Noname manuscript No.
(will be inserted by the editor)

Anonymous Subject Identification and Privacy Information Management in Video Surveillance

Ying Luo · Sen-ching S. Cheung · Riccardo Lazzeretti · Tommaso Pignata · Mauro Barni

Received: date / Accepted: date

Abstract The widespread deployment of surveillance cameras has raised serious privacy concerns and many privacy-enhancing schemes have been recently proposed to automatically redact images of selected individuals in the surveillance video for protection. Of equal importance are the privacy and efficiency of techniques to first, identify those individuals for privacy protection and second, provide access to original surveillance video contents for security analysis. In this paper, we propose an anonymous subject identification and privacy data management system to be used in privacy-aware video surveillance. The anonymous subject identification system uses iris patterns to identify individuals for privacy protection. Anonymity of the iris-matching process is

The first two authors acknowledge the support by the National Science Foundation under Grant No. 1018241. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation

Ying Luo
Department of Computer Information Technology and Graphics, Purdue University Northwest, Hammond, IN, USA 46323
E-mail: ying.luo@pnw.edu

Sen-ching S. Cheung
Department of Electrical and Computer Engineering, University of Kentucky, Lexington, KY, USA 40506
E-mail: sccheung@ieee.org

Riccardo Lazzeretti
Mathematics Department, University of Padova, Padova, Italy
E-mail: riccardo.lazzeretti@math.unipd.com

Tommaso Pignata
Information Engineering Department, University of Siena, Siena, Italy
E-mail: pignata.tommaso@gmail.com

Mauro Barni
Information Engineering Department, University of Siena, Siena, Italy
E-mail: barni@dii.unisi.it

guaranteed through the use of a garbled-circuit (GC) based iris matching protocol. A novel GC complexity reduction scheme is proposed by simplifying the iris masking process in the protocol. A user-centric privacy information management system is also proposed that allows subjects to anonymously access their privacy information via their iris patterns. The system is composed of two encrypted-domain protocols: the privacy information encryption protocol encrypts the original video records using the iris pattern acquired during the subject identification phase; the privacy information retrieval protocol allows the video records to be anonymously retrieved through a GC-based iris pattern matching process. Experimental results on a public iris biometric database demonstrate the validity of our framework.

Keywords Anonymous Subject Identification · Privacy Information Management · Privacy Protection · Video Surveillance · Garbled circuit

1 Introduction

In recent years, surveillance cameras have been widely used for preventing theft, collecting population data, and combating terrorism. Advances in pattern recognition algorithms such as searchable surveillance and automatic event/human recognition have turned the once labor-intensive processes into powerful automated systems that can quickly and accurately identify visual objects and events. Thus, it is unsurprising that the general public is increasingly wary about the possibility of privacy invasion with video surveillance systems. To mitigate these concerns and to facilitate continued development of surveillance technologies, it is imperative to make privacy protection a priority in current and future video surveillance systems.

Systematic study of privacy protection in video surveillance can be traced back to the PeopleVision system developed at IBM [59]. In the past decade, many algorithms and prototypes have been developed, and some of them will be reviewed in Section 2. Existing work in privacy protection mostly target towards applications in public places like airports or city streets which do not differentiate different individuals. Only mild obfuscation techniques such as blurring or pixelation are applied to every individual in the scene so that the resulting video is still useful for security purposes. Although such weak forms of protection cannot withstand privacy attacks [47,56], it is still considered beneficial as the expectation of privacy in public places is low.

However, as the applications of surveillance system broaden, the demands for privacy become more sophisticated. There are many situations in which different individuals in the environment may have different privacy requirements. Semi-public places like hospitals or schools may have legal and/or contractual responsibility to protect privacy of selected individuals. For example in U.S., privacy of students are protected under FERPA [69] and patients are protected under HIPAA [27]. Even for privately-owned facilities, tenants living in an apartment building may not want their every departure and return to be logged and observed [34], while employees may feel apprehensive if their

activities are constantly being monitored by their supervisors [61]. In order to balance the need of privacy and security, only a specific group of individuals in the environment, such as trusted employees of a corporation or patients in a hospital, are provided with privacy protection while transient visitors must be monitored at all time.

A key challenge of such a form of selective protection is a reliable and secure mechanism to identify whether a particular individual belongs to the privacy group. To reliably identify any individual, the best approach would be to rely on biometric signals like iris patterns which are convenient and highly discriminative. The use of biometric signals, however, provides *a direct link between the imagery to the true identity of an individual*. Such a linkage contradicts one of the fundamental tenets of privacy design discussed in the Common Criteria specification [48], namely the unlinkability property that ensures different usages by the same user cannot be linked together. If the security of the system is compromised, this extra information may pose a greater privacy risk that it purports to protect. It is thus important to sever this link between the biometric signal and the imagery to provide an anonymous identification mechanism for subject identification.

This anonymous subject identification service, however, should not interfere with the basic non-repudiation property of a surveillance system. As surveillance video is a security record, privacy protection system must provide secure mechanisms to preserve the original footage to show the presence/absence of any individual in the environment [15, 43, 50, 53, 54]. To retrieve the original videos, existing approaches often assume the most simplistic access control model in which a single user, usually the owner of the surveillance, has complete control of all contents. Such a centralized model is certainly not suitable for all situations, especially when there is a privacy conflict between the owner and the subject. A better approach is to *treat the privacy visual information of an individual in the same manner as any other privacy information such as personal financial or medical information – each access of the information must require a full consent from the corresponding user*. This is consistent with the basic premise in both the Privacy Act of U.S. and the Data Protection Directives of E.U. to obtain users' consent before accessing their sensitive information [71, Chapter 19]. On the other hand, implementing such a fine-grained access control posts a technical challenge because the anonymous surveillance system cannot associate the imagery with the unknown identity of the individual.

In this paper, we describe a complete design of a selective privacy protection surveillance system to simultaneously address the problems of anonymous identification and access control of privacy videos. Specifically, our proposed system has two main novel constructions. First, we propose an Anonymous Subject Identification (ASI) procedure that allows the surveillance system to use iris patterns in determining the privacy protection status of an individual without knowing his/her true identity. Our anonymous iris matching is based on garbled circuits [72] and we propose a novel masking technique that significantly reduces the complexity of the circuit. The new masking technique

stems from the statistical patterns of iris patterns and is independent from the specific GC implementation used. Second, we propose a novel Privacy Information Management (PIM) system that uses iris patterns in encrypting the privacy video. Two protocols, one for encryption and one for retrieval, are developed to ensure the privacy of both the plaintext biometric signals and privacy visual information. Preliminary results of this work have appeared in [39,41]. Thoroughly revised protocols and additional experimental results are presented here to demonstrate the effectiveness of our schemes.

The rest of the paper is organized as follows: related work are reviewed in Section 2 and an overview of our proposed system is in Section 3. The image processing component of our system is reviewed in Section 4. The design of the ASI module using GC is presented in Section 5. In Section 6, we describe the two protocols used in our privacy data management system and prove their validity in protecting privacy information. We have tested our system using a large collection of iris patterns and the results are shown in Section 7. We conclude the paper in Section 8.

2 Related Work

Most existing literature on privacy protected video surveillance focus on how to identify and protect sensitive information in the video. It is primarily a computer vision problem, aimed to identify, segment, and obfuscate sensitive visual information. Many schemes have been proposed ranging from the use of black boxes or large pixels [68], scrambling [15], face replacement [74], body replacement [67], to complete object removal [66] and a hybrid approach [49]. In addition, some work use variable strength filters such as inhomogeneous diffusion [51] to provide different degrees of protection based on the inferred sensitivity level of the pixel. The sensitivity can be determined based on distance [16,11], background details [16], and intention of the video as estimated by the camera motion and the scene structure [45,46]. A recent work has also analysed the trustworthiness of these techniques [54]. These works do not explicitly identify each individual to provide different levels of privacy protection. While they can provide the essential visual obfuscation function, they are not adequate in providing anonymity to selective privacy protected surveillance systems.

There are a number of privacy protection systems that address the issue of anonymous subject identification. In [58], individuals wore yellow hard hats to indicate that their identities need to be protected. The prominent yellow color on top of a person's head makes the hard hat an easy target for automatic identification. In [75], a color tag was used instead and triangulation through multiple cameras was employed to minimize the impact of self occlusion. Other sensors such as RFID have also been explored [70,53]. These methods provide anonymous identification as there is no link between the true identity and the physical marker used. However, the reliance on a physical marker makes the system vulnerable to physical attacks. The subjects will lose privacy protection

from the system if the marker is accidentally dropped. More importantly, if the marker is maliciously embezzled by unauthorized individuals, the system will protect the potential intruders and the security of the environment will be compromised.

A better approach for identification is based on biometric signals. There are a variety of biometric signals used for identification and the most common types include fingerprints, palmprints, voice, signature, face, iris and retinal patterns [71, Ch. 3]. In general, biometric signals excel in authenticating the subject's identity since they are based on who the subjects are. Face is perhaps the most commonly used one in surveillance as it can be recognized at long distance [23]. Iris recognition, on the other hand, is a far more accurate biometric system compared with face recognition [24]. While traditional iris readers require subjects to be in close range, emerging technologies can read iris in video surveillance environments [33,13]. As such, we have chosen iris recognition for our system based on the popular iriscode algorithm from [12]. Unlike those works that aim to extract iris patterns in the environment, our focus would be on developing protocols to provide anonymity to iris matching and to use iris for encrypting privacy information.

As described in Section 1, the use of biometric signals in selective privacy protection systems poses a security threat as it provides a direct link between the privacy information to the true identity of an individual. Anonymous biometric matching can be achieved using Secure Multi-Party Computation (SMPC) protocols [26], which can simultaneously guarantee the anonymity of the subject and the authentication of the biometric query [5]. The blind-vision face detector proposed in [1] is one such example. The authors use Oblivious Transfer (OT) to detect face without direct access to the raw pixel values. The privacy of the detector is guaranteed by having the entire table encrypted by a pre-computed set of public keys and transmitted to the server for detection. However, the large table size, the intensive encryption and decryption operations render the blind-vision detector computationally inefficient. A faster but less general approach is to use partial Homomorphic Encryption (HE) which preserves certain operations in the encrypted domain [19]. Paillier encryption, a type of HE systems, has been used to provide anonymous biometric matching in faces [17], iris patterns [73], and fingerprints [4].

