De-identification for Privacy Protection in Multimedia Content: A Survey

Slobodan Ribaric[1], Aladdin Ariyaeeinia[2], Nikola Pavesic[3]

[1]*University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia*

[2]*University of Hertfordshire, Hatfield, UK*

[3]*University of Ljubljana, Faculty of Electrical Engineering, Ljubljana, Slovenia*

ABSTRACT

Privacy is one of the most important social and political issues in our information society, characterized by a growing range of enabling and supporting technologies and services. Amongst these are communications, multimedia, biometrics, big data, cloud computing, data mining, internet, social networks, and audio-video surveillance. Each of these can potentially provide the means for privacy intrusion. De-identification is one of the main approaches to privacy protection in multimedia contents (text, still images, audio and video sequences and their combinations). It is a process for concealing or removing personal identifiers, or replacing them by surrogate personal identifiers in personal information in order to prevent the disclosure and use of data for purposes unrelated to the purpose for which the information was originally obtained. Based on the proposed taxonomy inspired by the Safe Harbour approach, the personal identifiers, i.e., the personal identifiable information, are classified as non-biometric, physiological and behavioural biometric, and soft biometric identifiers. In order to protect the privacy of an individual, all of the above identifiers will have to be de-identified in multimedia content. This paper presents a review of the concepts of privacy and the linkage among privacy, privacy protection, and the methods and technologies designed specifically for privacy protection in multimedia contents. The study provides an overview of de-identification approaches for non-biometric identifiers (text, hairstyle, dressing style, license plates), as well as for the physiological (face, fingerprint, iris, ear), behavioural (voice, gait, gesture) and soft-biometric (body silhouette, gender, age, race, tattoo) identifiers in multimedia documents.

Keywords: Privacy, Multimedia, De-identification, Biometric identifiers, Soft biometric identifiers, Non-biometric identifiers

1. Introduction

Recent advances in audio-recording devices, cameras, web technology and signal processing have greatly facilitated the efficacy of audio and video surveillance, primarily for the benefit of security and law enforcement. This technology is now widely exploited in a variety of scenarios to capture audio-

video recordings of people in public environments, either for immediate inspection (e.g., abnormal behaviour recognition, identification and tracking of people in real time) or for storage, and subsequent data analysis and sharing. Capabilities in the field are further supported through continued progress in a number of relevant areas, including smart, multi-camera networks [1], wireless networks of multispectral image sensors, drones equipped with camera, audio-sensor arrays, distributed intelligence and awareness, and distributed processing power [2].

Whilst it is clear that there are justifiable reasons for sharing multimedia data acquired in such ways (e.g. for law enforcement, forensics, bioterrorism surveillance, disaster prediction), there is also a strong need to protect the privacy of innocent individuals who are inevitably "captured" in the recordings. In order to recognise the growing scale of this surveillance and its effects on privacy, it is worth noting that, for instance, there are more than forty-eight hundred government surveillance cameras in Washington, D.C. [3] and over 4 million closed-circuit television (CCTV) cameras deployed in the United Kingdom. The average citizen in London is caught on CCTV cameras about 300 times a day [4]. The problem associated with this is further exacerbated by lack of compliance with the relevant data-protection legislation. According to a study in [5], this is the case for over 80% of the CCTV systems deployed in London's business space.

An additional and growing feature of the privacy problem in today's networked society is the advent of technologies such as "Google Street View" and "EveryScape", social networks, biometrics, multimedia, big data, and data mining. These provide an additional framework for the invasion of an individuals' privacy. In [6], J. Angwin analyzed relations among privacy, security and freedom in a world of relentless electronic surveillance - from Google to NSA. J. Angwin has concluded that we are living in the world of indiscriminate tracking where institutions are stockpiling data about individuals at an unprecedented pace. This indiscriminate tracking is powered by "the technology we love so much" - powerful desktops, laptops, tablets, smart-phones and web services.

In view of the above issues, considerable research has now been directed towards approaches for the preservation of privacy and personal information. The main facet of efforts in this area, which is also the focus of this paper, is concerned with the development of methods for the de-identification of

individuals captured in multimedia content (text, audio, still images, animation, video, and their combination). In order to provide an appropriate basis for the analysis presented here, the next section details the definition of privacy, and its social and legal aspects as well as its significance in today's society. The subsequent sections then present a survey of de-identification in multimedia content. The scope of the study is broad and covers methods for dealing with non-biometric, biometric physiological and behavioural identifiers, and soft biometric identifiers.

2. Privacy

There is no single definition of the term "privacy". The meaning of privacy depends on legal, political, societal, cultural and socio-technological contexts [7]. From the legal point of view, the first definition of privacy was given by Louis D. Brandeis and Samuel D. Warren more than 120 years ago [8]. They defined privacy as "the right to be let alone", with respect to the acquisition and dissemination of information concerning the person, particularly through unauthorized publication, photography or other media. Also, according to Brandeis and Warren, the person should be protected from investigation and seizures that invade a sphere of individual solitude deemed reasonable by society. Additionally, the person has "the right to be let alone" with respect to fundamental decisions concerning his or her intimate relationships or aspects of life.

