A Formal Treatment of Privacy in Video Data

Valerie Fetzer[1], Jörn Müller-Quade[1], and Tobias Nilges[2(✉)]

[1] Karlsruhe Institute of Technology, Karlsruhe, Germany
[2] Aarhus University, Aarhus, Denmark
tobias.nilges@cs.au.dk

Abstract. Video surveillance has become prevalent both in public spaces, e.g. to prevent crimes, and in private areas, e.g. in order to assist the staff in assisted living communities. This leads to privacy concerns regarding the ability of third parties to create profiles and track individuals, possibly across several services.

Usually, techniques such as pixelation and silhouettes are used to anonymize individuals. However, no formal treatment of privacy for video data has been proposed and current anonymization techniques are simply "best practice". To resolve this unsatisfactory state of affairs, we initiate a formal treatment of privacy in video data and propose a game-based notion for privacy in video data that is inspired by cryptographic security games.

We show for an exemplary video privacy scheme that this scheme satisfies our notion with good parameters. In order to evaluate these parameters, we conduct a user study where the users essentially play the role of the adversary in the privacy game. Our approach can be used as a blueprint to evaluate the privacy of other video privacy schemes.

1 Introduction

The advent of video mass surveillance [16,25] in public requires an increased effort to maintain the privacy of individuals, especially since all the collected data can be combined from different sources, allowing tracking of individuals over large time and space intervals [17]. It is obvious that this data can be used to solve crimes after they happened, but it can also be used as a deterrent to *prevent* crimes [31]. Nevertheless, innocent citizens are captured on video, and they should be protected against misuse of the collected video data by anonymization.

Video surveillance is not only used in public, but also as a helpful tool in assisted living homes and retirement homes [1,13]. In this setting, privacy is even more important, because the daily lives of the residents are monitored, including their private rooms. However, providing privacy in such an environment is even more difficult, since an operator of such a surveillance system might learn the habits of the residents and has a lot more side information about the individuals than an operator in public surveillance scenarios.

Over the years it has become apparent that it is notoriously difficult to define privacy. This difficulty spawns from the conflicting goals of releasing anonymous

I. Askoxylakis et al. (Eds.): ESORICS 2016, Part II, LNCS 9879, pp. 406–424, 2016.
DOI: 10.1007/978-3-319-45741-3_21

data: on the one hand the individuals want to remain private, on the other hand the released data should give the receiver of the data some utility. In the context of databases, notions such as k-anonymity [27], l-diversity [19] and t-closeness [23] tried to formally capture anonymity of individuals in a database, but all of these notions have some drawbacks [19,23]. The first formal guarantee for database privacy was presented by Dwork [12] with the notion of differential privacy, albeit only for real-valued data. Even in these scenarios, finding an adequate level of privacy and utility turns out to be non-trivial [20].

For privacy in video data, the state of the art is even worse: we largely rely on techniques such as bounding boxes, blurring, pixelation, edge detection and silhouettes [9,32] that *seem* to be good heuristics to give at least some privacy to individuals. Nevertheless, to the best of our knowledge, formal models for privacy in video data have not been proposed so far. This is at least in part due to the difficulty of formally defining the content of images or video data. Without a way to formally describe the features that are shown in the data, it is also very difficult to formally argue about the features that are hidden by the above mentioned heuristics.

Our contribution. We propose a formal security notion for privacy in video data that is applicable to many scenarios. In particular, this notion is also applicable to scenarios where the operator has a lot of side information about the individuals that are shown in the anonymized videos[1]. The privacy notion is inspired by cryptographic security notions, but evaluated empirically so that we can achieve a good understanding of the privacy of the anonymization algorithm/heuristic.

In more detail, we first define an anonymization scheme with respect to a set of features that it *does not* hide. This formally captures the utility of the anonymization scheme. Given such an anonymization scheme, we define a game-based security definition for privacy in video data, *indistinguishability under individual selection and anonymization* (IND-ISA), that is inspired by semantic security for encryption (IND-CPA) [14]. Overly simplified, we let an adversary choose two individuals from a set of individuals, for which he believes that he can distinguish the anonymized videos. The adversary is given an anonymized video of one of the individuals, and he has to decide which of the individuals is shown in the video. Privacy holds with respect to a parameter ε, i.e. an adversary wins if he decides correctly with probability greater than $\frac{1}{2} + \varepsilon$. The aforementioned notion is impossible to satisfy due to a trivial attack, but we propose a relaxed variant that is actually satisfiable.

We then exemplarily investigate the scenario of an assisted living community, where a possibly malicious operator controls several video cameras, similar to [18]. We focus on the case of individuals that fall to the floor, where the operator has to decide if the individual is actually falling down or just sitting down.

[1] We focus on hiding information about individuals in video data. This does *not* directly translate into anonymity. Instead, it limits the number of features that are available to identify individuals, which in turn can lead to (some form of) anonymity, depending on the scenario.

As a dataset, we use the publicly available [6]. Previously, a silhouette view has been proposed to provide privacy in this scenario [13]. We therefore implement an algorithm that creates a silhouette view and prepare a user study where the users essentially play the role of the adversary in the privacy game.

Our results from the evaluation of the user study show that our anonymization algorithm satisfies our privacy notion in the investigated scenario with $\varepsilon = 0.04$, i.e. the users' answers were essentially as good as random guesses[2]. Our analysis (although based on a very limited dataset) shows that the intuitive idea of using a silhouette view can provide privacy in complex scenarios. But we also show that great care has to be taken when applying the same algorithm to other situations. The same algorithm applied to videos of people sitting down allows to identify the individuals in the videos with significantly higher probability. Our analysis shows that this is due to the fact that the individuals in the anonymized video first take several steps and then sit down, whereas in the individuals in the other videos directly fall to the floor. These steps provides enough information to increase ε to 0.25, which means that the adversary can (with high probability) identify an individual with probability close to 75 %.