While HE is efficient for arithmetic operations, iris matching consists of mostly binary operations that are more suitable to be implemented by using Garbled Circuits (GC). GC provides a generic implementation of any binary function by having one party prepared an encrypted boolean circuit, and another party obliviously evaluated the circuit without accessing to intermediate values [72]. Blanton et al. propose a hybrid approach of GC and HE for efficient iris matching [8]. Our implementation is a pure GC solution which is more appealing for two reasons. First, GC relies on faster symmetric encryption while HE is based on asymmetric encryption. Recent research efforts have exploited such fact to significantly improve GC's efficiency [30,52,38,60]. Second, GC is characterized by shorter security parameters, whose impact on efficiency becomes pronounced in delivering medium and long term security [2]. In this

paper, we further exploit key characteristics of iris patterns to develop a novel computationally efficient GC-based iris matching circuit.

With anonymous subject identification, the surveillance system can anonymously determine privacy status of each individual for privacy protection. To ensure non-repudiation, the system must provide mechanism to access the original video. Existing approaches typically describe the access control via a privacy policy [18,37,11], enforced by a trusted server which is vulnerable to concerted attacks. A more distributed management system where the users and the server agents can anonymously exchange data, credential, and authorization information was proposed in [10]. However, anonymous access control is not implemented in the system and the server could still associate the encrypted videos with the identity of an individual. The privacy information management system proposed in this paper eliminates the trusted server by using SMPC protocols to authenticate and decrypt privacy information.

The key innovation comes from our approach in combining anonymity in biometric access control and privacy data management. Methods that use biometric to protect sensitive data are referred to as biometric cryptosystems [65]. They have been applied in a number of practical biometric systems [25,9,64, 36] in which a random key is protected by a biometric signal to produce a privacy template [25,9] or helper data [64,36]. Such a privacy template or helper data can only be decrypted by another biometric sample from the same individual. The purpose of their proposed protocols is to protect the security of the biometric system against the attack to central server by replacing the raw biometric samples with these templates. For our application, we use biometric cryptosystems to protect the AES keys that encrypt the privacy imagery. In [25], a key-binding iris template scheme is proposed that relies on error correction coding (ECC) to cope with small variations between different iris patterns from the same individual. While ECC-based techniques are efficient, ECC dictates the use of Hamming distance in measuring similarity between biometric signals and limits the threshold tolerance in the matching process. Our proposed approach uses a general GC-based SMPC protocol which is capable of complex similarity function and arbitrary choice of similarity tolerance.

3 System Overview and Security Model

The architecture of our system is shown in Figure 1. It has two networks that are separately administered. The inner network is the Privacy Protected Video Surveillance Network that consists of a set of cameras with the associated video processing units. The outer network is a Server network, housing two databases of encrypted videos and redacted videos, as well as a gallery of iriscodes of individuals whose privacy need to be protected in the surveillance environment.

We assume that there is an offline enrolment process to determine who need to be protected and to populate their iriscodes into the gallery. The two video databases are used to store redacted videos accessible by security personnel

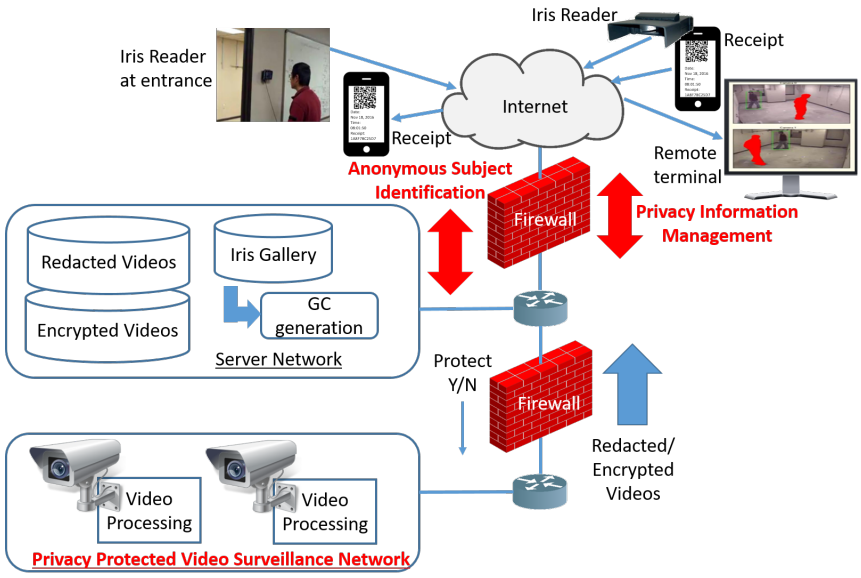


Fig. 1 System Architecture: the Server Network and Privacy Protected Video Surveillance Network interacting with the external iris readers at entrance and remote terminal. The three main groups of protocols are in red fonts.

with privacy information removed, and encrypted videos of the original footage retrievable only by the authorized members in the videos.

The gallery only interacts with external iris readers at various locations through an iris-matching garbled circuit, which is a one-time software that supports encrypted-domain matching with a query probe. Two types of external interactions are supported:

1. At any entrance to the surveillance environment, a user must provide an iris probe for the anonymous matching. A virtual receipt ticket will be provided back to the user as a QR code. The receipt contains time-of-entry and a retrieval code for later retrieval.
2. At a remote terminal, a user can view surveillance video with all privacy information redacted. A user can also provide an iris probe with the appropriate retrieval code to remove redaction of his/her own imagery.

All the functions of our proposed system are grouped into three main sets of protocols: the Anonymous System Identification module, the Privacy-Protected Video Surveillance Network, and the Privacy Information Management module. Their functions are described below, followed by the security model used throughout our system.

3.1 Anonymous Subject Identification

The Anonymous Subject Identification (ASI) module is jointly executed between the iris reader at the external entrance of the surveillance environment and a Iris-matching Garbled Circuit at the server. ASI has the following guarantees:

1. The server sends a decision bit to the surveillance network on whether the iris probe from the reader matches any signal represented in the garbled circuit. If there is a match, visual privacy of the incoming subject will be protected.
2. The server does not know the probe. Even if there is a match, the server will not know which signal in the garbled circuit matches the probe. This ensures the privacy of incoming subject.
3. The reader does not know anything about the signals represented in the garbled circuit. This protects the secrecy of the gallery.

The details of ASI can be found in Section 5.

3.2 Privacy-Protected Video Surveillance Network

The second component is the Privacy-Protected Video Surveillance Network. Due to the sensitively of raw video footage, this surveillance network is protected by two firewalls. The only input to the network is the decision bit from ASI and all output videos are either encrypted or redacted. If ASI identifies a match, the surveillance network will:

1. track the target individual across different camera views and handle possible occlusions with other individuals with different privacy status,
2. prepare a redacted video to be stored in the redacted video database with the target individual obfuscated, and
3. encrypt and store the original pixels in the encrypted video database to enable possible reversal of the obfuscation.

The implementation of the image processing steps in (1) and (2) will be discussed in Section 4. The encryption step (3) is part of the Privacy Information Management module, which will be discussed next.

3.3 Privacy Information Management

Privacy Information Management (PIM) module is jointly executed by all the databases in the server network and a remote user with an iris reader to decrypt his/her protected imagery. PIM has the following guarantees:

1. The user should have access to all of his/her original video segments.
2. The user does not have access to original video segments of any other users.

3. The server does not have access to any original video segments or gain any additional knowledge about the user's access patterns.
4. The server must authenticate the user's identity using his/her iris signal without any access to the signal in plaintext.

Details of PIM can be found in Section 6.

3.4 Security Model

Our security model assumes that both networks and external readers faithfully follow the protocols and do not collude with each other through side channels. As such, our focus is on preventing any adversarial behaviors to infer secret information from data exchanged among different components within the system. These secret information include:

1. sensitive video footage captured by the surveillance network,
2. iris signals represented in the garbled circuits, and
3. iris probe and identity of individuals at the reader.

For the security analysis of all our proposed protocols, we will thus assume a semi-honest adversarial behavior model to demonstrate that all messages exchanged in the protocol do not leak private information. We employ widely used efficient encryption such as AES and SHA that can provide long-term security using a long enough key size [3] to protect the system against brute force attacks. For the security analysis of Privacy Protected Video Surveillance Networks we remind to [66, 76, 10], while security of Amonymous Subject Identification protocol is analyzed in [39]. A discussion of the security of Privacy Imformation Management protocol is provided in Section 6. The no-collusion and semi-honest adversarial behavior model can also be difficult to guarantee in practice. In our prototype implementation, we monitor all the data traffic by firewalls and only specific types of data conformed to the protocols, as indicated in Figure 1, can pass through. This is accomplished by filtering out all but traffic to and from known addresses (within the internal networks), and specific ports. While such an approach cannot prevent collusion attacks within the network, it is effective in blocking external attacks. However we underline that the non-colluding semi-honest server can be used in many practical applications where server networks are cloud storage with no specific interest in the content of videos and biometric readers and cameras are simple IoT devices with limited available memory. Firewalls block online communication between the devices and memory constraints limit the possibility to store data on devices for later collusion.

4 Privacy Protected Video Surveillance Network

The privacy-protected video surveillance camera network is composed of a number of intelligent camera systems. The intelligent camera system is responsible for segmenting, tracking, encrypting, and obfuscating the visual imagery

corresponding to each individual based on the privacy bit from ASI. Figure 2 shows a screenshot of our system with redacted outputs from two cameras in our network, alongside with a video feed at the lower right from a separate camera showing the raw video scene. The details of the implementation can be found in a number of our prior works [66, 76, 10] and this section highlights a number of unique features in our design.

Segmentation: Segmentation identifies pixels in a single video frame that correspond to occupants. This is accomplished by a background subtraction algorithm in which each frame is compared with a background model of the environment to identify substantially different pixels as foreground. Background subtraction technologies are well-studied and our recent work showed that average precision can reach above 95% for well-controlled environments and 70% for challenging environments, including those with sudden changes in illumination, night camera, bad weather, and pan-tilt-zoom cameras [57]. Even though background subtraction on color cameras work well in our prototype, recently available RGB-depth cameras such as Microsoft Kinect have the potential to deliver even better segmentation performance due to their immunity to environmental noise.