Alan F. Westin defines privacy as the claim of an individual to determine what information about himself or herself should be known to others [9]. Based on the various usages of the word "privacy", there are many different conceptions of privacy and they can be classified into six general types [10]: (i) the right to be let alone; (ii) limited access to the self – the ability to protect oneself from unwanted access by others; (iii) secrecy – the concealment of certain matters from others; (iv) control over personal information; (v) personhood – the protection of one's personality, individuality and dignity; (vi) intimacy – control over, or limited access to, one's intimate aspects of life.

Depending on the social contexts and/or real life situations, privacy, in general, can be divided into a number of separate, but related, concepts [11]: (i) informational privacy – the right of the individual to limit access to personal information which could be used in any way to identify an individual; (ii) intentional privacy – the right of the individual to prevent or forbid further communication of observed

events or exposed features (e.g., publishing photos or video footage); (iii) decisional privacy – the right of the individual to make decisions regarding his life without any undue interference; (iv) spatial privacy – the right of the individual to have his own personal spaces which cannot be violated without his explicit consent. If we include some physical and socio-technological contexts in the above classification, we can talk about: (i) information privacy, which involves the establishment of rules governing the collection and handling of personal data such as medical and tax records and credit information; (ii) the privacy of communications, which covers the security and privacy of mail, telephone, e-mail and other forms of communication; (iii) bodily privacy, which concerns the protection of people's physical selves against invasive procedures such as genetic tests, drug testing and cavity searches; (iv) territorial privacy, which concerns the setting of limits on intrusion into domestic and other environments, such as the workplace or public space. This includes searches, video surveillance and ID checks.

An in-depth and comprehensive insight into the theory of privacy, existing attempts to conceptualize privacy and different definitions of privacy from the standpoint of jurists, philosophers and sociologists are given in the book [10].

Let us illustrate the need for privacy and personal data protection with three examples of privacy violation. Case 1 describes a situation in which privacy is violated due to the inadequate protection of the face as a biometric identifier. Case 2 describes a situation in which privacy is violated and abused due to the low level of protection of stored personal documents with biometric identifiers and other personal identifiable information. Case 3 deals with the potential abuse of a facial recognition system used in public places.

Case 1: A person attempted suicide by slitting his wrists with a knife in a street. A CCTV surveillance camera was recording him, and the person monitoring the camera notified the police. The person was saved and transported to hospital. Some months later, the Council issued two photographs of the person taken from the CCTV footage for publication in an article about the preventative benefits of CCTV. The person's face was not specifically masked and he could be identified by people who knew

him. Extracts from the CCTV footage were also shown on regional television in which the person's face had been masked at the Council's request.

Epilogue: The person sought judicial review of the Council's decision to release the CCTV footage without his consent. His application was rejected and this decision was upheld by the Court of Appeal with the explanation that there was no violation of privacy because "actions were already in the public domain" and revealing the footage "simply distributed a public event to a wider public." The applicant applied to the European Court of Human Rights and it concluded that "the disclosure by the Council therefore constituted a serious interference with his right to respect for private life. There were no relevant or sufficient reasons to justify the disclosure by the Council without obtaining the applicant's consent or ensuring as far as possible that his identity was masked." The Court therefore awarded him damages for his distress due to violation of his privacy [12].

Case 2: An identity thief using a stolen photocopy of an ID card and VAT number signed two contracts in a web shop with a mobile service provider and picked up two smart-phones. The person whose identity was stolen reported the case to the police and the Personal Data Protection Agency (PDPA).

Epilogue: PDPA made an inspection and requested contracts, delivery reports and a copy of the submitted ID. After discrepancies were found in the contracts (a fake signature) and negligence in the delivery procedures (the ID was not checked), the mobile service provider admitted its mistakes and cancelled the contracts. Police caught the gang with this *modus operandi*. One of the gang members was an insider in the mobile service provider company.

Case 3: In 2001, the police in Tampa, USA, used face scanning and facial recognition software to scan and capture images of football fans at the Super Bowl, without the knowledge of the people involved [13].

Epilogue: The use of facial recognition systems in public places was banned. Why? Different organizations could use faces captured by a facial recognition system to discover places that a person had visited or to scan different large databases in order to profile and/or socially control a person.

Privacy violations described in Cases 1 and 3 could be prevented by de-identification of biometric identifiers, while violation in Case 2 could be prevented by storing personal documents in appropriate safe manner.