Related work. Automatic surveillance systems have seen a rise in the last years, an overview of automated visual surveillance systems can be found in [29]. Senior et al. [26] develop the *PrivacyCam*, where a processing unit in the camera itself takes measures to ensure privacy before sending the video data to the central server. Similarly, Cavallera [5] proposes an architecture of a privacy-preserving video surveillance system, which segments the video stream into privacy-preserving behavioral video data and personal video data. Thorpe et al. [28] also develop a model to split a video stream into an anonymized and a non-anonymized stream. Winkler and Rinner [33] invent the *TrustCam*, a camera with hardware security support, to ensure privacy protection in hardware. The same authors [34] also give an overview about security requirements and privacy protection techniques in the context of visual sensor networks.

Chen et al. [8] present a work that is concerned with the effectiveness of anonymization by only replacing the face of a person with a black box. The authors come to the conclusion that this is by no means satisfactory. A similar result is obtained by Neustaedter et al. [24]. They study a scenario where the privacy of a person working at home should be preserved in a video conference with colleagues at work through blurring the image. The authors find that the blur has to be so intense that the video conference is not worthwhile.

A wide range of techniques for video anonymization have been developed, e.g. [2,3,7,10]. However, there have been very few studies about the quality of anonymization techniques. One of the few studies is the work of Birnstill et al. [4]. They conduct a user study to empirically evaluate different anonymization techniques. The techniques silhouette, pixelization, gray blurring, color blurring and

[2] Please note that we only had a limited dataset available. In real scenarios, the operators of surveillance systems should be able to obtain a larger dataset and more representative results.

edge detection are evaluated with regard to utility and privacy protection. It is empirically shown that the utility of all techniques is very high and most of the techniques achieve acceptable privacy protection. The quality of the anonymization techniques, however, is only indicated through the opinions of the participants of the user study, no formal definition of video anonymization quality has been introduced.

2 Defining Privacy for Video Anonymization Algorithms

Our main goal is to define a security notion for anonymization algorithms on video data. Towards this goal, we first have to define an abstract notion of such an algorithm. Lacking a better term, in the following we will use "anonymize" to describe the process of applying an algorithm to video data that removes certain features. This "anonymization" does not in itself provide *anonymity* for the individuals in the video data.

In general, the motivation behind anonymization of video data is to give privacy to the individual(s) that are shown in the video, while still allowing a third party to use the anonymized video for a specific purpose. Put differently, an anonymization method finds a trade-off between the conflicting goals of providing complete anonymity to the individual on the one hand and providing optimal utility of the video data to a third party on the other. Utility of video data is very hard to formalize, but it is possible to test whether a video allows to learn a certain set of features (e.g. height, physique, etc.).

Thus, for our formalization, we take into account both the privacy and the utility of the data. We do so by explicitly stating a set of features that the anonymization algorithm will *not* hide[3]. A similar approach was recently taken in the context of database privacy by Kifer and Machanavajjhala [21]. However, they explicitly state the data that is supposed to be *hidden*, while we define the features that are not hidden. In most countries, collecting video data must already be justified by a specific purpose, e.g. observing thefts. This purpose implicitly specifies the features that are necessary to accomplish the task. Contrary to database privacy, where the utility of the data is unclear at the time of anonymization, all other information in the video data can be removed by the anonymization process.

Explicitly stating the features of an algorithm that remain unchanged allows to decide whether a specific algorithm is suitable for the given scenario, both with respect to the privacy of the individuals and the utility for the third party. Please note the fundamental difference between e.g. an encryption scheme and an anonymization scheme: the anonymization scheme provides anonymity only in certain scenarios, where the features that remain in the clear do not jeopardize anonymity. In particular, the same algorithm cannot be used blindly in related scenarios without careful consideration (cf. Sect. 4.4).

[3] One might argue that it is very difficult to explicitly state all features that are not hidden, but we show in Sect. 4.4 that a good approximation seems to be sufficient in most cases.

Each anonymization scheme is defined with respect to a set of features that will not be hidden after anonymizing a video.

Definition 1. *A set of features $\mathcal{F}$ is defined as a set of tuples $(f, \mathcal{D})$, where f denotes the name of the feature, and $\mathcal{D}$ the domain of this specific feature. We write $\mathcal{F}(v)$ to denote the specific manifestations of the features in video v.*

Consider for example the features `Age` and `Height`. Then $\mathcal{D}_{\texttt{Age}}$ could be $[0, 100]$ and $\mathcal{D}_{\texttt{Height}}$ could be $[30, 220]$. For a specific video v, $\mathcal{F}(v)$ could be e.g. $\{(\texttt{Age}, 32), (\texttt{Height}, 178)\}$. In the following we define the abstract notion of an anonymization scheme. The anonymization algorithm has to preserve the features in the set $\mathcal{F}_{\mathsf{AS}}$, which implicitly defines the utility of the scheme.

Definition 2. *An* anonymization scheme AS *consists of two algorithms* $(\mathsf{PGen}, \mathsf{Anonymize})$ *and an associated feature set $\mathcal{F}_{\mathsf{AS}}$.*

- $\mathsf{PGen}()$ *outputs a set of public parameters* pp.
- $\mathsf{Anonymize}(\mathsf{pp}, v)$ *outputs an anonymized video $\tilde{v}$ such that $\mathcal{F}_{\mathsf{AS}}(\tilde{v}) = \mathcal{F}_{\mathsf{AS}}(v)$.*

In some cases the anonymization algorithm might add some noise to the preserved features, e.g. the age of an individual might be harder to decide in the anonymized video. This would mean that the equality of $\mathcal{F}_{\mathsf{AS}}(\tilde{v})$ and $\mathcal{F}_{\mathsf{AS}}(v)$ does not hold, instead the features in $\mathcal{F}_{\mathsf{AS}}(\tilde{v})$ are from a smaller domain than $\mathcal{F}_{\mathsf{AS}}(v)$. This influences both utility and privacy of the anonymization scheme. On the one hand, partially hiding some of the features in $\mathcal{F}_{\mathsf{AS}}$ will increase the privacy. On the other hand, the utility of the anonymized video decreases. We do not consider this in our formal definition, since the exact loss of utility/gain of privacy is very hard to quantify, which in turn might make it a lot harder to work with a definition covering this case. One can interpret our definition as a "better safe than sorry" variant, that gives more privacy at the cost of some utility. But verifying the utility of an anonymization scheme in a real world scenario should be fairly simple.