Occlusion Handling: After segmentation, foreground pixels need to be assigned to individual occupants and a key challenge is occlusion. Environmental occlusion can typically be minimized by a careful placement of cameras. The dominant occlusion problem is between multiple occupants within a camera view. It is possible to build appearance models for different occupants using color, texture and motion, and then use these models for pixel assignment [66]. However, such an approach can be unreliable due to variations in appearance of an individual across the environment, and similarity in appearance among different individuals. As such, we have adopted a simple and conservative approach – raw pixels from the occluded regions are encrypted using the credentials from ALL individuals in the regions. Such an approach cannot provide full access of privacy video as stated in goal 1 in 3.3. Nevertheless, the alternative of letting each individual to have access may violate privacy guarantee (goal 2 in 3.3) and is deemed a more serious offense.

Tracking: Tracking refers to the process of linking pixels corresponding to the same individual across time and camera views. The problem for tracking a lone person in a camera view is trivial. It is important to point out that we enforce this single-person view for the camera at the entrance when a person first passes through the biometric reader. An example is shown in the top left picture of Figure 1. This allows the surveillance network to unmistakably relate the decision bit from ASI to the visual appearance of one person. Situations for the rest of the cameras are more complicated. Tracking of multiple individuals without occlusion in a single camera view is still straightforward - visual objects in consecutive frames are assigned to the same individual based on their spatial proximity. When occlusion occurs, tracking is unnecessary as our occlusion handling does not differ-

entiate different individuals. However, after occlusion, it is challenging to re-establish tracking of separate individuals using only appearance models. To increase robustness, we exploit overlap between adjacent camera views to resolve ambiguity after occlusion. Figure 2 shows an example with two persons occluding each other in one camera view but not at all in the other view. Tracking information from the non-occluded view can be related to the occluded one in resolving ambiguity. In fact, overlapping view between different cameras is crucial in relaying tracking status from one camera to another in covering a large surveillance environment. In [76], we have developed an optimal camera placement strategy to compute the minimum number of cameras and their locations so that every point within an arbitrary-shaped surveillance environment can be observed by at least two cameras [76]. Occlusions involving $n > 2$ individuals can theoretically be resolved when every point can be observed by n or more cameras. However, the number of cameras required can be significantly higher and may not be practical in most situations.

Obfuscation and Preservation: Segmentation and tracking allow us to identify all the raw pixels across different camera views that correspond to a specific individual. If this individual is marked by ASI for protection, these pixels will need to be redacted and their original values securely preserved. For preservation, those raw pixel values at each frame are padded with black background to make a rectangular frame and compressed using a standard video compression software. The compressed video stream will be further encrypted and details of the encryption are provided in Section 6.1. To redact these sensitive regions so as to provide a publicly accessible surveillance video, a myriad of different obfuscation schemes are provided ranging from black box, colored silhouette to full object removal [66]. Some examples are shown in Figure 2.

5 Anonymous Subject Identification

The Anonymous Subject Identification (ASI) protocol is a SMPC protocol that supports anonymous matching of biometric signals between the biometric reader and the iris-database server. In this section, we first review the basics of garbled circuits (GC) and provide a simple GC implementation of the iriscode matching procedure. Then, we describe a novel iris masking technique that simplifies the on-line operations on the encrypted data.

5.1 Fundemental GC primitives

Garbled Circuit (GC) approach [72] is an efficient computationally-secure method for evaluating a binary circuit with private input bits from two parties. While the circuit itself is public, GC guarantees that neither parties can discover the intermediate value and the final answer is available to at least

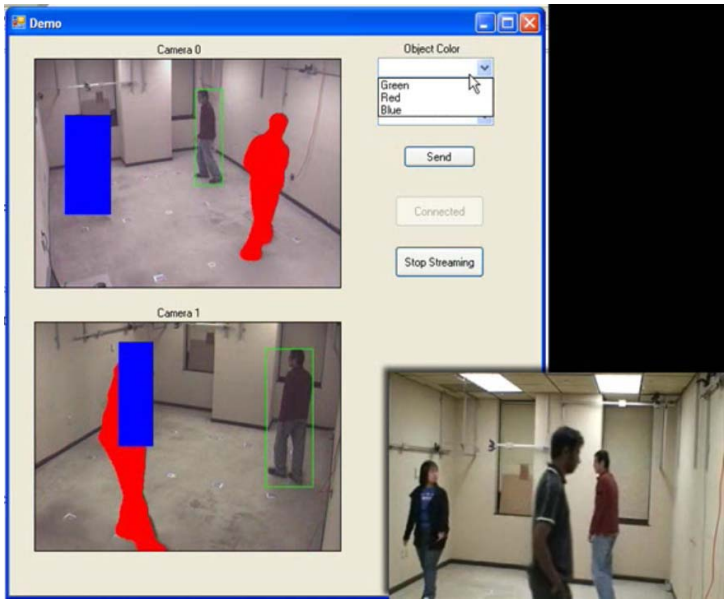


Fig. 2 The user interface on the left shows two camera views from our system. Two individuals (blue box and red silhouette) are redacted - note that the two different camera views allow accurate tracking despite the occlusion in the bottom view. The actual scene is shown in the lower right corner - it is captured by a hand-held camera that is not part of the system.

one party. There are different variations of GC but we will focus on the simple case where one party, Garbler, prepares a *garbled circuit* and sends it to the other party, Evaluator, for evaluation. Garbler implements the binary circuit and encrypts its private inputs by representing different logical states of each wire with *garbled values*. Specifically, for each wire W_i of the circuit, the garbler randomly chooses a pair of complementary *garbled values* consisting of two secrets, $w_i^0, w_i^1 \in \{0,1\}^t$, where w_i^j is the garbled value corresponding to logical value j on wire W_i and t is the security parameter. It is crucial to note that the knowledge of w_i^j does not reveal j.

Before describing the protocol for building and evaluating a garbled circuit, we first describe an interactive protocol called 1-out-of-2 Oblivious Transfer (OT), denoted as $w^r := OT(r; w^0, w^1)$. OT uses a public-key cipher to map Evaluator's private binary input $r \in \{0,1\}$ to a garbled value w^r on the input wire. OT guarantees that Garbler knows nothing about r, and Evaluator gets w^r but knows nothing about w^{1-r}, the complementary garbled value. Details of OT are given in Protocol 1.

We provide a brief explanation of Protocol 1 here and detailed security analysis and implementations can be found in [31, ch. 4]. After the initialization in steps 1 and 2, Garbler sends two public keys pk^0 and pk^1 corresponding to 0 and 1 to Evaluator in step 3. In step 4, Evaluator chooses pk^r based on

Require
Garbler: garbled values w^0, w^1;
Evaluator: input $r \in \{0, 1\}$
$\left(Enc_{pk}(\cdot), Dec_{sk}(\cdot)\right)$ is a public-key cipher with public key pk and private key sk.
Ensure
Evaluator gets w^r

1. Garbler prepares two random key pairs (pk^0, sk^0) and (pk^1, sk^1) to represent logical values 0 and 1 respectively.
2. Evaluator prepares a random key pair (pk, sk).
3. Garbler sends public keys pk^0 and pk^1 to Evaluator.
4. Evaluator computes $c := Enc_{pk^r}(pk)$ based on its private input r and sends c to Garbler.
5. Garbler decrypts c using both of its secret keys: $\widehat{pk}^0 := Dec_{sk^0}(c)$ and $\widehat{pk}^1 := Dec_{sk^1}(c)$.
6. Garbler encrypts the garbled values using the corresponding decrypted results: $c^0 := Enc_{\widehat{pk}^0}(w^0)$ and $c^1 := Enc_{\widehat{pk}^1}(w^1)$.
7. Garbler sends (c^0, c^1) to Evaluator.
8. Evaluator decrypts $w^r := Dec_{sk}(c^r)$.

Protocol 1: 1-out-of-2 Oblivious Transfer $w_r := OT(r; w^0, w^1)$

her secret and uses it to encrypt her chosen public key pk for Garbler. In step 5, Garbler gets $\widehat{pk}^r$ due to the matching keys, but without the knowledge of pk, it is unable to identify which one is correct. The remaining steps show how Garbler transfers the encrypted garbled values to Evaluator, which eventually decrypt the chosen w^r. The extra information c^{1-r} does not reveal any information about w^{1-r} because Evaluator does not have the secret key corresponding to $\widehat{pk}^{1-r}$.

The expensive public key operations can be treated as pre-computation [31]: during the setup phase, the full OT procedure is executed based on randomly chosen values r', w'^0, and w'^1 so that Evaluator obtains $w'^{r'}$. During the online phase, Evaluator sends to Garbler the masked bit $b := r' \oplus r$ based on the real private input r exclusive-or ($\oplus$) with r'. If b is 0, Garbler knows that Evaluator has the same input as in the setup phase and sends $(w'^0 \oplus w^0, w'^1 \oplus w^1)$. If b is 1, Garbler sends $(w'^1 \oplus w^0, w'^0 \oplus w^1)$. Finally, Evaluator XORs the r-th value in the returned duple with the stored $w'^{r'}$ to obtain w^r. The online stage requires only XOR operations.

Using the OT protocol, the GC protocol is fully specified in Protocol 2. In step 1, Garbler randomly generates all the garbled values for all wires. The garbled values corresponding to the private inputs of Garbler can be directly sent to Evaluator (step 2) while those corresponding to the private inputs of Evaluator need to be revealed with a 1-out-of-2 OT (steps 5-6). Steps 3-4 and 7 form the core of GC's encrypting and evaluating the circuit. The garbled table $G(f_l)$ in step 3 encrypts the output garbled value as $\text{SHA}(w_i^r \parallel w_j^s \parallel l) \oplus w_k^{f_l(r,s)}$, which can only be decrypted if Evaluator has the correct input garbled values. As the same set of inputs can be used by multiple gates, we include the gate index l and collision-resistant hash function SHA to make the encryption of different gates independent from each other. The random

Require
Garbler: private input $X \in \{0,1\}^{|\mathbf{W}_G|}$;
Evaluator: private input $Y \in \{0,1\}^{|\mathbf{W}_E|}$;
Ensure
Garbler gets $\mathbf{C}(X,Y)$ where $\mathbf{C}$ is a public circuit consisting of a set of binary gates $\mathbf{GA}$, a set of wires $\mathbf{W}_G$ for input X, wires $\mathbf{W}_E$ for input Y, and connecting wires $\mathbf{W}_I$.