The main focus of this paper is the de-identification of biometric identifiers in multimedia documents for privacy protection. It is therefore interesting to view some of the main concerns related to the use of biometrics [11]: (i) biometric data can be collected and shared without the user's knowledge and permission; (ii) biometric data which have been collected for some specific purposes can later be used for other unintended or unauthorized purposes. This is referred to as "functional creep"; (iii) biometric data can be copied or removed from the user and used for secondary purposes; (iv) biometric data can be used to reveal sensitive personal information, such as gender, race, and ethnicity, but also mental and health status; (v) biometric data can be used to pinpoint, locate and track individuals. Even more, by associating biometric data with non-biometric identifiers (name, address, ID and passport number) it can lead to covert surveillance, profiling and social control; (vi) biometric data can be exposed to external attacks due to improper storage and/or transmission.

The biometric templates of an individual may be stolen, modified and shared, and privacy and security may be compromised. There are three aspects of privacy protection of individuals regarding biometric template protection [14]: (i) irreversibility – it should be computationally hard to reconstruct the original biometric template from the stored reference data; (ii) unlinkability – different biometric templates cannot be linked to each other or to the individual who is the source of both; and (iii) confidentiality – protection of the user's biometric template against unauthorized access or disclosure. Recently, efforts have been made to standardize biometric template protection. There are four main biometric template protection schemes: (i) extracting and storing a mathematical sketch of a biometric template; (ii) fuzzy commitment in which a biometric feature vector is bound to a secret message; (iii) encrypting the biometric features at enrolment; and (iv) cancellable or revocable biometrics where the template is transformed using a secret transformation at enrolment, and stored in the system. Recognition is based on matching between a test template which is obtained by using the correct transformation and the transformed version of the enrolment template. Cancellable biometric includes

cancellable face [15], fingerprint [16], iris [17], voice [18] and other biometric modalities. A detailed and comprehensive overview of cancellable biometrics and biometric cryptosystems is given in [19, 20].

Privacy issues and ethical and legal issues related to privacy and multimedia in different contexts, environments and scenarios are subjected to detailed discussion [21, 22]. In [21], privacy protection based on reversible cryptographic obscuration is presented. Additionally, privacy issues in scenarios with multimedia (video and audio) surveillance are considered. The author describes a scenario where a surveillance device intercepts sound and the surveillance constitutes a search. In such a case, the police or government institutions must first obtain a warrant prior to the installation of the device (according to US Title I of the Electronic Communications Privacy Act). Bharucha et al. [22] discuss the ethical implications of real-time multimedia surveillance technology for the privacy and dignity of long-term care residents, personnel and care processes. The authors de-identified privacy sensitive data (face and voice) of all stakeholders (residents, professional and non-professional staff, administrative staff, families and visitors), but only after the filming was completed. This is a weak point of the approach, because third parties may gain access to the recordings before the participants are de-identified.

2.3 Phases of contemporary privacy development

After consideration of privacy at the political and socio-cultural and organizational level and describing a privacy baseline (period 1945–1960), A. F. Westin [9] introduced three phases of contemporary privacy development as follows.

i) The first era of contemporary privacy development, (period 1961–1979), which is characterized by the rise of information privacy as an explicit social, political, and legal issue of the high-technology age. In 1973, a US government advisory committee initially proposed a set of principles to protect the privacy of personal data in recordkeeping systems named Fair Information Practices (FIPs) [23]. The six basic principles of Fair Information Practices are: (i) the existence of personal data collections should be public knowledge; (ii) individuals have the right to review and correct information related to them; (iii) the minimum information necessary should be collected, and, where appropriate, the

consent of the included individuals should be obtained; (iv) personal data should be accurate and complete and retained only for a given time period; (v) data should only be used for the purpose originally intended; and (vi) data should be protected by security safeguards against unauthorized access, modification or use.

In 1970s, European countries began to enact privacy laws applicable to the public and private sectors, beginning with Sweden (1973), the Federal Republic of Germany (1977), and France (1978) [23]. These laws were consistent with FIPs.

ii) The second era of contemporary privacy development, (period 1980–1989). Technologically, this was a period of enhanced computer and telecommunications performance, but without fundamental changes in information-society relationships bearing on privacy;

iii) The third era of contemporary privacy development, (period 1990–now). This is the period when privacy became a first-level social and political issue in Europe and the US, assumed global proportions, and was impacted by 9/11 and its aftermath.

The main framework for privacy and personal data protection in the European Union is The 1995 Data Protection Directive of the European Union (Directive 95/46/EC) [24]. It is an operating basic model for handling personal data that demands the deployment of appropriate technical and organisational measures to protect private information in the course of transferring or processing personal data. This legal requirement along with ethical responsibilities has restricted data sharing and utilisation, while various organisations may require the use of such data for research, business, academic, security and many other purposes. In July 2008, the Information Commissioner's Office (ICO) commissioned a review of the 1995 EU Data Protection Directive (95/46/EC) [25]. This was motivated by the fact that since the introduction of the Directive, the world had witnessed dramatic changes in the way personal data was accessed, processed and used. At the same time, the general public had become increasingly aware of the potential for their personal data to be abused.

