduce the potential for privacy intrusion. These privacy-protection algorithms, even when operating imperfectly, will serve the purpose of making it harder, if not impossible, to run automatic algorithms to extract privacy-intrusive information, and making abuses by human operators more difficult or costly.

27.6.2 Will Privacy Technology Be Used?

The techniques described in this chapter could be considered as optional additions to systems that display images—that will cost more and risk impinging on the usefulness of the systems, while the privacy protection benefits may accrue to stakeholders other than the service provider or the primary users. We must then ask why providers of image-based services will choose to bear the extra burden of implementing privacy protection technologies, even when the technologies are fast and accurate enough to be practically deployed. Clearly in many cases companies will choose to implement them as being the "right thing" to do, out of concern for protecting privacy, and for guarding their good name. Others may be pressured by the public, shareholders or customers to apply such technologies, or be asked to do so by privacy ombudsmen. Finally explicit legislation may be implemented to require such technologies, though creating manageable legislation for the nebulous area of privacy is extremely difficult. Existing legislation in some jurisdictions may already require the deployment of these techniques in domains such as surveillance as soon as they become feasible and commercially available.

Even when privacy protection methods are mandated, compliance and enforcement are still open to question, particularly in private systems such as medical images and surveillance. McCahill and Norris [33] estimated that nearly 80% of CCTV systems in London's business space did not comply with current data protection legislation, which specifies privacy protection controls such as preventing unauthorized

people from viewing CCTV monitors. Legislating public access to surveillance systems as proposed by Brin [10] is one solution, but that still begs the question—are there are additional video feeds that are not available for public scrutiny? A potential solution that we have proposed [52] is certification and registration of systems, along the lines of the TRUSTe system that evolved for Internet privacy. Vendors of video systems might invite certification of their privacy-protection system by some independent body. (In the US, the Federal Trade Commission Act[5] has the power to enforce companies' privacy policies.) For purpose-built devices with a dedicated camera sensor (like PrivacyCam, Sect. 27.4.2), this would suffice. Individual surveillance installations could also be certified for compliance with installation and operating procedures, with a certification of the privacy protection offered by the surveillance site prominently displayed on the equipment and CCTV advisory notices. Such notices might include a site (or even camera) identification number and the URL or SMS number of the surveillance privacy registrar where the site can be looked up to confirm the certification of the surveillance system. Consumer complaints would invoke investigations by the registrar, and conscientious companies could invite voluntary inspections.

27.7 Conclusions

As cameras and networking have become cheaper and ubiquitous, there has been an explosion in the dissemination of images, for new and traditional applications from photo-sharing to surveillance and medical imaging. With this explosion, there has been a corresponding increase in the potential for privacy-intrusive uses of those images. Thus far, controls on such privacy intrusions have been very limited. We have examined how images in different domains can contain sensitive information, particularly images of faces that allow individuals to be identified. We have described ways in which that information can be obscured by redaction, based on computer vision techniques to identify regions of interest, and image processing techniques to carry out the redaction in a secure, possibly invertible, manner. Finally, we have described three particular systems that have been used to apply privacy preserving techniques to face images and explored ways in which such privacy protection techniques can be deployed and might become more widespread.

References

1. Acquisti, A.: Privacy and security of personal information: Economic incentives and technological solutions. In: Camp, J., Lewis, R. (eds.) The Economics of Information Security. Kluwer, Dordrecht (2004)
2. Associated Press. Swiss official demands shutdown of Google Street View. New York Times (2009)

[5] http://www.ftc.gov/privacy.