It is obvious that the above definition of an anonymization scheme does not give any privacy. Instead, we want to define a security notion for anonymization schemes analogous to the notion of indistinguishability under chosen plaintext-attack (IND-CPA) in the context of semantically secure encryption [14]. Informally, IND-CPA states that an adversary cannot even learn a single plaintext bit of an encrypted message. For the purpose of defining a formal security notion for privacy in video data, a similar approach would yield a desirable level of privacy: an adversary cannot identify an individual in an anonymized video. However, in contrast to an encrypted message, an anonymized video still has to provide some non-trivial utility (otherwise we could just encrypt it). We will show that this enforces a weaker notion of security, compared to a direct translation of IND-CPA to the scenario of video anonymization.

First, let us briefly recall the IND-CPA security game for an encryption scheme ES. The experiment executes the key generation of the encryption scheme (depending on the security parameter k) and sends the resulting public key pk

to the adversary $\mathcal{A}$. The adversary is now allowed to choose two messages m_0 and m_1, for which he believes that he can distinguish the respective ciphertexts. He sends both messages to the experiment, which in turn randomly selects a bit b and encrypts m_b. This encrypted message $c = \mathsf{Enc}(\mathsf{pk}, m_b)$ is then given to the adversary. The adversary now has to output a bit b' indicating whether the encrypted message is m_0 or m_1. The adversary wins this game if $b = b'$. This game is formalized in Fig. 1.

$$
\begin{aligned}
&\textbf{Experiment } \mathsf{Exp}^{\mathsf{ind-cpa}}_{\mathsf{ES},\mathcal{A}}(k)\\
&\qquad (\mathsf{pk}, \mathsf{sk}) \leftarrow \mathsf{ES.KeyGen}(1^k)\\
&\qquad (m_0, m_1) \leftarrow \mathcal{A}(\mathsf{pk}, 1^k)\\
&\qquad b \leftarrow \{0,1\}\\
&\qquad c \leftarrow \mathsf{ES.Enc}(\mathsf{pk}, m_b)\\
&\qquad b' \leftarrow \mathcal{A}(\mathsf{pk}, c)\\
&\text{return } b = b'
\end{aligned}
$$

Fig. 1. The IND-CPA security game for encryption schemes.

Obviously, an adversary can always guess the random bit. Therefore the security requirement states that an adversary only breaks the security of an encryption scheme if his chance of winning the game is bigger or equal to $\frac{1}{2}+\epsilon(k)$, where $\epsilon(k)$ is a non-negligible function in the security parameter.

2.1 From IND-CPA to IND-ISA

Our goal is to provide a security notion that basically states that an adversary cannot guess which individual is shown in the anonymized video. As indicated above, we cannot directly apply the approach of the IND-CPA notion to privacy in video data, as we will show in the following. Imagine a security game where the experiment first samples some parameters for the anonymization algorithm, then the adversary is allowed to choose two arbitrary videos. He receives an anonymized version of one of the videos, and then has to decide what the underlying video was. This is basically the IND-CPA game with video data. The crucial difference between encryption and anonymization is that we require some utility from the anonymized video. In particular, there has to exist at least one predicate or feature in both videos that has to be recognizable (otherwise we would have no utility). Now an adversary can just select two videos that differ in this predicate, which will result in the adversary winning the game with probability 1.

We thus have to restrict the adversary's power to prevent this trivial attack. This means in particular that the adversary may only indirectly select the video: the video that the experiment anonymizes differs from all the videos that the adversary has at his disposal. While this might seem to be a strong restriction, we

argue that this is the case in all scenarios where anonymization is required. The adversary might have some knowledge–not just in the form of video data–about the individual(s) that appear in the anonymized video, but he should obviously not have access to the exact same video without anonymization. The adversary can still choose videos based on individuals for which he believes that he can distinguish the anonymizations, e.g. due to the height, clothing or physique of the individuals.

Additionally, we have to restrict the choice of videos with respect to the feature set $\mathcal{F}_{\mathsf{AS}}$ of the anonymization algorithm. To obtain a meaningful notion, we require that the experiment picks only videos that are indistinguishable with respect to these features; the videos can be completely arbitrary with respect to all other features. An anonymization algorithm satisfying our notion thus implies that the adversary cannot distinguish the anonymized videos depending on all the other features, i.e. these features are effectively removed by the anonymization algorithm.

We model the requirement of videos that are indistinguishable with respect to $\mathcal{F}_{\mathsf{AS}}$ by defining an extraction algorithm Ext. This algorithm extracts the features included in $\mathcal{F}_{\mathsf{AS}}$ from a given video.

Definition 3. *An* extraction algorithm Ext *for a feature set* $\mathcal{F}_{\mathsf{AS}}$ *gets as input a video* v *and returns the manifestations* $\mathcal{F}_{\mathsf{AS}}(v)$.

We are now ready to propose the notion of *indistinguishability under individual selection and anonymization*, or IND-ISA. The adversary selects two individuals from a set of individuals, the experiment selects two videos containing the individuals and uses the extractor Ext to verify that the videos are indeed indistinguishable with respect to $\mathcal{F}_{\mathsf{AS}}$. It then anonymizes one video at random and sends it to the adversary, who has to guess which individual is shown in the video. This notion captures the goal of hiding the identity of an individual, even if the adversary is allowed to select the individuals that the experiment has to choose from. The game is depicted in Fig. 2. Let $\mathcal{A}$ denote the adversary, I denote the set of individuals and $\mathcal{V}$ the set of videos from which the experiment can choose to create an anonymized challenge. We write $\mathcal{V}_{|i}$ to denote the subset of $\mathcal{V}$ showing individual $i \in I$.