The terrorist attacks on September 11, 2001 have had significant impacts on privacy, information law and its practice in the US [26]. Here is the list of the main important acts: USA Patriot Act (2001),

Homeland Security Act (2002), Intelligence Reform and Terrorism Prevention Act (2004.), Real ID Act (2005) and NSA Warrantless Surveillance (2005).

There is an everlasting debate between experts in the field of security and privacy experts about security-privacy balance. They are all aware that there must be a balance between privacy and security because it guarantees foundations of our freedom and democracy. In contemporary times, the balance has shifted towards the security side of scale [3]. The intensity of electronic (dragnet) surveillance at the US state level and local levels, after September 11, 2011, may be illustrated by increasing the budget of Federal intelligence agency from $27 billion (prior to the attacks) to $75 billion in 2013 [6].

A comparison of US and European approaches to privacy legislation is given in [27 - 29]. Summarizing the comparison, we can state that: (i) while data protection and privacy are fundamental rights in the EU and are also applicable in the law enforcement context, there is no equivalent protection in the US [29]; (ii) the basic EU data protection principles such as restrictions on the further use and dissemination of data collected in a law enforcement context, purpose limitation or time limits on data retention do not exist at all or exist only rudimentarily in the US; (iii) in EU law, fundamental rights cover all persons targeted by law enforcement and surveillance measures, regardless of their nationality while US law distinguishes between US and non-US citizens.

Note that in October 2015, the European Court of Justice struck down a 15-year-old agreement known as the Safe Harbour, which was an attempt to bridge differing approaches to data protection in Europe and the US. The Court concluded the data of Europeans are exposed to allegedly indiscriminate surveillance by the US government. The General Data Protection Regulation [30], adopted by the European Parliament in April 2016, represents the reform of EU data protection rules and covers the following main areas: protection of personal data, data transfers outside the EU, data protection on social networks and Big Data services. It was an essential step to strengthen citizens' fundamental rights in the digital age and facilitate business by simplifying rules for companies in the Digital Single Market.

The time period from 2001 until now is characterized by technologies such as internet, wireless communications, data-mining software based on large data-warehousing applications, cloud

computing, drones with video camera and other sensors, the increased use of law-enforcement video-camera systems in public places, and along with the adoption of biometric identification systems by many governments and private organizations.

2.4 Common Criteria for Information Technology Security Evaluation and Privacy-Enhancing Technologies

There is a strong linkage among privacy, privacy protection and technologies designed specifically for privacy protection. The common framework for privacy, privacy protection and technologies is the multipart standard Common Criteria for Information Technology Security Evaluation [31] and Privacy-Enhancing Technologies [32, 33]. Privacy-Enhancing Technologies (PETs) have been developed to protect internally stored personal data that might be privacy–sensitive. It stands for a coherent system of information and communications technology (ICT) measures that protect privacy by eliminating or reducing personal data, or by preventing unnecessary and/or undesired processing of personal data, all without losing the functionality of the information system [32]. An extension of PETs has resulted in a more substantial approach called Privacy by Design (PbD). PbD is a concept developed in the 90s [34]. It combines the principles of Fair Information Practices and a proactive approach to protecting privacy by embedding it into the design specifications of technologies, business practices, and physical infrastructures. A typical example of a system to which PbD has been applied is the De-Identification Camera [35] (Section 5.2).

For the benefit of discussions in this paper, below, we provide the definition of a set of key terms.

i) *personal information* is any information relating to a person,

ii) *personal identifiable information* (or *personal identifiers*) is the personal information, which allow his or her identification,

iii) *privacy concerns* exist wherever personal information containing personal identifiers is captured in *multimedia content* (text, still images, audio and video sequences, and their combination), and

iv) *preservation of the privacy* of persons captured in multimedia content necessitates the de-identification of all of their personal identifiers (we use the term *a personal identifier recognition* to denote biometric-based person identification or verification based on a personal

identifier), e.g. gate recognition means gait-based person identification or verification. Modern computer technologies such as biometrics, cloud computing, ambient intelligence, data-mining, internet services, social networks and audio-video surveillance are privacy intrusive because they allow collecting, extracting, observing, transferring and storing of *personal identifiers.*

3. De-identification and irreversible de-identification

De-identification in multimedia content is defined as the process of concealing or removing personal identifiers, or replacing them with surrogate personal identifiers in multimedia content, in order to prevent the disclosure and use of data for purposes unrelated to the purpose for which the information was originally obtained. It is no doubts that de-identification is one of the basic methods for protecting privacy, while permitting other uses of personal information.