3. Avidan, S., Butman, M.: Blind vision. In: European Conference on Computer Vision, vol. 3953, pp. 1–13 (2006)
4. Avidan, S., Butman, M.: Efficient methods for privacy preserving face detection. In: NIPS, pp. 57–64 (2006)
5. Bentham, J.: Panopticon Letters. London (1787). http://cartome.org/panopticon2.htm
6. Black, J., Ellis, T.: Multi camera image tracking. In: International Workshop on Performance Evaluation of Tracking and Surveillance (2001)
7. Bolle, R.M., Connell, J.H., Pankanti, S., Ratha, N.K., Senior, A.W.: Guide to Biometrics: Selection and Use. Springer, New York (2003)
8. Boult, T., Micheals, R.J., Gao, X., Eckmann, M.: Into the woods: Visual surveillance of non-cooperative and camouflaged targets in complex outdoor settings. Proc. IEEE **89**(10), 1382–1402 (2001)
9. Brassil, J.: Using mobile communications to assert privacy from video surveillance. In: 19th IEEE International Parallel and Distributed Processing Symposium, p. 290a (2005)
10. Brin, D.: The Transparent Society: Will Technology Force Us to Choose Between Privacy and Freedom. Perseus, Cambridge (1999)
11. Caloyannides, M.: Society cannot function without privacy. IEEE Secur. Priv. Mag., May/June 2003
12. Carrillo, P., Kalva, H., Magliveras, S.: Compression independent reversible encryption for privacy in video surveillance. EURASIP J. Inf. Secur.
13. Congestion charging: Enforcement technology. BBC LDN (2003). http://www.bbc.co.uk/london/congestion/technology.shtml
14. Cootes, T., Edwards, G., Taylor, C.: Active appearance models. IEEE Trans. Pattern Anal. Mach. Intell. **23**(6), 492–7 (2001)
15. Coutaz, J., Bérard, F., Carraux, E., Astier, W., Crowley, J.L.: CoMedi: Using computer vision to support awareness and privacy in mediaspaces. In: CHI, pp. 13–14. ACM Press, New York (1999)
16. Danielson, P.: Video surveillance for the rest of us: Proliferation, privacy, and ethics education. In: International Symposium on Technology and Society, 6–8 June 2002, pp. 162–167 (2002)
17. Davies, S.: The loose cannon: An overview of campaigns of opposition to national identity card proposals. In: e-ID: Securing the Mobility of Citizens and Commerce in a Greater Europe. Unisys, February 2004
18. Dufaux, F., Ebrahimi, T.: Scrambling for video surveillance with privacy. In: Proceedings of Computer Vision and Pattern Recognition, June 2006, p. 160 (2006)
19. Erkin, Z., Franz, M., Guajardo, J., Katzenbeisser, S., Lagendijk, I., Toft, T.: Privacy-preserving face recognition. In: Privacy Enhancing Technologies Symposium (2009)
20. Facebook press room, 21 April 2010
21. Flores, A., Belongie, S.: Removing pedestrians from Google Street View images. In: International Workshop on Mobile Vision, June 2010
22. Frome, A., Cheung, G., Abdulkader, A., Zennaro, M., Wu, B., Bissacco, A., Adam, H., Neven, H., Vincent, L.: Large-scale privacy protection in Google Street View. In: Proceedings of Computer Vision and Pattern Recognition (2009)
23. Gross, R., Airoldi, E., Malin, B., Sweeney, L.: Integrating utility into face de-identification. In: Workshop on Privacy Enhancing Technologies. CMU (2005)
24. Gross, R., Sweeney, L., de la Torre, F., Baker, S.: Model-based face de-identification. In: Workshop on Privacy Research in Vision. IEEE, New York (2006)
25. Gross, R., Sweeney, L., Cohn, J., de la Torre, F., Baker, S.: Face de-identification. In: Senior, A.W. (ed.) Protecting Privacy in Video Surveillance. Springer, Berlin (2009)
26. Hampapur, A., Brown, L., Connell, J., Ekin, A., Lu, M., Merkl, H., Pankanti, S., Senior, A., Tian, Y.L.: Multi-scale tracking for smart video surveillance. IEEE Trans. Signal Process. (2005)
27. Hudson, S., Smith, I.: Techniques for addressing fundamental privacy and distribution trade-offs in awareness support systems. In: CSCW, pp. 248–257 (1996)
28. Ito, I., Kiya, H.: One-time key based phase scrambling for phase-only correlation between visually protected images. EURASIP J. Inf. Secur.