1. Pre: Garbler randomly selects a pair of garbled values corresponding to 0 and 1 for each wire:

$$\mathbf{GW} := \{(w_i^0, w_i^1) \in \{0,1\}^t \times \{0,1\}^t \text{ for each } W_i \in \mathbf{W}_G \bigcup \mathbf{W}_E \bigcup \mathbf{W}_I\}.$$

2. Pre: Garbler commits its private inputs to the garbled values and sends to Evaluator:

$$\mathbf{GW}_G := \{w_i^{x_i} \text{ from } (w_i^0, w_i^1) \in \mathbf{GW} \text{ corresponding to each } x_i \in X\}.$$

3. Pre: Each gate $f_l(W_i, W_j) = W_k \in \mathbf{GA}$ is a truth-table with 4 entries. To garble each entry $f_l(r,s)$ where $r,s \in \{0,1\}$, Garbler creates the following tuple (based on [42]):

$$\Big(\text{SHA}(w_i^r \parallel w_j^s \parallel l) \oplus w_k^{f_l(r,s)}, \text{SHA}(w_k^{f_l(r,s)})\Big) \tag{1}$$

   where SHA denotes the SHA-256 hash function, $\parallel$ means concatenation and $\oplus$ is exclusive-or. Garbler forms the garbled gate $G(f_l)$ which is a random permutation of (1) for all $r,s \in \{0,1\}$.
4. Pre: Garbler sends the collection of garbled gates $\mathbf{GGA} := \{G(f_l), \forall f_l \in \mathbf{GA}\}$ to Evaluator.
5. Pre: Garbler and Evaluator start the pre-computation steps of OT for Evaluator to learn the garble values corresponding to its private inputs:

   $\mathbf{GW}_E := \{w_i^{y_i} := OT(y_i; w_i^0, w_i^1)$ where $(w_i^0, w_i^1) \in \mathbf{GW}$ corresponding to each $y_i \in Y\}$.

6. Garbler and Evaluator complete OT. Evaluator gets $\mathbf{GW}_E$.
7. Evaluator proceeds to recursively evaluate each garbled gate $G(f_l) \in \mathbf{GGA}$ starting with those based on the inputs. Using the gate index l and previously obtained input garbled values w_i' and w_j' where $w_i' \in \mathbf{GW}_G$ and $w_j' \in \mathbf{GW}_E$, Evaluator decrypts the first element in (1) for each of the 4 entries in the truth table as follows:

$$\text{SHA}(w_i^r \parallel w_j^s \parallel l) \oplus w_k^{f_l(r,s)} \oplus \text{SHA}(w_i' \parallel w_j' \parallel l) \tag{2}$$

   and obtains the unique output garbled value by hashing (2) and validating it with the second element in (1).
8. Following the same procedure, Evaluator can evaluate the entire $\mathbf{GGA}$ in a breadth-first search order to obtain the final garbled value $w^{\mathbf{C}(X,Y)}$.
9. Depending on whether the output is private to Garbler, Garbler can either decrypt $w^{\mathbf{C}(X,Y)}$ for Evaluator or provide the output table to Evaluator with each entry encrypted by the corresponding garbled value $w^{\mathbf{C}(X,Y)}$.

Protocol 2: Garbled Circuit $\mathbf{C}(X,Y) := GC_{\mathbf{C}}(X;Y)$

permutation ensures that no consistent patterns are revealed to Evaluator. The second hash value in (1) ensures that Evaluator can identify the correct output garbled value during evaluation in step 7. The entire circuit can then be recursively evaluated in step 8 following a similar manner. The final output value can be either shared with or withheld from Garbler as described in step 9.

All the steps in Protocol 2 that do not depend on the private inputs from Evaluator are marked as pre-computation. This classification is important to complexity reduction in our application. In our iriscode matching circuit, the biometric server plays the role of Garbler that encrypts a matching circuit based on the entire gallery. The biometric reader is Evaluator waiting for the input of a biometric probe. The pre-computation steps can be completed during the system setup time. The complexity reduction effort will thus concentrate on the online portion when a user is actually waiting. In addition, the optimization effort will focus on reducing the number of non-XOR gates because of the *free-XOR* gate technique introduced in [30]. The idea is based on the observation that complementary garbled values of a wire need to be statistically independent. As such, it is sufficient to have $w^1 := w^0 \oplus R$, where R is a secret random number at the garbler. If the garbled values of both inputs to a XOR gate are related by the same R, the evaluation of the garbled table of the gate (step 7 of Protocol 2) can be replaced by a single XOR function on the input garbled values, effectively rendering it "free" [30]. As the XOR gate is not universal, there could be other binary non-XOR gates remained. Thus, the goal of optimization is to implement our circuit with as few non-XOR gates as possible.

Another important consideration is that this GC implementation is not reusable in the sense that a new circuit is required for each query probe. This is because multiple different probes from the same user can reveal the complementary garbled values of input wires that can fully/partially decrypt the associated gates. There are recent work on reusable GC [22] but it could lead to a significant increase in circuit size [20]. Luckily, with multiple circuits preloaded to the reader during the setup phase, this lack of reusability has only a minor impact on the time complexity of the online computation.

5.2 GC-based iriscode matching

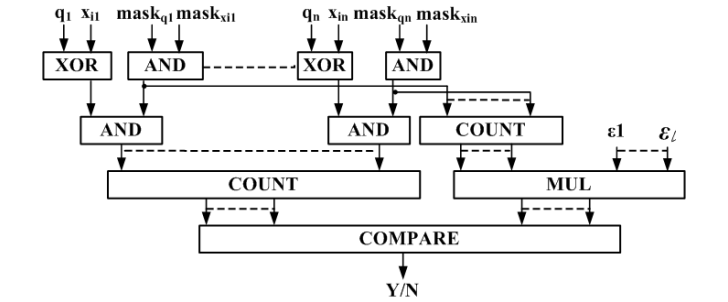
In our proposed system, the biometric server has an N-member iriscode gallery $DB := \{(\mathbf{x}_1, mask_{\mathbf{x}_1}) \ldots, (\mathbf{x}_N, mask_{\mathbf{x}_N})\}$. $\mathbf{x}_i$ denotes the n-bit iris feature of member i and $mask_{\mathbf{x}_i}$ is the corresponding binary mask that zeros out the unusable portion of the irises due to occlusion by eyelids and eyelash, specular refections, boundary artifacts of lenses, or poor signal-to-noise ratio. The reader captures a probe $(\mathbf{q}, mask_{\mathbf{q}})$ from the user and evaluates the GC, which produces a match if there exists at least an $\mathbf{x}_i \in DB$ such that $d(\mathbf{q}, \mathbf{x}_i) < \epsilon$ for a similarity threshold ϵ. $d(\mathbf{q}, \mathbf{x}_i)$ is a modified Hamming Distance (HD) defined below [12]:

$$d(\mathbf{q}, \mathbf{x}_i) := \frac{D(\mathbf{q}, \mathbf{x}_i)}{M(\mathbf{q}, \mathbf{x}_i)} = \frac{\| (\mathbf{q} \oplus \mathbf{x}_i) \cap mask_{\mathbf{q}} \cap mask_{\mathbf{x}_i} \|}{|mask_{\mathbf{q}} \cap mask_{\mathbf{x}_i} \|} \tag{3}$$

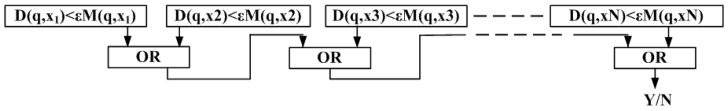
where $\cap$ and $\| \cdot \|$ denote logical-and and the norm of the binary vector respectively.

Our GC implementation has the server as Garbler with private inputs DB and ϵ, and the reader as the evaluator with private iriscode $(\mathbf{q}, mask_{\mathbf{q}})$. As

Protocol 2 is applicable to any boolean circuits, we focus on the schematic of our iriscode matching circuit here. Figure 3(a) shows the circuit for private iriscode matching between the probe $\mathbf{q}$ and the entry $\mathbf{x}_i$ in the database. It uses the basic garbled circuits (XOR, AND, and MULtiplication), a COUNT circuit to compute the number of ones in its input [6], and a COMPARE circuit to check if the first input is lower than the second input [29]. Given the fact that division in (3) is a complicated circuit [32] and multiplication involves fewer gates than division [28], we roll the denominator $M(\mathbf{q}, \mathbf{x}_i)$ of (3) into the similarity threshold ϵ and test whether $D(\mathbf{q}, \mathbf{x}_i) < \epsilon \cdot M(\mathbf{q}, \mathbf{x}_i)$. Since all computation should be computed over integers and ϵ is a decimal number in the range $[0, 1]$, we pre-multiply ϵ by 2^l and round it to an integer in the range $[0, 2^l]$ before taking part in the multiplication circuit with $M(\mathbf{q}, \mathbf{x}_i)$. Also, $D(\mathbf{q}, \mathbf{x}_i)$ is left shifted by l bits so the COMPARE function checks the result of $D(\mathbf{q}, \mathbf{x}_i) \cdot 2^l < (\epsilon \cdot 2^l) \cdot M(\mathbf{q}, \mathbf{x}_i)$. In order to highlight the overall structure of the circuit, we hide the scale-up processing and use $D(\mathbf{q}, \mathbf{x}_i)$ and ϵ instead of $D(\mathbf{q}, \mathbf{x}_i) \cdot 2^l$ and $\epsilon \cdot 2^l$ in Figure 3(a).



(a) Sub-circuit: $D(\mathbf{q}, \mathbf{x}_i) < \epsilon M(\mathbf{q}, \mathbf{x}_i)$



(b) Entire matching circuit with N sub-circuits from (a) followed by a disjunctive aggregation

Fig. 3 Circuit design for private iriscode matching

The output of the sub-circuit $D(\mathbf{q}, \mathbf{x}_i) < \epsilon \cdot M(\mathbf{q}, \mathbf{x}_i)$ cannot be made available to the server in plaintext, otherwise the server will know the exact entry that matches the probe and reveal the user's identity. In Figure 3(b), we use OR gates to connect the outputs of all COMPARE sub-circuits $D(\mathbf{q}, \mathbf{x}_i) < \epsilon \cdot M(\mathbf{q}, \mathbf{x}_i)$ for $i \in \{1, \ldots, N\}$ together. In the end, only the final output of all OR gates will be decoded and shared by two parties.