The terms *de-identification* and *anonymization* are often used interchangeably, but some experts make the difference between them. De-identification refers to the reversible process of removing or obscuring any personally identifiable information from individual records in a way that minimizes the risk of unintended disclosure of the identity of individuals and information about them. It involves the provision of additional information to enable the extraction of the original identifiers by, for instance, an authorized body. Anonymization refers to the process of data de-identification that produces data where individual records cannot be linked back to an original as they do not include the required translation variables to do so [36]. It is a one-directional (irreversible) process and does not allow the original identifiers to be obtained from de-identified data. In this paper we use the term de-identification for both approaches, but in some cases we emphasize whether it is a case of reversible or irreversible process. In either case, the de-identification process is required to be of sufficient effectiveness, regardless of whether the recognition attempts are made by humans or by machines. Moreover, in many cases, the process of de-identification has also to preserve the data utility, naturalness and intelligibility [37, 38].

3.1 Taxonomy of the identifiers in multimedia content

The following proposed taxonomy of the identifiers in multimedia content that have to be de-identified in order to protect privacy is inspired by the Safe Harbour approach [39]. According to this approach, which constitutes the guiding principles for de-identification in healthcare applications, there are 18 types of identifiers that have to be de-identified in order to cover the identity of the recipients of health-care services (patients). These are names; all geographic subdivisions smaller than a state; all elements of dates (except year) for dates directly related to an individual; telephone and facsimile numbers; electronic-mail addresses; social security numbers; medical record numbers; health-plan beneficiary numbers; account numbers; certificate/license numbers; vehicle identifiers and serial numbers including license-plate numbers; device identifiers and serial numbers; internet universal resource locators (URLs); internet protocol (IP) address numbers; biometric identifiers; including fingerprints and voiceprints; full-face photographic images and any comparable images; and any other unique identifying number, characteristic, or code, unless otherwise permitted by the *Privacy Rule for re-identification* [40].

Based on the above types of personal identifiers, the identity information extracted from multimedia content can be classified as follows.

i) *Non-biometric identifiers* including text context, speech context, licence plate, specific socio-political and environmental context, dressing style, and hairstyle;

ii) *Biometric identifiers* are the distinctive, measurable, generally unique and permanent personal characteristics used to identify individuals. In the following, they are usually categorized as *physiological* (face, iris, ear, fingerprint) versus *behavioural* (voice, gait, gesture, lip-motion, stile of typing),

iii) *Soft biometric identifiers* provide some vague physical, behavioural or adhered human characteristic that is not necessarily permanent or distinctive (height, weight, eye colour, silhouette, age, gender, race, moles, tattoos, birthmarks, scars) [41, 42]. In most cases *soft biometric identifiers* alone cannot provide a reliable personal identification, but they can be used for improving the performance of recognition [42, 43], or to classify people into particular categories, which is also

privacy intrusive. Figure 1. shows the taxonomy of identifiers in multimedia content, which is adopted as a logical basis for structuring discussions in the remainder of this paper.

It is worth noting that very often multimedia content may simultaneously include biometric, soft-biometric and non-biometric identifiers, which all have to be de-identified in order to protect the privacy of individuals. This can be referred to as *multimodal de-identification*.

Detecting and concealing or removing or replacing personal identifiers in multimedia content is an interdisciplinary challenge that incorporates such scientific areas as natural-language processing, text processing, image processing, pattern recognition, machine learning, speech analysis, video tracking and biometrics.

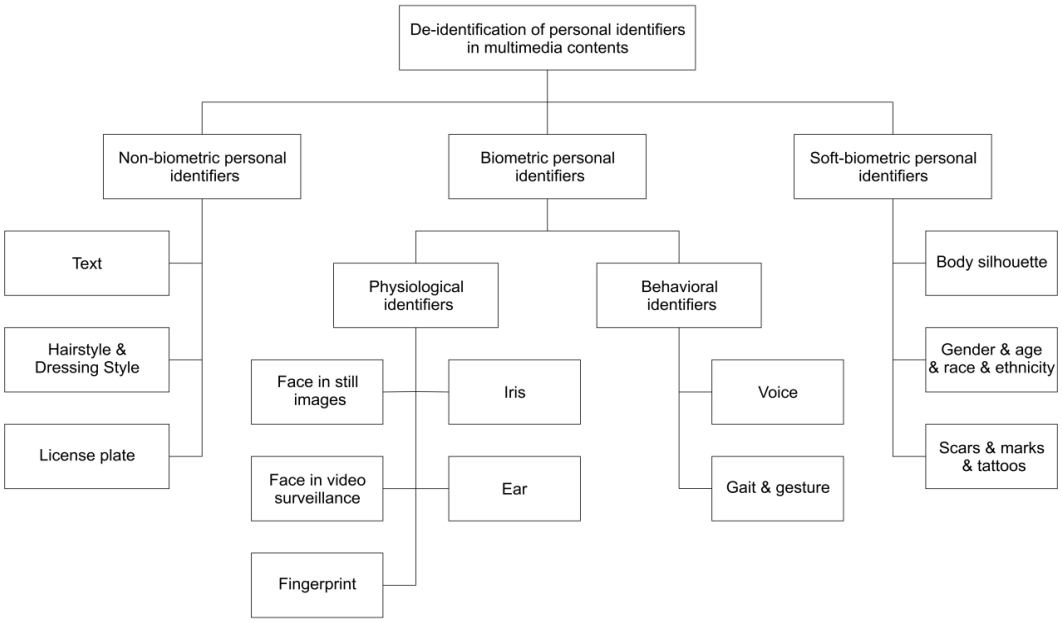


Figure 1. Taxonomy of identifiers in multimedia content.