29. Khan, S., Shah, M.: Tracking people in presence of occlusion. In: Asian Conference on Computer Vision (2000)
30. Koshimizu, T., Toriyama, T., Babaguchi, N.: Factors on the sense of privacy in video surveillance. In: Proceedings of the 3rd ACM Workshop on Continuous Archival and Retrieval of Personal Experiences, pp. 35–44. ACM, New York (2006)
31. Koshimizu, T., Umata, I., Toriyama, T., Babaguchi, N.: Psychological study for designing privacy protected video surveillance system: PriSurv. In: Senior, A.W. (ed.) Protecting Privacy in Video Surveillance. Springer, Berlin (2009)
32. Lipton, A.J., Clark, J.I.W., Thompson, B., Myers, G., Zhang, Z., Titus, S., Venetianer, P.: The intelligent vision sensor: Turning video into information. In: Advanced Video and Signal-based Surveillance. IEEE, New York (2007)
33. McCahill, M., Norris, C.: CCTV. Perpetuity Press, Leicester (2003)
34. McKenna, S., Jabri, J.S., Duran, Z., Wechsler, H.: Tracking interacting people. In: International Conference on Face and Gesture Recognition, March 2000, pp. 348–53 (2000)
35. Neustaedter, C., Greenberg, S., Boyle, M.: Blur filtration fails to preserve privacy for home-based video conferencing. ACM Trans. Comput. Hum. Interact. (2006)
36. Newton, E., Sweeney, L., Malin, B.: Preserving privacy by de-identifying facial images. Technical Report CMU-CS-03-119, Carnegie Mellon University, School of Computer Science, Pittsburgh (2003)
37. Newton, E., Sweeney, L., Malin, B.: Preserving privacy by de-identifying facial images. IEEE Trans. Knowl. Data Eng. **2**(17), 232–243 (2005)
38. New York city police department releases draft of public security privacy guidelines for public comment. NYPD Press Release, February 2009
39. Palmer, M.: Google debates face recognition technology. Financial Times, 19 May 2010
40. Paruchuri, J.K., Cheung, S.S., Hail, M.W.: Video data hiding for managing privacy information in surveillance systems. EURASIP J. Inf. Secur.
41. Patel, S.N., Summet, J.W., Truong, K.N.: Blindspot: Creating capture-resistant spaces. In: Senior, A.W. (ed.) Protecting Privacy in Video Surveillance, pp. 185–201. Springer, Berlin (2009)
42. Phillips, P.J., Scruggs, W.T., O'Toole, A.J., Flynn, P.J., Bowyer, K.W., Schott, C.L., Sharpe, M.: FRVT 2006 and ICE 2006 large-scale results. Technical Report NISTIR 7408, NIST, Gaithersburg, MD 20899, March 2006
43. Poindexter, J.: Overview of the information awareness office, August 2002. http://www.fas.org/irp/agency/dod/poindexter.html
44. Poseideon. http://www.poseidon-tech.com/
45. Potamianos, G., Neti, C., Gravier, G., Garg, A., Senior, A.W.: Recent advances in the automatic recognition of audiovisual speech. Proc. IEEE (2003)
46. Privacy fears force search giant to block facial recognition application on Google goggles. The Daily Mail Online, December 2009
47. Sadeghi, A.R., Schneider, T., Wehrenberg, I.: Efficient privacy-preserving face recognition. In: 12th International Conference on Information Security and Cryptology (2010)
48. Sargin, M.E., Aradhye, H., Moreno, P., Zhao, M.: Audiovisual celebrity recognition in unconstrained web videos. In: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, April 2009
49. Schiff, J., Meingast, M., Mulligan, D.K., Sastry, S., Goldberg, K.: Respectful cameras: Detecting visual markers in real-time to address privacy concerns. In: Senior, A.W. (ed.) Protecting Privacy in Video Surveillance. Springer, Berlin (2009)
50. Senior, A.W.: Privacy protection in a video surveillance system. In: Senior, A.W. (ed.) Protecting Privacy in Video Surveillance. Springer, Berlin (2009)
51. Senior, A.W., Pankanti, S., Hampapur, A., Brown, L., Tian, Y.-L., Ekin, A.: Blinkering surveillance: Enabling video privacy through computer vision. Technical Report RC22886, IBM T.J.Watson Research Center, NY 10598, August 2003
52. Senior, A.W., Pankanti, S., Hampapur, A., Brown, L., Tian, Y.-L., Ekin, A.: Enabling video privacy through computer vision. IEEE Secur. Priv. **3**(5), 50–57 (2004)

53. Spiekermann, S., Grossklags, J., Berendt, B.: E-Privacy in 2nd Generation E-Commerce: Privacy Preferences Versus Actual Behavior, pp. 38–47. ACM Press, New York (2001)
54. Stana, R.: Video surveillance. Technical Report GAO–03–748, United States General Accounting Office, June 2003
55. Stauffer, C., Grimson, W.E.L.: Learning patterns of activity using real-time tracking. IEEE Trans. Pattern Anal. Mach. Intell. **22**(8), 747–757 (2000)
56. Sweeney, L., Gross, R.: Mining images in publicly-available cameras for homeland security. In: AAAI Spring Symposium on AI technologies for Homeland Security (2005)
57. Swiss Federal Data Protection Commissioner. Leaflet on video surveillance by private individuals. 3003 Bern, January 2003
58. The health insurance portability and accountability act (HIPAA) privacy and security rules (1996)
59. Turk, M., Pentland, A.: Eigenfaces for recognition. J. Cogn. Neurosci. **3**(1), 71–86 (1991)
60. Venetianer, P., Zhang, Z., Scanlon, A., Hu, Y., Lipton, A.: Video verification of point of sale transactions. In: AVSS (2007)
61. Weaver, M.: Canoe mystery man arrested for fraud. The Guardian, December 2007
62. Wickramasuriya, J., Alhazzazi, M., Datt, M., Mehrotra, S., Venkatasubramanian, N.: Privacy-protecting video surveillance. In: SPIE International Symposium on Electronic Imaging (2005)
63. Yu, T., Lim, S.-N., Patwardhan, K., Krahnstoever, N.: Monitoring, recognizing and discovering social networks. In: Proceedings of Computer Vision and Pattern Recognition (2009)
64. Zhang, W., Cheung, S.-C., Chen, M.: Hiding privacy information in video surveillance system. In: Proc. International Conference on Image Processing (2005)