5.3 Reduction of GC gates' amount through iris masks

The online complexity of our GC circuit depends on two sets of parameters: the size of the circuit and the security parameters for the basic OT and GC protocols as discussed in Section 5.1. Our focus is to reduce the size of the circuit, while relying on publicly-available software library to provide state-of-the-arts performance on the OT and GC protocols. Externally, the size of the circuit is determined by the size of the gallery N, the length of the iriscode n, and the precision parameter l. While these are design parameters based on applications, we can optimize the circuit by reducing the number of gates to carry out the same operations.

As explained in Section 5.1, our goal is to minimize the number of non-XOR gates used in our circuit. Using the implementation of basic blocks from [30], the numbers of non-XOR gates used by different functions in our circuit implementation depicted in Figure 3 are tabulated in Table 1. As expected, the complexity of all functions grows linearly with the number of iriscodes N in the gallery. For iriscode matching, the size of the iriscode n is typically much larger than the precision parameter l. As such, the AND gates and COUNT functions dominate the complexity over other functions.

Table 1 Complexity of different components in Figure 3

Function	Number of non-XOR gates
AND	$2Nn$
COUNT	$2N(n - \log(n+1))$
COMPARE	$Nl\log(n)$
MUL	Nl
OR	N

As shown in Figure 3, the abundance of AND gates is due to the incorporation of masks in the distance computation. If the masks were treated as public information, all the AND gates and one COUNT function could be eliminated and the complexity could be significantly reduced. While the iris feature is obviously private data, it is unclear if the mask itself contains any sensitive information for identification. Prior schemes such as that in [40] treated masks as public information without quantifying the possible loss in privacy. There are other studies such as [35] that showed eyelashes positions, which made up a significant portion of the mask, had inherent correlation and could be used to infer important ethnic information about an individual. To the best of our knowledge, the privacy leakage through iris masks has not been statistically quantified in any previous studies. Using a publicly-available iriscode database CASIA [62], which contains multiple iriscodes for more than 290 individuals, we statistically measure the difference between the Hamming distances of iris masks for the same individuals and for different individuals. We found that iris masks from the same individual demonstrate correlations that are not present across different individuals. As such, *iris masks should be considered as private*

information at the server and not shared with the external biometric reader. The details of the experimental results are provided in Section 7.1.

Even though the actual masks should not be shared, we consider a different approach to simplify the usage of masks in our GC implementation. A typical mask contains information about eyelashes, eyelids, specular reflections, or other noise, and most of their positions do not vary significantly from one individual to another. As such, it is possible that the positions of these imperfections do not significantly affect the distance computation. A common iris mask can therefore be designed to replace the individual masks without much loss in precision. The common mask can be created by ORing all the available masks in the database. Our experiments in Section 7.2 show that using such a common mask on CASIA results in only less than 1% drop in recognition performance when compared with using individual masks. While it is our ongoing work to see if such a conclusion can be scaled up to a much larger database, we present here the design of a simplified iriscode matching circuit based on a publicly-available mask common among all iriscodes. The simplified GC sub-circuit for $D(\mathbf{q}, \mathbf{x}_i) < \epsilon M(\mathbf{q}, \mathbf{x}_i)$ is shown in Figure 4. We use MASK to denote the common mask and highlight all the function blocks that can be pre-computed. Similar to the scenario in which the masks are treated as public information, this circuit eliminates all the AND gates and one COMPARE function. Instead, MASK_FILTER functions are added which only accept the iriscodes into the matching processing with the set of corresponding masks. As demonstrated in our experiments in Section 7, the use of a common mask results in a speedup factor of over 3.7.

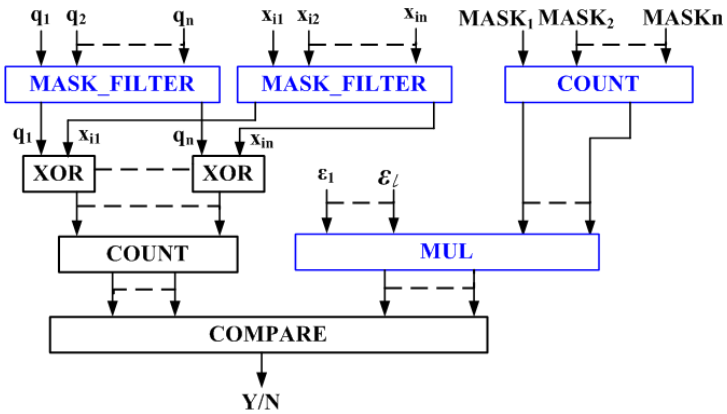


Fig. 4 Simplified GC sub-circuit for $D(q, X_i) < \epsilon M(q, X_i)$

6 Privacy Information Management (PIM)

In this section, we describe the Privacy Information Management (PIM) system that supports anonymous authentication of privacy information retrieval using biometric signals. Our proposed design consists of two protocols: the first one is the encryption of the privacy imagery in video, based on the biometric signal obtained during the ASI process as described in Section 5. The second protocol is invoked during the retrieval process where the decryption is performed using the biometric signal from a remote location. These two protocols are described below.

6.1 Privacy Information Encryption

The privacy information encryption protocol is executed right after the server has ascertained, via the ASI module from Section 5, that the subject entering the surveillance needs to be protected. The protocol involves three parties: the biometric reader (Reader), the server (Server), and the camera network (Camera) as described in Section 4. The objective is to let Camera use the biometric information from Reader to encrypt the privacy videos, which are then stored in Server for later retrieval. The details of the privacy information encryption protocol are given in Protocol 3, where we assume all transmission are made through secure channels, i.e. eavesdroppers are not able to observe transmitted messages and attackers cannot modify or erase the messages.

Require
Reader: Biometric probe $\mathbf{q}$;
Camera: Video segment v containing the protected subject;
Server: No private inputs.
Ensure
The following record is stored at Server:

$$(\mathrm{AES}_k(v), k \oplus r, \mathbf{q} \oplus s) \tag{4}$$

where k is a random AES key (256-bit), $\mathrm{AES}_k(v)$ is the AES-encrypted video v with key k, r is a 256-bit random number, and s is a random number as long as the iris probe q. In addition, (r, s) are provided by the Reader to the user as the retrieval code.

1. Upon entry of a new user whose identity has been anonymously authenticated in ASI module, Reader randomly prepares an AES key k and two random numbers r and s. (r, s) along with the timestamp are provided back to the user for later retrieval.
2. Reader sends k to Camera, and sends $(k \oplus r, \mathbf{q} \oplus s)$ to Server.
3. Camera computes $\mathrm{AES}_k(v)$, and sends $\mathrm{AES}_k(v)$ to Server.
4. Server indexes the record (4) by the time-of-the-day and stores it in the encrypted video database

Protocol 3: Privacy Information Encryption

Security Analysis of Protocol 3

As Protocol 3 is a storage protocol with no specific output, its ideal functionality is a no-op and the goal of the analysis is to ensure that the information exchanged is indistinguishable from uniform random data.

Reader: No data is sent to Reader so it does not learn any private information from Camera or Server.

Camera: It receives k from Reader. k is a one-time randomly chosen AES key that will be associated with the individual while he/she is in the surveillance area. As it is a one-time encryption key, the Camera will not be able to use this information to relate to any previous recordings of the same individual.

Server: It receives $\text{AES}_k(v)$ from Camera, and $(k \oplus r, \mathbf{q} \oplus s)$ from Reader. As the Server has databases of videos and iriscode of the same individual, it is important to ensure that the new message does not allow the Server to infer the identify.

- For the AES encrypted video $\text{AES}_k(v)$, we shall assume *the 256-bit key length is robust against brute-force attacks on* k. The key k is a one-time encryption key so previous recordings of the same individual at the Server are independent.
- The security of $k \oplus r$ is ensured by the one-time pad with a private key r that is only available at the Reader. The query probe $\mathbf{q}$ is encrypted with another one-time randomly generated private key s, which is either 2048 or 9600 bit long. As such, brute-force attacks on r and s are not possible.
- However, Server has an iriscode $\mathbf{q}'$ in its gallery that is similar or possibly identical to $\mathbf{q}$ and it can use $\mathbf{q}'$ to attack s. However, the retrieval code (r, s) is only available at the Reader, making it impossible for Server to ascertain if there is a match. Also, as s changes every time, different records from the same individual are independent from each other. A drawback of this approach is that the user needs the correct retrieval code for retrieval. Details of the retrieval procedure can be found in Protocol 4.

6.2 Privacy Information Retrieval

The privacy information retrieval protocol, described in Protocol 4, is used when a user (Requester) wants to retrieve his/her privacy videos from the video database stored at Server. We assume that communications are made through secure channels. The security goals in Section 3 state that the user must be authenticated with the biometric and only allowed to access the associated records. The encryption protocol introduced in Section 6.1 further requires the user to possess the correct retrieval code in addition to the iris biometric to complete the process.

Correctness of Protocol 4

In steps 1 and 2, Server identifies all the relevant records pertinent to Requester's query. For each record, Server generates a random key t_i in step

Require
Requester: Probe $\mathbf{q}'$ and retrieval code (r', s');
Server: A set of records in the form of (4);
Ensure
Requester obtains the associated private video.

1. Requester sends time-stamp query to Server.
2. Server identifies all records with matching time. Denote this set of records as

$$A := \{\big(\mathrm{AES}_{k_i}(v_i), k_i \oplus r_i, \mathbf{q_i} \oplus s_i\big) : i = 1, 2, \ldots, N\} \tag{5}$$

   where N is the number of matching records.
3. For each record in A,
   (a) Server generates a 256-bit uniformly random number t_i.
   (b) Server creates a garbled circuit $GC(\mathbf{q_i} \oplus s_i; \mathbf{x})$ that returns t_i if $d(\mathbf{q_i} \oplus s_i, \mathbf{x}) < \epsilon$ and sends the GC to Requester.
   (c) Requester uses her private input $\mathbf{x} := \mathbf{q}' \oplus s'$ and jointly evaluates the garbled circuit with Server to obtain the output t'_i, which is available only to the Requester.
   (d) Server sends the following to Requester:

$$\big(\mathrm{AES}_{k_i}(v_i), k_i \oplus r_i \oplus t_i\big) \tag{6}$$

   (e) Requester uses her own r' and GC's output t'_i to compute:

$$k'_i := k_i \oplus r_i \oplus t_i \oplus r' \oplus t'_i$$

   (f) Requester tries to decrypt $\mathrm{AES}_{k_i}(v_i)$ with k'_i.