Remark 1. In some scenarios, it might be possible to allow the adversary to actually choose two videos directly and only require that the features that are not hidden have to be identical. While this yields a stronger notion of privacy, we believe that such a notion is hard to achieve in general. Our notion on the other hand provides a (weaker) guarantee, but is applicable in most realistic scenarios.

Remark 2. Obviously, IND-ISA only states that given two videos with the same manifestations of features regarding $\mathcal{F}_{\mathsf{AS}}$ an adversary cannot distinguish the anonymized videos. This does not necessarily guarantee anonymity: in a scenario where one individual has a specific set of feature manifestations that no other individual shares, the anonymized video might still leak the identity. Thus, in order to achieve anonymity of individuals in a specific scenario, it has to be

Experiment $\mathsf{Exp}^{\mathsf{ind-isa}}_{\mathsf{AS},\mathcal{A}}$

$(\mathsf{pp}) \leftarrow \mathsf{AS.PGen}()$

$(i_0, i_1) \leftarrow \mathcal{A}(\mathsf{pp})$

return 0 if $i_0 = i_1$ or $i_0, i_1 \notin I$

$v_0 \leftarrow \mathcal{V}_{|i_0}$

$v_1 \leftarrow \mathcal{V}_{|i_1}$

return 0 if $\mathsf{Ext}(\mathcal{F}_{\mathsf{AS}}, v_0) \neq \mathsf{Ext}(\mathcal{F}_{\mathsf{AS}}, v_1)$

$b \leftarrow \{0, 1\}$

$\tilde{v} \leftarrow \mathsf{AS.Anonymize}(\mathsf{pp}, v_b)$

$b' \leftarrow \mathcal{A}(\mathsf{pp}, \tilde{v})$

return $b = b'$

Fig. 2. The IND-ISA privacy game for video anonymization schemes.

ensured that at least two individuals with the same feature manifestations actually exist.

In comparison to standard cryptographic notions, it is at the very least unclear how to achieve asymptotic security. Instead, we measure the success probability of an adversary with a statistical parameter ε that is supposed to indicate his success probability in comparison to simply guessing the result. We defer a detailed discussion on the parameter ε to Sect. 3, where we propose a method to determine ε. Given the definition of IND-ISA, we are able to define the privacy of an anonymization scheme.

Definition 4. *We say an anonymization scheme* AS *is* $\varepsilon-$*IND-ISA-secure, if any adversary wins* $\mathsf{Exp}^{\mathsf{ind-isa}}_{\mathsf{AS},\mathcal{A}}$ *with probability at most* $\frac{1}{2} + \varepsilon$ *averaged over all videos in* $\mathcal{V}$*.*

At first sight, the above definition of privacy for anonymization schemes might seem very weak, because we only require an *average* distinguishing advantage. In particular, there might exist an adversary that identifies an individual in a certain video with very high probability. This lies in stark contrast to the classical cryptographic security notions, but it still gives us a meaningful measure of the quality of an anonymization scheme.

3 On Obtaining IND-ISA-Secure Anonymization Schemes

Before we describe our anonymization scheme and discuss the results, let us first elaborate on how one can actually obtain IND-ISA security. The main problem that we face is that we have no underlying cryptographic assumptions or statistical data on the image to work with, on which we could base the security for an algorithm.

Our idea is to take an empirical approach to verify the security of a specific algorithm. Instead of assuming hardness of some underlying assumption, we actually implement the IND-ISA game with videos and an anonymization scheme. In contrast to cryptographic games, where we make no assumptions about the adversary apart from possibly polynomial efficiency, we now use a "constructive" approach.

The approach that we use for our example implementation is as follows: We create a user study, in which users essentially take the role of the adversary (a human adversary is the most likely case in real world scenarios). Using a statistical analysis, it is possible to measure the indistinguishability of anonymized videos in a meaningful way. Given enough participants for such a study, we can make a fairly good assumption on the IND-ISA security of the anonymization algorithm. We propose the following method to determine the value ε. The user study is basically a Bernoulli process, i.e. the game that the user plays is independent of the games that other users played. This means that the results should be distributed according to a binomial distribution. We can thus sum up all the answers from the study and compute the Clopper-Pearson interval [11] with confidence e.g. 95 % and $p = \frac{1}{2}$. This results in two values $[a, b]$, and we set $\varepsilon = \max\{b - \frac{1}{2}, 0\}$. Taking the value b of the interval corresponds to a "worst-case" choice for the parameter ε.

Another approach based on the same idea is to use learning algorithms to play the IND-ISA game, e.g. PAC learning [30]. This would yield several advantages over the approach with user studies: first of all it is much cheaper and less time consuming than a user study while additionally shedding light on the question whether an adversary can learn to distinguish the anonymized videos over time. We leave this research direction as an open question for future work.

In the following section, we give two examples of such studies with somewhat surprising results.

4 An IND-ISA-Private Anonymization Scheme

In the following we will present an exemplary instantiation of an anonymization scheme applied to two real world scenarios to illustrate our approach. In order to do so, we first have to define an anonymization scheme according to Definition 2 and a scenario to apply it to.

4.1 Scenario

In recent years, supporting the staff of an assisted living community with privacy preserving video surveillance has been the focus of research, both with respect to feature recognition and privacy [1,13,18,22,35]. In a little more detail, the idea is to enhance the apartments of an assisted living community with video cameras that help the staff detecting accidents and emergencies. One line of research focuses on fully automated systems that incorporate e.g. fall detection algorithms [13]. However, the detection rate is not perfect and in the case of

an emergency a falling resident *must* be detected, which inevitably leads to possibly many false positive alarms. Another aspect is the problem that in some countries fully automated systems are not allowed to make decisions without human verification. Thus, in a realistic scenario an operator has to access the video data and can evaluate whether an emergency arises. This operator can observe the individuals in their private rooms, therefore anonymization methods have to be applied to ensure privacy of the residents.