In the next sections we provide an overview of de-identification of non-biometric identifiers (Section 4), physiological biometric identifiers (Section 5), behavioural biometric identifiers (Section 6), and soft-biometric identifiers (Section 7). Besides the solutions for de-identification, we also discuss the unsolved problems and challenges related to de-identification, assessment of privacy level protection, naturalness and usability of de-identified multimodal contents.

4. De-identification of non-biometric identifiers

4.1 Text de-identification

Research on de-identification was initiated with text-based personal healthcare records (PHRs). The approach in this application area involves the removal of a number of specific categories of information from the text file, and replacing them with realistic surrogate information [44 - 46]. The automated de-identification of text-based PHRs is focused on both highly-structured type-specific records and/or free-text medical records with a highly variable structure. The de-identification methods are based on templates and specialized knowledge of the context for replacing personal health information (PHI) in medical records, or on a complex combination of dictionaries and text-analysis algorithms. Recently, approaches based on a combination of machine learning, heuristics and statistical methods, as well as pattern-matching are used [44].

Reversible de-identification is commonly used in the protection of personal data in health-care and biomedical research [47]. An overview of this de-identification challenge of PHR, the data and the annotation process, the evaluation metrics, and a discussion on the nature of the de-identification systems and the identification of directions for future research are given in [48]. In the context of text de-identification, it is worth noting that medical imagery, which consists of header information, typically in a DICOM (Digital Imaging and Communications in Medicine) format, and image data generated by imaging devices, contains privacy sensitive information in both header and image data. Privacy sensitive information of medical image data can be illustrated by the fact that it is possible to reconstruct a person's face using three-dimensional models generated from computed tomography (CT) and magnetic resonance (MR) imaging [49]. By using a multimodal de-identification approach, the text sensitive information in the DICOM header has to be removed or replaced with surrogate information, while image data have to be de-identified by methods based on reversible privacy filters (see Sections 5.2 and 5.3).

4.2. Hairstyle and dressing style de-identification

Hairstyle and dressing style carry identity-revealing information [50 - 53] and they can be used to classify people into different categories. There is also the problem called "a pair-wise constraint"

identification [54], which means that people can determine that two de-identified face portraits in a video belong to the same person by using clothing, hairstyle, dress style or other cues as alternative information, and so there is a risk of exposing a person's identity. Alternative information that can be useful for identity revealing includes speech context, specific social and political context, and the environment. Relatively little research work has been done in the area of removing or hiding hairstyle and dressing style, as well the above mentioned contexts for de-identification purposes [55, 56].

4.3 License plate de-identification

Web services like Google Street View and EveryScape systematically gather and share large-scale images of public places. The gathered images of public places in their original forms contain privacy sensitive information, such as the faces of individuals and car license numbers on license plates. According to the Safe Harbour approach, this information is among 18 types of identifiers that have to be de-identified in order to conceal the identity of an individual. In [57], the authors focus on the detection of faces and license plates in Google Street View footage, while the de-identifications are simply done by blurring the detected locations (see Section 5.1). A simplified version of the face detector based on a fast sliding-window approach over a range of window sizes is used for the detection of license plates. The detector employs the linear combination of a heterogeneous set of feature detectors, which are based on families of features of varying complexity, encompassing simple but fast features such as bit features, as well as more expensive but more informative features such as Gabor wavelets. The separated detectors for US and EU plates are trained by minimizing the objective function. They belong to a large family of sliding window detectors, such as Schneiderman-Kanade [58] and Viola-Jones detectors [59]. The authors report that a completely automatic system has detected and sufficiently blurred 94 – 96% of the license plates in evaluation sets sampled from Google Street View imagery.