Protocol 4: Privacy Information Retrieval

3(a) to encrypt $k_i \oplus r_i$ as in (6). Server also creates a new garbled circuit $GC(\mathbf{q_i} \oplus s_i; \mathbf{x})$ in step 3(b) to encrypt t_i and sends it to Requester, which provides her private input $\mathbf{x} := \mathbf{q}' \oplus s'$. If Requester has a matching iriscode, i.e. $d(\mathbf{q_i}, \mathbf{q}') < \epsilon$, and a matching second part of the retrieval code, $s_i = s'_i$, then joint execution of the GC will result in $d(\mathbf{q_i} \oplus s_i, \mathbf{q}' \oplus s') < \epsilon$ as the XOR does not affect the hamming distance calculation and the GC will return $t'_i = t_i$ in step 3(c). If the inputs do not match, GC will return a random output. In step 3(d), Server sends the encrypted video $\mathrm{AES}_{k_i}(v_i)$ and encrypted key $k_i \oplus r_i \oplus t_i$ to Requester, who can easily retrieve the AES key k_i in step 3(e) with a matching first part of the retrieval code $r_i = r'$ and the correct GC output. The video can then be decrypted with the correct AES key in step 3(f).

Security Analysis of Protocol 4

The ideal functionality for Requester is to decrypt her videos but nothing else, and the ideal functionality for Server is no-op.

Requester: Based on steps 3(b) and (d), the following information is received by Requester:

$$B := \{(\mathrm{AES}_{k_i}(v_i), k_i \oplus r_i \oplus t_i, GC(\mathbf{q_i} \oplus s_i; \mathbf{x})) : i = 1, 2, \ldots, N\} \tag{7}$$

For a given record i, we claim that it is indistinguishable from uniform random data unless $r = r_i$, $s = s_i$, and $d(\mathbf{q_i}, \mathbf{q}) < \epsilon$. The reason is as follows:

first, if either of the last two conditions are not met, the garbled circuit will return an uniformly random value independent of t_i, which makes $k_i \oplus r_i \oplus t_i$ independent of $\text{AES}_{k_i}(v_i)$. This applies even if the retrieval code is stolen, i.e. $(r, s) = (r_i, s_i)$, the record still does not reveal any information without a matching biometric $\mathbf{q}'$. Second, if the two conditions are met but $r \neq r'$, the garbled circuit will produce the correct t_i but the unknown r_i still keeps $k_i \oplus r_i$ and $\text{AES}_{k_i}(v_i)$ secret. Finally, as k_i, t_i, and r_i are independent for different i's and so are the garbled circuits, Requester's capability of decrypting one record has no impact on others, thereby protecting privacy of videos not belonging to her.

Server: Server receives the time-stamp query from Requester in step 1. This step can reveal the possible time frame Requester is interested in. Alternatively, Server can send the entire database to Requester to plug this small information leak, but such an approach is not practical for large video databases. It is important to note that Server also receives additional information in step 3(c) during the joint evaluation of the garbled circuits. However, the result of the evaluation is not available to Server and the security of garbled circuits guarantees that Server does not know which videos Requester can successfully decrypt.

7 Implementation Details and Experiments

All our experiments are based on a subset of the CASIA Iris database, the CASIA-IrisV3-Lamp, from the Chinese Academy of Sciences Institute of Automation (CASIA) [62]. To ensure that we begin with a good quality set of samples, we remove erroneous samples which cannot produce accurate iriscode based on the Matlab feature generation code from [44]. As a result, 3763 samples from 292 individuals are included in our dataset. Our proposed system is implemented in Java and runs on an Intel Core i7-4790 CPU @3.60GHz 3.60GHz with 12GB RAM on 64-bit Windows 8.1 Enterprise.

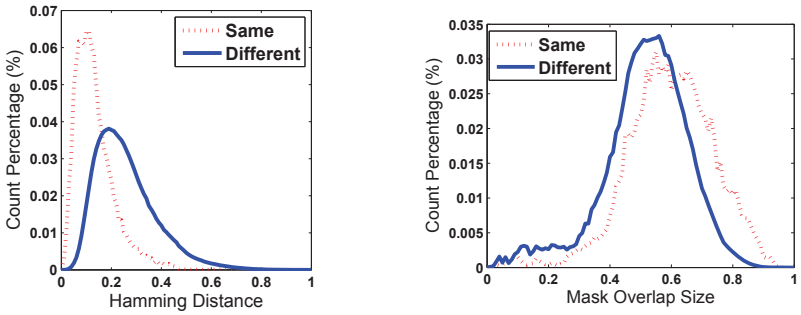
7.1 Privacy and similarity among iris masks

In this section, we study if the mask alone in an iriscode can leak privacy information. While there are different ways to measure privacy leakage, we argue that masks contain privacy information if two masks from the same individual are more similar than the masks from two randomly selected individuals. As there are no well established methods to measure similarity between masks, we consider two generic similarity measurements between binary strings: normalized Hamming distance (HD) and overlap size (OS) defined below:

$$d_{HD}(\mathbf{x}, \mathbf{y}) := \frac{1}{n}||\mathbf{x} \oplus \mathbf{y}|| \quad \text{and} \quad s_{OS} := \frac{1}{n}||\mathbf{x} \cap \mathbf{y}|| \tag{8}$$

Treating these measures on masks from same individuals and different individuals as random variables, we can formulate statistical tests to determine

if they are distinguishable. In our experiments, we have computed these measures for 28,006 pairs of masks between the same individuals and 7,050,197 pairs between different individuals. The distributions of the two measures for these two sets are shown in Figure 5(a) and (b).



(a) Hamming Distances (b) Masks Overlap Sizes

Fig. 5 Mask distance distributions

To test if the difference of these measures between from the same and different individuals are statistically significant, we utilize the distribution-free Wilcoxon Rank-Sum Test between these two samples [14, Ch.15]. We focus the analysis on d_{HD} here as that of s_{OS} is the same and leads to the same conclusion. We hypothesize that masks from the same individual are smaller in distance compared to those from different individuals. If this hypothesis is accepted, there is no identity information leaked through masks. In our test, the samples from the same individuals are labeled as X and the samples from different individuals as Y. Let μ_x and μ_y be the averages of X and Y respectively. The null hypotheses is $H_0 : \mu_x - \mu_y = 0$ and the alternative hypothesis is $H_a : \mu_x - \mu_y \neq 0$. When the samples from X and from Y are pooled into a combined sample of size $m + n$, these observations are sorted from smallest (rank 1) to largest (rank $m + n$). Then, the sum of ranks of all samples from X is considered as our test statistic W, i.e. $W = \sum_{i=1}^{m} R_i$ where R_i is the rank for the i-th sample of X. Due to the large sample size, the distribution of the test statistic $z = (W - \mu_W)/\sigma_W$ can be approximated by a standard normal distribution if H_0 is true where

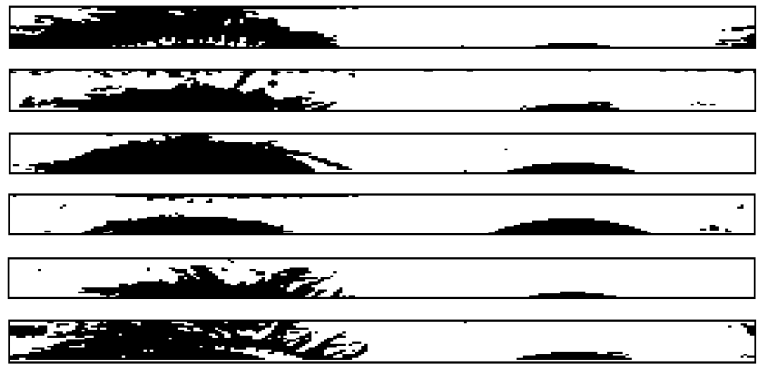
$$\mu_W = \frac{m(m+n+1)}{2} = 9.91 \times 10^{10}$$
$$\sigma_W^2 = \frac{mn(m+n+1)}{12} = 1.16 \times 10^{17}$$

At the confident level of 99%, H_0 is rejected if either $z \geq 2.58$ or $z \leq -2.58$. In our experiments, $W = 5.19 \times 10^8$ which implies that $z = -288.91$. The null hypothesis is therefore rejected. Based on these results, we conclude that

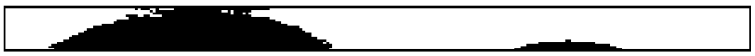
masks have inter-correlation among each individual, and therefore, should not be shared between the server and the biometric reader.

7.2 Common mask

Samples of masks from different individuals are shown in Figure 6(a). We can observe that there are a great deal of similarity among masks even from different individuals. Also, our earlier experiments depicted in Figure 5(a) indicate that there could be up to 50% bit difference even between masks from the same individual.



(a) Real masks from database



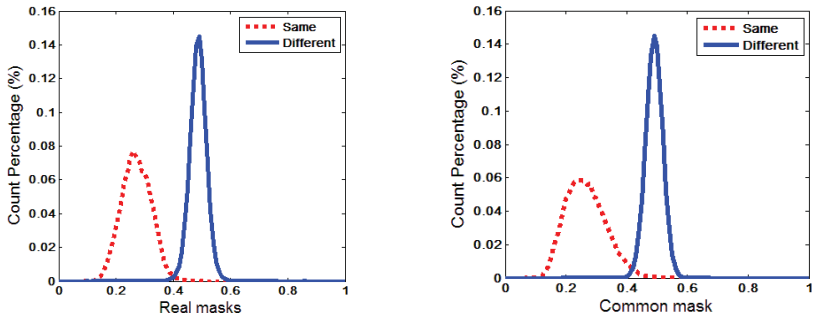
(b) Common mask

Fig. 6 Real masks and common mask

As such, it is conceivable to use a common mask to replace individual masks without much loss in precision. As we have pointed out in Section 5.3, the use of a common mask can significantly reduce the complexity of our GC circuits. To test this hypothesis, we use the following method to derive the common mask: first, we pre-align all iriscodes in the database with respect to a randomly chosen iriscode $\mathbf{x}$ by shifting each of them to a position that minimizes its distance from $\mathbf{x}$. The common iris mask is set to '1' at all bit positions where the percentages of the pre-aligned masks being '1' at those positions exceed an empirically-determined threshold λ. The common mask obtained from the CASIA iris database is shown in Figure 6(b).