We study two situations that are closely related: an individual falling to the floor and an individual sitting down. These situations are difficult to distinguish by an algorithm and therefore sometimes require human verification. We assume that the operator knows the individuals that are shown by the surveillance system, but he only sees the anonymized video of the presumed fall. The video data recorded before the incident and afterwards cannot be accessed by the operator. Further, the operator is supposed to verify that the individual fell or recognize an error by the algorithm. This implies that the anonymization scheme must provide enough utility to make this distinction.

4.2 Anonymization Scheme

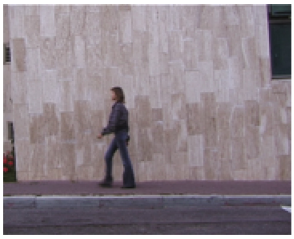
In the literature, the anonymization technique of choice for this scenario is a silhouette view. On the one hand, it still allows to discern the movements of an individual, on the other hand most features like colors and environment are removed. However, there is no formal treatment regarding the validity of this approach. We thus implemented an anonymization algorithm that realizes a blurred silhouette view of a video. The blurring is added to remove the exact outline of the individual, which might give away too much information. The implementation is based on the OpenCV library and written in `C++`.

The generation of the public parameters for our anonymization scheme $\mathsf{AS}_{\mathsf{sil}}$ is implicit, they are included in the anonymization algorithm $\mathsf{Anonymize}$. It proceeds in three steps:

1. For each frame: calculate the silhouette view by means of background subtraction
2. For each frame: blur the image with median filter and Gaussian blur
3. Normalize the whole video by zooming to the relevant section and then scaling the video to a fixed width

The public parameters $\mathsf{pp}_{\mathsf{sil}}$ are the parameters used for each step: in the first step any parameters that are needed to produce the silhouette view, for the second step the parameters for the median filter and Gaussian blur and in the last step the width the video is scaled to. Figure 3 shows the intermediate results of the algorithm when applied to an example video.

In accordance with our definition of anonymization schemes, our algorithm must be defined with respect to a feature set $\mathcal{F}_{\mathsf{AS}_{\mathsf{sil}}}$. Finding a formal specification for the feature set is highly nontrivial, because some features are a superset of other features. These features are therefore also distinguishable after



(a) The raw image

(b) After silhouette extraction

(c) After blurring

(d) After normalizing

Fig. 3. An example for the use of $\mathsf{AS}_{\mathsf{sil}}$. In (a) the original image can be seen, picture (b) and (c) are intermediate steps and the finished anonymization is shown in (d). The video data is taken from Gorelick et al. [15].

anonymization. Thus, we define a rather coarse feature set that should cover all distinguishable features, but might include some features that are actually hidden by the scheme $\mathsf{AS}_{\mathsf{sil}}$. In the following, let

$$\mathcal{F}_{\mathsf{AS}_{\mathsf{sil}}} = \{\texttt{physique}, \texttt{movements}, \texttt{accessories}\}.$$

Note that it is e.g. possible to derive knowledge about the age from the sequence of movements and the physique of an individual. By accessories we mean objects like an umbrella, a bag or a walking aid.

4.3 Dataset

We use a dataset due to Charfi et al. [6]. It contains 191 videos of individuals walking, falling down and sitting down. The resolution of the videos is 320×240 px and the frame rate is 25 frames/second. Only 60 of these videos were suitable for our study. The amount was further limited because the number of videos of each individual fluctuates. There are only four videos of one individual falling down, and 3 videos of another individual sitting down. The videos of falling and sitting individuals were used as the set $\mathcal{V}$ in the IND-ISA game. Videos of individuals walking through the room were given the users to get some background information on the individuals. Therefore the four individuals of the dataset add up to the set I in the IND-ISA game.

The camera is placed in a corner of the room, comparable to the installation in a real assisted living community. The room contains a table, a chair and a sofa. In some videos, a carpet or blanket is put on the floor. Based on this dataset we can test several aspects.

– Is it possible to discern whether an individual is falling down or sitting down? This is one situation that might occur if an automated system misclassifies a situation and describes the utility of the data.
– Is it possible to identify an individual if one knows only the video of the person falling down? This reflects our definition of privacy in video data.

Our user study covers both of these aspects. We cut the videos that we wanted to anonymize such that they only show the respective action, i.e. an individual falling down or sitting down[4].

4.4 User Study

We prepared a user study that evaluated two aspects. On the one hand, users had to decide whether an individual was falling down or sitting down in an anonymized video. On the other hand, we let each user play the IND-ISA privacy game for the anonymization scheme $\mathsf{AS}_{\mathsf{sil}}$ based on the aforementioned video dataset. The study was conducted similar to our proposal from Sect. 3. Each user was shown four videos of different individuals that walk through a room (cf. Fig. 4), i.e. these individuals represent the set I of the IND-ISA game. These individuals were (supposedly) chosen in a way that they were indistinguishable with respect to $\mathcal{F}_{\mathsf{AS}_{\mathsf{sil}}}$ (because we did not have a feature extraction that could have been used). The user had to select two of the four videos, i.e. individuals, for which he believed he could distinguish an anonymized version.

After the selection the user was shown one anonymized video of one of the selected persons falling down or sitting down, respectively. Then the user had to decide which person was anonymized in the video. Answer options were "Person X", "Person Y" and "I don't know", where X and Y were replaced with the letters corresponding to the chosen persons. We added the "I don't know"-option to get a better understanding of how the users felt when seeing the anonymized video. For the analysis, we assumed that in the case of the "I don't know"-option an IND-ISA-adversary would just guess an option with probability $\frac{1}{2}$. Additionally, the user had the option to describe the feature(s) that helped him identify the individual.