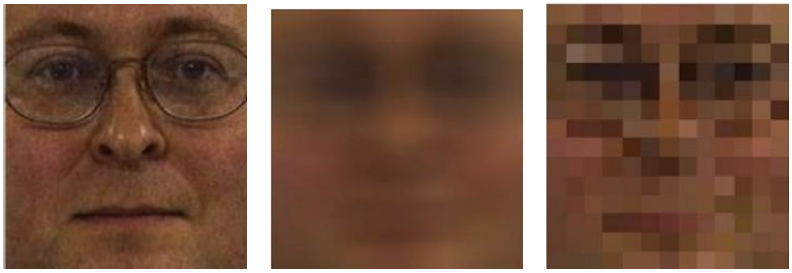
In [60], a method named inhomogeneous principal component blur (IPCB) is proposed. It adaptively blurs different pixels of a license plate by taking into account the prior distribution of sensitive information. Based on the assumption that not all information in the license plate region is privacy sensitive, the authors propose a preservative license plate de-identification method to balance privacy

protection and quality preservation. For example, the state name is usually less sensitive than the license numbers, so only the plate's area with the license numbers should be de-identified. Therefore, selectively blurring or masking only the license number area minimizes the unwanted degradation of the original image and improves its naturalness. The blurring is based on the Principal Component Analysis (PCA) approach - the original plate's area is substituted by a reconstructed area that is obtained by applying a smaller number of eigenvectors. The proposed method is reversible: a de-identified plate can be recovered by knowing the coefficients of each principal component.

5. De-identification of physiological biometric identifiers

5.1 Face de-identification in still images

The main physiological biometric identifier in multimedia content, requiring de-identification for privacy preservation is the *face* [61]. The early research into face de-identification was focused on *face still images*, and recommended the use of ad-hoc approaches such as "black box", “blurring” and “pixelation” of the image region occupied by the face [62, 63]. In *the black-box approach*, after the face detection and face localization in the image, the face region is simply substituted by a black (or white) rectangle, elliptical or circular cover. *Blurring* (Figure 2b); the experiments were performed on the cmu-pie-database [64] is a simple method based on smoothing the face in an image with Gaussian filters using a variety of sufficiently large variances. By applying different variances, different levels of blurred images of the face are obtained [62]. *Pixelation* (Figure 2c) consists of reducing the resolution (sub-sampling) of a face region. Naive methods such as blurring and pixelation might prevent a human from recognising subjects in the image, but they cannot thwart recognition systems.



a) b) c)

Figure 2. Naive methods of face de-identification: a) Original image; b) Blurring: $\sigma^2 = 18$; c) Pixelation: parameter p = 12 [65].

An effective approach that subverts naive de-identification methods is called *parrot recognition* [66]. Instead of comparing the de-identified images to the original images, parrot recognition is based on comparing probe (de-identified) images with gallery images, where the same distortion is applied as in the probe images. It is shown that such an approach drastically improves the recognition rate, i.e. it reduces the level of privacy protection [66]. To achieve an improved level of privacy protection, more sophisticated approaches have been proposed. In [67], an eigenvector-based de-identification method is described. The original face is substituted by a reconstructed face that is obtained by applying a smaller number of eigenfaces. As a result, the face details are lost and the de-identified image becomes harder to recognise. In the same paper, the privacy-operating characteristic (POC) is introduced and used to show, quantitatively, the trade-off between privacy and security. The eigenvector-based method easily produces very unnatural images, but still keeps some of the facial characteristics that can be used for automatic recognition.

In recent years, advances in biometric identification have inspired researchers in the field of de-identification. Examples are the face de-identification methods referred to as *k-Same* [68], *k-Same-Select* [69] and *Model-based k-Same* [70]. By applying the *k-Same algorithm*, to the given person-specific set of images, where each person is represented by no more than one image, a set of de-identified images is computed. Each de-identified image is represented by an average face image of the *k* closest face images from the person-specific set of images. The *k* closest face images in the person specific set are replaced by the same *k* de-identified face images. The *k-Same* algorithm selects

the *k* closest images based on Euclidean distances in the image space or in the PCA coefficient space. Figure 3. illustrates the *k-Same* algorithm (*k* = 4) where for a person-specific set of face images I (which consists of 12 original images), the set of de-identified face images D is computed. The set D consists of 12/*k* identical face images, where each image is represented as an average of the *k* = 4 closest original images.

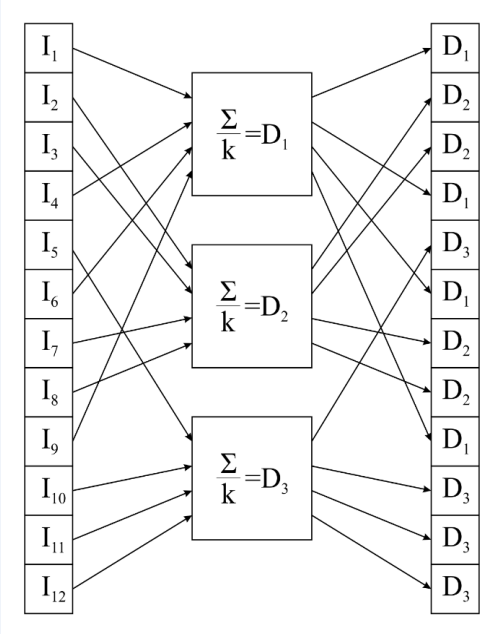


Figure 3. Illustration of *k-Same* algorithm (modified from [68]). As an example it should be noted that the original images I_1, I_4, I_6 and I_9 are represented with the same de-identified face image D_1; I - a person-specific set of face images; D - a set of de-identified face images; ∑ - a sum of the *k* closest face images from a person-specific set of images I.