Figure 7 shows the distribution of HDs using both real masks and the common mask. With $\epsilon = 0.41$, False Accept Rate (FAR) is 0.53% while False Reject Rate is 0.54% for the distribution computed with real masks. The best

FAR and FRR are 1.44% and 1.47% at $\epsilon = 0.43$ for the distribution with the common mask, based on setting λ to 80%. We can see that the accuracy in the case of common mask is reduced by less than 1%.



(a) Real masks (b) Common mask

Fig. 7 HD distributions

7.3 GC-based Anonymous Iris-code matching

We analyze the computation performance of our implementation of ASI using two sets of iriscodes: the first one is the 9600-bit code generated by the scheme in [44] and the second one is a downsampled version of the first set to conform with the 2048-bit code described in [12]. The reason of using the second set is to have a fair comparison with the scheme described in [8]. Downsampling should not affect the complexity measurements in this section and is used here as there is no public implementation of the scheme described in [12]. The basic building blocks are based on [29]. Other GC implementation can be used as our focus is on simplifying the GC circuit for iriscode matching instead of basic GC blocks. The complexity measurements of the two sets using different security parameters are shown in Table 2.

We do not include circuit construction and circuit transmission in our complexity measurements as they can be computed offline [7]. The setup time is based on OT precomputation which is needed every time when a new iris probe is presented to the reader. The setup time is independent of the size of the iris database but is related to the length of the iriscode, as shown in Table 2. The online time is the real computation time measured at the reader and the server for the entire circuit evaluation process after the pre-computation is completed. The overall time includes, in addition to the online time, the time for loading the circuits from the files, constructing the garbled table for each gate, and communicating between the two parties. All the results in Table 2 are measured by averaging the comparisons of 100 pairs of iriscodes in the database.

Table 2 Number of non-XOR gates, runtime (ms) and bandwidth (KB) based on different security parameters (bit)

n-bit	# non -XOR	Security Parameters	Setup Time	Online Time		Overall Time	Band-width
				Reader	Server		
Individual Masks							
2048	8349	80	17,127	15	62	142	571.5
		112	18,557	17	69	151	754.0
		128	18,798	18	75	159	845.7
9600	38654	80	87,033	62	285	611	2655.0
		112	88,256	63	287	641	3503.2
		128	93,278	76	323	686	3828.5
Common Mask							
2048	2059	80	8,850	7	19	38	133.7
		112	9,153	9	21	46	176.5
		128	9,289	10	24	49	197.9
9600	9641	80	43,337	18	64	162	626.1
		112	43,490	19	65	176	826.5
		128	43,696	22	74	189	926.7

The performance of the totally GC-based private iris-code matching is quite efficient: when we adopt an 80-bit security parameter, it takes 142 ms to compare two 2048-bit iris-codes with private iris features and masks. If the common mask is used, a speedup factor of up to 3.7 or 38 ms per comparison can be achieved. This is comparable to 14 ms as reported in [8]. Considering that longer cyphertexts will be required to provide longer-term security, we also measure the processing time using longer security parameters of 112 and 128 bits in Table 2. The execution time is increased by 12% for the individual masks and 29% for the common mask at the most. These are much smaller than the 62% increase for the hybrid protocol as reported in [55]. As such, our GC-only protocol is clearly preferred in the cases when longer term security is needed.

7.4 Privacy Information Management

In this section, we focus on the implementation details and performance analysis of the PIM system as described in Section 6. The privacy information encryption protocol described in Protocol 3 relies on an AES cipher for protecting the original video imageries. Specifically, we employed the video encryption model in [63, Ch.5] to encrypt the H.263 video bitstream pertinent to each individual. The encryption time of the 256-bit AES key is on average 10.16 ms. The decryption time is 9.37 ms. The main computation and storage burden of the entire protocol are dominated by the AES encryption the private video, which varies depending on the number of protected individuals and the time duration of each protected individual inside the surveillance perimeter. Since the AES implementation is not our original work, we do not analyze the computation and communication complexity of video encryption. The privacy information retrieval protocol described in Protocol 4 needs an additional

GC circuit to match the live probe from the requester with the stored data. We adopted 80-bit security parameter for our GC implementation. The result GC circuit has 2071 non-XOR gates with total runtime measured at 47ms for comparing a pair of 2048-bit iriscode combined with the same length random number and 9655 non-XOR gates at 188ms for 9600-bit inputs.

8 CONCLUSION

In this paper, we have proposed a framework for anonymous subject identification (ASI) and privacy information management (PIM) with biometric signals in a privacy-aware video surveillance system. Our ASI system uses iris patterns to determine the privacy protection status of an incoming individual. By capitalizing on the recent advancements in garbled-circuit (GC) based secure multi-party protocols, a novel GC-based implementation of the ASI system has been proposed in the first part of the paper. We have discovered that the complexity of the GC-based ASI system heavily depends on the use of individual iris masks. Our experiments have demonstrated that while making the masks public as suggested by other works can leak privacy information, using a common mask for all comparisons can significantly reduce the complexity with negligible loss in recognition accuracy. In the second part of the paper, we have designed a PIM system that protects all surveillance videos with privacy information and allows any user to anonymously access his/her own imageries. The proposed system uses the user's biometric signal and a retrieval code obtained during the ASI process to encrypt a secret key for unlocking the original video imagery. The retrieval process is based on a simple GC to authenticate the identity of the user, while guaranteeing that the user cannot gain any information about other users, and the server knows nothing about the identity of the user or the actual video contents. Future works include validation of the common mask assumption with a larger database, improved performance in GC-based similar iris search through hierarchical clustering of data, and a distributed implementation of the PIM system in a large camera network.

References

1. Avidan, S., Moshe, B.: Blind vision. In: Proceedings of the 9th European Conference on Computer Vision, pp. 1–13 (2006)
2. Barker, E., Barker, W., Burr, W., Polk, W., Smid, M.: Recommendation for key management. NIST Special Publication 800-57 (2012)
3. Barker, E.B.: Sp 800-57 rev. 4. recommendation for key management. part 1: General (2016)
4. Barni, M., Bianchi, T., Catalano, D., Raimondo, M.D., Labati, R.D., Failla, P., Fiore, D., Lazzeretti, R., Piuri, V., Piva, A., et al.: A privacy-compliant fingerprint recognition system based on homomorphic encryption and fingercode templates. In: Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on, pp. 1–7. IEEE (2010)
5. Barni, M., Droandi, G., Lazzeretti, R.: Privacy protection in biometric-based recognition systems: A marriage between cryptography and signal processing. Signal Processing Magazine, IEEE **32**(5), 66–76 (2015)

6. Barni, M., Guajardo, J., Lazzeretti, R.: Privacy preserving evaluation of signal quality with application to ECG analysis. In: Information Forensics and Security (WIFS), 2010 IEEE International Workshop on, pp. 1–6. IEEE (2010)
7. Beaver, D.: Precomputing oblivious transfer. Advances in CryptologyCRYPT095 pp. 97–109 (1995)
8. Blanton, M., Gasti, P.: Secure and efficient protocols for iris and fingerprint identification. Tech. rep., Cryptology ePrint Archive, Report 2010/627, 2010. http://eprint. iacr. org (2010)
9. Cavoukian, A., Stoianov, A.: Biometric encryption: A positive-sum technology that achieves strong authentication, security and privacy. Information and Privacy Commissioner, Ontario, Canada (2007)
10. Cheung, S.C., Venkatesh, M.V., Paruchuri, J., Zhao, J., Nguyen, T.: Protecting and managing privacy information in video surveillance systems. In: A. Senior (ed.) Protecting Privacy in Video Surveillance. Springer (2009)
11. Chinomi, K., Nitta, N., Ito, Y., Babaguchi, N.: Prisurv: Privacy protected video surveillance system using adaptive visual abstraction. In: Advances in Multimedia Modeling. Springer, pp. 144–154 (2008)
12. Daugman, J.: How iris recognition works. IEEE Transactions on Circuits and Systems for Video Technology **4**, 21–30 (2004)
13. De Marsico, M., Nappi, M., Riccio, D., Wechsler, H.: Mobile iris challenge evaluation (MICHE)-I, biometric iris dataset and protocols. Pattern Recognition Letters **57**, 17–23 (2015)
14. Devore, J.: Probability and Statistics for Engineering and the Science, Brooks/Cole Pub. Co., Monterey, California **704** (1991)
15. Dufaux, F., Ebrahimi, T.: Scrambling for video surveillance with privacy. 2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06) p. 160 (2006). DOI http://doi.ieeecomputersociety.org/10.1109/CVPRW.2006.184
16. Elezovikj, S., Ling, H., Chen, X.: Foreground and scene structure preserved visual privacy protection using depth information. In: Multimedia and Expo Workshops (ICMEW), 2013 IEEE International Conference on, pp. 1–4. IEEE (2013)
17. Erkin, Z., Franz, M., Guajardo, J., Katzenbeisser, S., Lagendijk, I., Toft, T.: Privacy-preserving face recognition. In: Privacy Enhancing Technologies, pp. 235–253. Springer (2009)
18. Fidaleo, D.A., Nguyen, H.A., Trivedi, M.: The networked sensor tapestry (nest): a privacy enhanced software architecture for interactive analysis of data in video-sensor networks. In: VSSN '04: Proceedings of the ACM 2nd international workshop on Video surveillance & sensor networks, pp. 46–53. ACM Press, New York, NY, USA (2004). DOI http://doi.acm.org/10.1145/1026799.1026809
19. Fontaine, C., Galand, F.: A survey of homomorphic encryption for nonspecialists. EURASIP Journal on Information Security **2007** (2007)
20. Gentry, C., Gorbunov, S., Halevi, S., Vaikuntanathan, V., Vinayagamurthy, D.: How to compress (reusable) garbled circuits. IACR Cryptology ePrint Archive **2013**, 687 (2013)
21. Goldreich, O.: Foundations of Cryptography: Volume II Basic Applications. Cambridge (2004)
22. Goldwasser, S., Kalai, Y., Popa, R.A., Vaikuntanathan, V., Zeldovich, N.: Reusable garbled circuits and succinct functional encryption. In: Proceedings of the forty-fifth annual ACM symposium on Theory of computing, pp. 555–564. ACM (2013)
23. Grother, P., Ngan, M.: Face recognition vendor test (frvt) performance of face identification algorithms. NIST Interagency Report **8009**, 2 (2014)
24. Grother, P., Quinn, G., Matey, J., Ngan, M., Salamon, W., Fiumara, G., Watson, C.: Irex iii performance of iris identification algorithms. Interagency report **7836** (2012)
25. Hao, F., Anderson, R., Daugman, J.: Combining cryptography with biometrics effectively. IEEE Transactions on Computers **55**(9), 1081–1088 (2006)
26. Hazay, C., Lindell, Y.: Efficient secure two-party protocols: Techniques and constructions. Springer (2010)
27. of Health, U.D., Services, H., et al.: Summary of the HIPAA privacy rule. Washington, DC: Department of Health and Human Services (2003)