We let each user play several games (four, respectively two) with new anonymized videos to increase the number of samples that we could use, which in turn breaks the independence of the sequential games. Thus, if we use the approach from Sect. 3 to compute the value ε, the result is less accurate, because we introduce learning effects. Due to the structure and low sample size of our user study, however, it is not possible to quantify these learning effects in a

[4] Our discussion in Sect. 4.5 elaborates on the problems that arise when cutting the videos.

Fig. 4. Parts of the videos that provide the user of the study with some knowledge about the individuals which he has to distinguish. Each video shows the corresponding individual walking around.

meaningful way, so we stick to the same approach to compute ε in this example implementation. Usually, one would have to calculate the multidimensional confidence interval with respect to several identically distributed games.

All in all, we obtained results from 248 users, but only 103 users answered the complete study. The results show that the anonymization method is suitable to discern between individuals falling down and individuals sitting down. The detection rate of individuals sitting down in the anonymized videos was 100 %, and the detection rate of individuals falling down was between 97 % and 99 %. Our results concerning the privacy of the individuals are summarized in the following two paragraphs.

Evaluation of the Fall Detection Study. Due to the small amount of available videos, we let each user play four games with individuals falling down. Figure 5 shows images from the anonymized videos. The set of videos that was used for the anonymization is disjoint from the set of videos that the users saw in order to select their "challenge" identity in the IND-ISA game.

An analysis of the data shows that, generally, the users had difficulties in identifying the correct individual in the anonymized video. Averaged over all four games, 60 % of the users selected the "I don't know"-option, with values ranging from 52 to 69 %. Of all the users that identified an individual, averaged over all four games, the answers were close to uniformly distributed (19 % selected the

Fig. 5. Images of some of the anonymized videos used in the fall detection study. Each anonymized video shows the corresponding individual falling down.

Individuals	Game 1			Game 2			Game 3			Game 4		
	[1]	[2]	IDK	[1]	[2]	IDK	[1]	[2]	IDK	[1]	[2]	IDK
A,B	3	**3**	4	**4**	12	5	**3**	8	9	2	**3**	12
A,C	9	**10**	41	**6**	20	22	**10**	8	20	3	**4**	31
A,D	0	**7**	2	**2**	1	6	**4**	5	6	4	**3**	6
B,C	3	**6**	28	**5**	6	23	**5**	2	23	3	**2**	25
B,D	0	**1**	1	**1**	1	3	**0**	1	4	2	**1**	5
C,D	4	**6**	2	**1**	1	8	**6**	1	11	7	**3**	5
Total	19	**33**	78	**19**	41	67	**28**	25	73	21	**16**	84

Fig. 6. This table shows the number of answers for each of the four games. Each game is split into 6 separate games due to the choices of individuals by the users. [1] and [2] specify the first individual and the second individual from the first column, respectively. Correct answers are written in bold font.

correct answer, 21 % the wrong answer, 60 % did not recognize the individual). A more detailed breakdown is shown in Fig. 6.

While the first game might be interpreted as an indication that it is possible to identify the individuals, we have to consider that the sample size is very small. When we look at the actual detection rates for each possible pair of choices by the users, we observe that except for one pair of individuals in the first game (Individuals A and D), the answers are nearly uniformly distributed. The users that selected this constellation of individuals described that the gait and speed of movements of the individuals convinced them that the anonymized video showed the correct individual. Upon inspection of the videos, we noticed that the anonymized video shows the individual taking two steps before falling to the floor.

There are several possible interpretations. On the one hand, the sample size is far too small to rule out a statistical fluke (9 users, 7 correct answers). Additionally, the users' answers with regard to the identifiable feature have to be taken with a grain of salt, because many users claimed to have identified individuals according to e.g. gait, even in cases where they chose the wrong individual. On the other hand, in combination with the results from the videos of individuals sitting down (cf. Sect. 4.4), we believe that the two steps shown in the video are indeed enough information to correctly identify an individual.

We now want to fix a value for ε. As mentioned above, we will ignore learning effects and just calculate ε as described in Sect. 3. We split the "I don't know"-answers evenly between correct and incorrect guesses, which results in a success probability of 49.01 % based on our study, i.e. the wrong answer is given with 50.99 %. By computing the Clopper-Pearson confidence interval [11] for the success probability we get the interval [44.56 % , 53.47 %]. Thus, $\varepsilon = 0.5347 - 0.5 = 0.0347$, which rounds up to $\varepsilon = 0.04$.

Evaluation of the Sitting Detection Study. In comparison to the fall detection study, we had less videos available so that we let each user play only

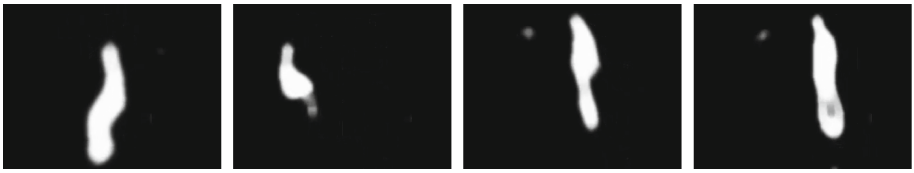


Fig. 7. Parts of some of the anonymized videos used in the sitting detection study. Each video shows the corresponding individual sitting down in anonymized form.

two games. Apart from this and the fact that the individuals sit down in the anonymized videos, the setup is identical to the fall detection study. Examples for the anonymized videos used in the sitting detection study are shown in Fig. 7.

Our results are shown in Fig. 8. On average, 55 % of the users identified the correct individual, but only 18 % identified the wrong individual, leaving 27 % of undecided users. Looking closer into the data, it seems as if the first game was more or less uniformly distributed averaged over all 6 possible choices of the users, but in the second game 75 % of the users selected the correct individual. Even worse, a closer look at the individual games shows that the users identified some individuals with probability (close to) 1, even for certain choices of individuals in game 1.

Most users claimed to have identified the individuals due to the walking style in the videos. In comparison to the videos that show individuals falling down, nearly all anonymized videos include two steps taken by the individual before sitting down. As in the previous study, this seems to give the users a high probability in identifying the anonymized individual.

Obviously, the value ε is much worse than in the previous case. Again, we distribute the "I don't know"-answers evenly between the right and the wrong answers. The success probability of an IND-ISA-adversary is therefore 69 % and

Individuals	Game 1			Game 2		
	[1]	[2]	IDK	[1]	[2]	IDK
A,B	**4**	6	11	2	**12**	3
A,C	**4**	20	21	2	**32**	8
A,D	**2**	1	3	0	**5**	0
B,C	**17**	7	5	0	**27**	7
B,D	**1**	0	3	4	**2**	1
C,D	**12**	0	2	0	**11**	2
Total	**40**	34	46	8	**89**	21

Fig. 8. This table shows the number of answers for both games. Each game is split into 6 separate games due to the choices of individuals by the users. [1] and [2] specify the first individual and the second individual from the first column, respectively. Correct answers are written in bold font.

the error probability 31 %. The calculation of the Clopper-Pearson interval yields [62.02 %, 74.22 %]. Thus, $\varepsilon = 0.7422 - 0.5 = 0.2422$, which rounds up to $\varepsilon = 0.25$.

4.5 Discussion

We believe that the main reason that users could identify an individual in an anonymized video was that the anonymized videos included one or two steps of the individual. We cut the videos in this way to help the users identify the action, but our results concerning the privacy show that this allows the identification of anonymized individuals with high probability.

Intuitively, this already follows from the definition of the privacy game IND-ISA and our anonymization scheme: the anonymization scheme does not hide sequences of movements, therefore the IND-ISA game would require us to select videos that are indistinguishable with respect to these movements. From an adversarial point of view, seeing an individual walk will not give him much information about an individual falling down (in our scenario). However, if the anonymized video includes, apart from the falling or sitting individual, some steps of the individual, the adversary can use his knowledge to identify this individual. When we selected videos that include both walking and falling down, we did not (manually) check whether the individuals were indistinguishable with respect to the walking movements. Our results clearly show that the individuals are not indistinguishable in that regard. We believe that removing the part of the videos that shows the steps of the individuals would greatly reduce the users' ability to identify the individuals and yield a smaller value for ε. In turn, the measured utility of the anonymization might drop a bit.

This highlights a very important aspect of anonymizing video data: even if an anonymization method is very good in one scenario, it is not possible to incorporate it into another scenario without considering all aspects of the anonymization process.

To finish our evaluation, we want to discuss a value ε_{t} that describes the anonymity of individuals in the above described scenario, i.e. an operator is shown anonymized videos of alleged falls of individuals. The overall identification probability in this scenario is dependent on the accuracy of the fall detection system. If we assume the detection algorithms of Charfi et al. [6], a fall is correctly detected with probability 99.6 %. Thus, despite the poor performance of the anonymization due to videos showing the walking movements, the privacy of the individuals in the scenario is very close to $\varepsilon_{\mathsf{t}} = 0.04$.

5 Conclusion and Future Work

Currently, privacy in video data is only argued on a very informal level, and the applied anonymization algorithms are more or less chosen because they are "best practice". As our second example showed, this can be a dangerous way to approach privacy in complex systems. We started a formal treatment of the problem of privacy in video data and laid a formal foundation for future

research in this area. We believe that an interdisciplinary effort is required to find widely applicable anonymization algorithms, together with good formal guarantees. This is due to the fact that for anonymization methods, we need knowledge from security/cryptography, feature recognition and machine learning. Additionally, a large dataset for the evaluation of the anonymization schemes seems essential, because it is not possible to create large amounts of realistic videos in software.

Using the approach that we presented, it is possible to define a set of candidates for IND-ISA secure anonymization algorithms that cover a wide range of applications. Being able to abstract from the anonymization method allows the design of *provably private* video surveillance systems. The system itself can be proven private, while the anonymization algorithm is then chosen according to the scenario in which the surveillance system is used. Thus, the security analysis is simplified and modularized.

References

1. Abowd, G.D., Bobick, A.F., Essa, I.A., Mynatt, E.D., Rogers, W.A.: The aware home: a living laboratory for technologies for successful aging. In: Proceedings of AAAI Workshop and Automation as a Care Giver, pp. 1–7 (2002)
2. Bamba, B., Liu, L.: Privacygrid: supporting anonymous location queries in mobile environments. Technical report, Defense Technical Information Center (2007)
3. Berger, A.M.: Privacy mode for acquisition cameras and camcorders. US Patent 6,067,399 (2000)
4. Birnstill, P., Ren, D., Beyerer, J.: A user study on anonymization techniques for smart video surveillance. In: 2015 12th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), pp. 1–6. IEEE (2015)
5. Cavallaro, A.: Adding privacy constraints to video-based applications. In: European Workshop on the Integration of Knowledge, Semantics and Digital Media Technology (2004)
6. Charfi, I., Miteran, J., Dubois, J., Atri, M., Tourki, R.: Definition and performance evaluation of a robust SVM based fall detection solution. In: Eighth International Conference on Signal Image Technology and Internet Based Systems (SITIS), pp. 218–224. IEEE (2012)
7. Chen, D., Chang, Y., Yan, R., Yang, J.: Tools for protecting the privacy of specific individuals in video. EURASIP J. Adv. Sig. Process. **2007**(1), 075427 (2007)
8. Chen, D., Chang, Y., Yan, R., Yang, J.: Protecting personal identification in video. In: Senior, A. (ed.) Protecting Privacy in Video Surveillance, pp. 115–128. Springer, London (2009)
9. Chinomi, K., Nitta, N., Ito, Y., Babaguchi, N.: Prisurv: privacy protected video surveillance system using adaptive visual abstraction. In: Satoh, S., Nack, F., Etoh, M. (eds.) MMM 2008. LNCS, vol. 4903, pp. 144–154. Springer, Heidelberg (2008)
10. Cichowski, J., Czyzewski, A.: Reversible video stream anonymization for video surveillance systems based on pixels relocation and watermarking. In: IEEE International Conference on Computer Vision Workshops (ICCV Workshops), pp. 1971–1977 (2011)
11. Clopper, C.J., Pearson, E.S.: The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika **26**(4), 404–413 (1934)