28. Kolesnikov, V., Sadeghi, A., Schneider, T.: How to Combine Homomorphic Encryption and Garbled Circuits. Signal Processing in the Encrypted Domain pp. 100–121 (2009)
29. Kolesnikov, V., Sadeghi, A., Schneider, T.: Improved garbled circuit building blocks and applications to auctions and computing minima. Cryptology and Network Security pp. 1–20 (2009)
30. Kolesnikov, V., Schneider, T.: Improved garbled circuit: Free xor gates and applications. Automata, Languages and Programming pp. 486–498 (2008)
31. Lazzeretti, R.: Privacy preserving processing of biomedical signals with application to remote healthcare systems. Ph.D. thesis, Ph. D. thesis, PhD school of the University of Siena, Information Engineering and Mathematical Science Department (2012)
32. Lazzeretti, R., Barni, M.: Division between encrypted integers by means of garbled circuits. The 2011 IEEE Intl. Workshop on Information Forensics and Security (WIFS'11) (2011)
33. Lee, Y., Micheals, R.J., Filliben, J.J., Phillips, P.J.: Vasir: An open-source research platform for advanced iris recognition technologies. Journal of Research of the National Institute of Standards and Technology **118**, 218 (2013)
34. Lewis, S.M.: The fourth amendment in the hallway: Do tenants have a constitutionally protected privacy interest in the locked common areas of their apartment buildings? Michigan Law Review pp. 273–310 (2002)
35. Li, Y., Savvides, M., Chen, T.: Investigating useful and distinguishing features around the eyelash region. In: 2008 37th IEEE Applied Imagery Pattern Recognition Workshop. IEEE (2008)
36. Linnartz, J., Tuyls, P.: New shielding functions to enhance privacy and prevent misuse of biometric templates. In: Audio-and Video-Based Biometric Person Authentication, pp. 1059–1059. Springer (2003)
37. Lioudakis, G.V., Koutsoloukas, E.A., Dellas, N.L., Tselikas, N., Kapellaki, S., Prezerakos, G.N., Kaklamani, D.I., Venieris, I.S.: A middleware architecture for privacy protection. Computer Networks **51**(16), 4679–4696 (2007)
38. Liu, C., Wang, X.S., Nayak, K., Huang, Y., Shi, E.: ObliVM: A generic, customizable, and reusable secure computation architecture. In: IEEE Security and Privacy (2015)
39. Luo, Y., Cheung, S.C., Pignata, T., Lazzeretti, R., Barni, M.: An efficient protocol for private iris-code matching using garbled circuits. In: IEEE International Conference on Image Processing (ICIP 2012). Orlando, Florida, USA (2012)
40. Luo, Y., Cheung, S.C.S., Ye, S.: Anonymous biometric access control based on homomorphic encryption. In: IEEE International Conference on Multimedia & Expo. Cancun, Mexico (2009)
41. Luo, Y., Sen-ching, S.C.: Privacy information management for video surveillance. In: SPIE Defense, Security, and Sensing, pp. 871,207–871,207. International Society for Optics and Photonics (2013)
42. Malkhi, D., Nisan, N., Pinkas, B., Sella, Y.: Fairplay — a secure two-party computation system. In: USENIX (2004). http://www.cs.huji.ac.il/project/Fairplay
43. Martin, K., Plataniotis, K.N.: Privacy protected surveillance using secure visual object coding. Circuits and Systems for Video Technology, IEEE Transactions on **18**(8), 1152–1162 (2008)
44. Masek, L., Kovesi, P.: Matlab source code for a biometric identification system based on iris patterns. Tech. rep., The School of Computer Science and Software Engineering, The University of Western Australia (2003)
45. Nakashima, Y., Babaguchi, N., Fan, J.: Automatically protecting privacy in consumer generated videos using intended human object detector. In: Proceedings of the international conference on Multimedia, pp. 1135–1138. ACM (2010)
46. Nakashima, Y., Babaguchi, N., Fan, J.: Automatic generation of privacy-protected videos using background estimation. In: Multimedia and Expo (ICME), 2011 IEEE International Conference on, pp. 1–6. IEEE (2011)
47. Newton, E.N., Sweeney, L., Main, B.: Preserving privacy by de-identifying face images. IEEE Transactions on Knowledge and Data Engineering **17**(2), 232–243 (2005)
48. Organisations, C.C.P.S.: Common criteria for information technology security evaluation, part 2: Security functional components. Tech. Rep. CCIMB-2012-09-002 (2012)

49. Padilla-López, J.R., Chaaraoui, A.A., Gu, F., Flórez-Revuelta, F.: Visual privacy by context: proposal and evaluation of a level-based visualisation scheme. Sensors **15**(6), 12,959–12,982 (2015)
50. Paruchuri, J.K., Cheung, S.C.S., Hail, M.W.: Video data hiding for managing privacy information in surveillance systems. EURASIP Journal on Information Security **2009**(236139) (2009)
51. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. Pattern Analysis and Machine Intelligence, IEEE Transactions on **12**(7), 629–639 (1990)
52. Pinkas, B., Schneider, T., Smart, N., Williams, S.: Secure two-party computation is practical. Advances in Cryptology–ASIACRYPT 2009 pp. 250–267 (2009)
53. Qureshi, F.Z.: Object-video streams for preserving privacy in video surveillance. In: Advanced Video and Signal Based Surveillance, 2009. AVSS'09. Sixth IEEE International Conference on, pp. 442–447. IEEE (2009)
54. Rashwan, H.A., Solanas, A., Puig, D., Martínez-Ballesté, A.: Understanding trust in privacy-aware video surveillance systems. International Journal of Information Security pp. 1–10 (2015)
55. Sadeghi, A., Schneider, T., Wehrenberg, I.: Efficient privacy-preserving face recognition. Information, Security and Cryptology–ICISC 2009 pp. 229–244 (2010)
56. Saini, M., Atrey, P.K., Mehrotra, S., Kankanhalli, M.: Adaptive transformation for robust privacy protection in video surveillance. Advances in Multimedia **2012**(4) (2012)
57. Sajid, H., Cheung, S.C.S.: Background subtraction for static & moving camera. In: Image Processing (ICIP), 2015 IEEE International Conference on, pp. 4530–4534. IEEE (2015)
58. Schiff, J., Meingast, M., Mulligan, D., Sastry, S., Goldberg, K.: Respectful cameras: Detecting visual markers in real-time to address privacy concerns. In: International Conference on Intelligent Robots and Systems (IROS), pp. 971–978. Springer (2007)
59. Senior, A., Pankanti, S., Hampapur, A., L. Brown, Y.L.T., Ekin, A.: Blinkering surveillance: Enabling video privacy through computer vision. IEEE Security and Privacy **3**, 50–57 (2005)
60. Songhori, E.M., Hussain, S.U., Sadeghi, A.R., Schneider, T., Koushanfar, F.: Tinygarble: Highly compressed and scalable sequential garbled circuits. In: IEEE S & P (2015)
61. Sundstrom, E., Burt, R.E., Kamp, D.: Privacy at work: Architectural correlates of job satisfaction and job performance. Academy of Management Journal **23**(1), 101–117 (1980)
62. Tan, T., Sun, Z.: Casia-irisv3. Tech. rep., Chinese Academy of Sciences Institute of Automation, http://www.cbsr.ia.ac.cn/IrisDatabase.htm (2005)
63. Uhl, A., Pommer, A.: Image and Video Encryption: From Digital Rights Management to Secured Personal Communication, vol. 15. Springer (2004)
64. Uludag, U., Jain, A.: Securing fingerprint template: Fuzzy vault with helper data. Proceedings of CVPR Workshop on Privacy Research In Vision (2006)
65. Uludag, U., Pankanti, S., Prabhakar, S., Jain, A.: Biometric cryptosystems: issues and challenges. Proceedings of the IEEE **92**(6), 948–960 (2004)
66. Venkatesh, M.V., Cheung, S.C., Zhao, J.: Efficient object-based video inpainting. Pattern Recognition Letters : Special issue on Video-based Object and Event Analysis (2008). DOI doi:10.1016/j.patrec.2008.03.011
67. Wactlar, H., Stevens, S., Ng, T.: Enabling Personal Privacy Protection Preferences in Collaborative Video Observation. NSF Award Abstract 0534625, http://www.nsf.gov/awardsearch/showAward.do?award Number=0534625
68. Wada, J., Kaiyama, K., Ikoma, K., Kogane, H.: Monitor camera system and method of displaying picture from monitor camera thereof. Matsushita Electric Industrial Co. Ltd. (2001)
69. Wasson, J.C.: Ferpa in the age of computer logging: school discretion at the cost of student privacy (2002)
70. Wickramasuriya, J., Datt, M., Mehrotra, S., Venkatasubramanian, N.: Privacy protecting data collection in media spaces. In: ACM International Conference on Multimedia. New York, NY (2004)
71. William, S., Lawrie, B.: Computer Security: Principles And Practice, third edn. Pearson Education, Inc. (2015)

72. Yao, A.C.: Protocols for secure computations. In: Proceedings of the 23rd Annual IEEE Symposium on Foundations of computer science (1982)
73. Ye, S., Luo, Y., Zhao, J., Cheung, S.: Anonymous Biometric Access Control. EURASIP Journal on Information Security **2009**(Article ID 865259, 17 pages) (2009)
74. Yu, X., Babaguchi, N.: Privacy preserving: Hiding a face in a face. In: ACCV, pp. 651–661 (2007)
75. Zhao, J., Cheung, S.: Multi-camera surveillance with visual tagging and generic camera placement. In: Proceedings of ACM/IEEE International Conference on Distributed Smart Cameras (2007)
76. Zhao, J., Cheung, S.C., Nguyen, T.: Optimal camera network configurations for visual tagging. IEEE Journal on Selected Topics of Signal Processing **2**(4) (2008)