12. Dwork, C.: Differential privacy. In: Bugliesi, M., Preneel, B., Sassone, V., Wegener, I. (eds.) ICALP 2006. LNCS, vol. 4052, pp. 1–12. Springer, Heidelberg (2006)
13. Fleck, S., Straßer, W.: Smart camera based monitoring system and its application to assisted living. Proc. IEEE **96**(10), 1698–1714 (2008)
14. Goldwasser, S., Micali, S.: Probabilistic encryption. J. Comput. Syst. Sci. **28**(2), 270–299 (1984)
15. Gorelick, L., Blank, M., Shechtman, E., Irani, M., Basri, R.: Actions as space-time shapes. Trans. Pattern Anal. Mach. Intell. **29**(12), 2247–2253 (2007)
16. Haering, N., Venetianer, P.L., Lipton, A.: The evolution of video surveillance: an overview. Mach. Vis. Appl. **19**(5–6), 279–290 (2008)
17. Hampapur, A., Brown, L., Connell, J., Ekin, A., Haas, N., Lu, M., Pankanti, S.: Smart video surveillance: exploring the concept of multiscale spatiotemporal tracking. IEEE Sig. Process. Mag. **22**(2), 38–51 (2005)
18. Huber, M., Müller-Quade, J., Nilges, T., Thal, C.: A provably privacy preserving video surveillance architecture for an assisted living community. In: GI-Jahrestagung, pp. 563–574 (2014)
19. Kifer, D., Gehrke, J.: L-diversity: privacy beyond k-anonymity. In: ICDE, p. 24 (2006)
20. Kifer, D., Machanavajjhala, A.: No free lunch in data privacy. In: Proceedings of the 2011 ACM SIGMOD International Conference on Management of Data, SIGMOD 2011, pp. 193–204. ACM, New York (2011)
21. Kifer, D., Machanavajjhala, A.: Pufferfish: a framework for mathematical privacy definitions. ACM Trans. Database Syst. **39**(1), 3:1–3:36 (2014)
22. Kleinberger, T., Becker, M., Ras, E., Holzinger, A., Müller, P.: Ambient intelligence in assisted living: enable elderly people to handle future interfaces. In: Stephanidis, C. (ed.) UAHCI 2007 (Part II). LNCS, vol. 4555, pp. 103–112. Springer, Heidelberg (2007)
23. Li, N., Li, T., Venkatasubramanian, S.: t-Closeness: privacy beyond k-anonymity and l-diversity. In: ICDE, pp. 106–115 (2007)
24. Neustaedter, C., Greenberg, S., Boyle, M.: Blur filtration fails to preserve privacy for home-based video conferencing. ACM Trans. Comput. Hum. Interact. (TOCHI) **13**(1), 1–36 (2006)
25. Norris, C., McCahill, M., Wood, D.: The growth of CCTV: a global perspective on the international diffusion of video surveillance in publicly accessible space. Surveill. Soc. **2**(2/3), 110–135 (2002)
26. Senior, A., Pankanti, S., Hampapur, A., Brown, L., Tian, Y.L., Ekin, A., Connell, J., Shu, C.F., Lu, M.: Enabling video privacy through computer vision. IEEE Secur. Priv. **3**(3), 50–57 (2005)
27. Sweeney, L.: k-anonymity: a model for protecting privacy. Int. J. Uncertainty Fuzziness Knowl. Based Syst. **10**(5), 557–570 (2002)
28. Thorpe, C., Li, F., Li, Z., Yu, Z., Saunders, D., Yu, J.: A coprime blur scheme for data security in video surveillance. IEEE Trans. Pattern Anal. Mach. Intell. **35**(12), 3066–3072 (2013)
29. Valera, M., Velastin, S.A.: Intelligent distributed surveillance systems: a review. In: IEEE Proceedings - Vision, Image and Signal Processing, vol. 152, pp. 192–204. IET (2005)
30. Valiant, L.G.: A theory of the learnable. Commun. ACM **27**(11), 1134–1142 (1984)
31. Welsh, B.C., Farrington, D.P.: Evidence-based crime prevention: the effectiveness of CCTV. Crime Prev. Community Saf. **6**(2), 21–33 (2004)

32. Wickramasuriya, J., Datt, M., Mehrotra, S., Venkatasubramanian, N.: Privacy protecting data collection in media spaces. In: Proceedings of the 12th Annual ACM International Conference on Multimedia, pp. 48–55. ACM (2004)
33. Winkler, T., Rinner, B.: Privacy and security in video surveillance. In: Atrey, P.K., Kankanhalli, M.S., Cavallaro, A. (eds.) Intelligent Multimedia Surveillance, pp. 37–66. Springer, Heidelberg (2013)
34. Winkler, T., Rinner, B.: Security and privacy protection in visual sensor networks: a survey. ACM Comput. Surv. (CSUR) **47**(1), 2 (2014)
35. Wood, A.D., Stankovic, J.A., Virone, G., Selavo, L., He, Z., Cao, Q., Doan, T., Wu, Y., Fang, L., Stoleru, R.: Context-aware wireless sensor networks for assisted living and residential monitoring. IEEE Netw. **22**(4), 26–33 (2008)