Figure 4. gives an example of *k-Same* de-identification for value k = 6.

a) b)

Figure 4. *k-Same* de-identification: a) Original image; b) De-identified image for *k* = 6; [65].

It has been shown that the best-possible success rate for a face-recognition algorithm linking a de-identified face image to the correct face image in the set I is 1/*k* [68]. The procedure based on the *k-Same* algorithm is irreversible, guarantees probable privacy (1/*k*), but very often results in "ghosting" artefacts in de-identified images due to image misalignment or an expression variant of the faces present in the *k* images from set I. In order to improve the data utility and the naturalness of the de-identified face images, the *k-Same-Select* is proposed [69]. The algorithm partitions the input set of face images into mutually exclusive subsets using the data-utility function and applies the *k*-Same algorithm independently to the different subsets. The data utility function is usually selected to preserve the gender or a facial expression in the de-identified Due to the use of the *k-Same* algorithm, *k-Same-Select* guarantees that the resulting face set is *k*-anonymized [71]. For both algorithms, there are two main problems: they operate on a closed set I, and the determination of the proper privacy constraint *k*. In order to produce de-identified images of much better quality and preserve the data utility, the *Model-based k-Same algorithms* [70] are proposed – one of which is based on Active Appearance Models (AAMs) [72] and another based on the model that is the result of mixtures of identity and non-identity components obtained by factorizing the input images. Modifications to the *k-Same Select algorithm,* in order to improve the naturalness of the de-identified face images (by retaining face expression) and privacy protection, are proposed in [73, 74].

In [75], the authors proposed a reversible privacy-preserving photo sharing architecture which ensures privacy and preserves the usability and convenience of online photo sharing. The architecture takes into account the content and context of a photo and utilizes a Secure JPEG framework. Visual privacy in a JPEG can be protected by using: (i) naive de-identification where the reconstruction of an original image is performed by extracting from a JPEG header, decrypting and placing back the original pixels; (ii) scrambling, which modifies the original values of the pixels and the discrete cosine transform (DCT) coefficients in a reversible way. The proposed architecture is convenient for privacy protection in social networks and photo hosting platforms (Facebook, Pinterest, Instagram).

In [76], a morphing-based visual privacy protection method is described. The morphing is performed by using a set of face key points (eyes, nose, mouth), both original source and target images, the

interpolation of some pixels between the key points, and dividing both images using Delaunay triangulation. Subsequently, for each pixel in the final (morphed) face image, the pixel's value is computed as a weighted sum of intensities between the corresponding pixels in both images. By using an inverse of morphing (unmorphing), the protected face image can be recovered. The method was tested on a subset of a FERET database and demonstrates that morphed faces retain the likeness of a face while making them unrecognizable. The same authors [77] used a geometrical transformation or warping for face de-identification. The warping is performed in the following steps: (i) select a set of key points (facial features) in the face image (eyes, nose, mouth) and several points around the detected facial features and the sides of the face; (ii) change the coordinates of these points to the destination coordinates by adding or subtracting a random value with a weight which determines the warping strength; (iii) compute the transformation matrix based on the original and destination coordinates. By using the inverse transformation, the original face can be estimated. The warping was tested on a Yale dataset (165 faces of 15 subjects). The test showed that the naturalness and privacy level of protection depend on the warping strength.

5.2 Face de-identification in video surveillance systems

Due to tremendous development and use of visual technologies such as CCTVs, visual sensor networks and camera phones, the term *the visual privacy* is introduced. It determines relationship between collection and dissemination of visual information, the public expectation of privacy, and the legal and ethical issues surrounding them.

A valuable review of visual privacy and visual privacy protection methods is given in [38]. Authors classified the methods for privacy protection of individuals appearing in videos into five large categories: (i) intervention - preventing someone to capture private visual data from the environment; (ii) blind vision - image or video processing in an anonymous way; (iii) secure processing – process visual information in a privacy respectful way; (iv) redaction – methods based on image filtration, encryption and *k*-same family algorithms, object/people removal, visual abstraction/object replacement, and (v) data hiding – steganography and watermarking-based methods.

Most of the described methods in Section 5.1 are applicable for the de-identification of still, frontal facial images or facial images in a television broadcast, but not necessarily suitable for use with video-surveillance systems. The reasons are: (i) such privacy-protection schemes degrade the visual quality needed for security; (ii) they do not preserve the naturalness of the de-identified moving images; (iii) most of them modify the surveillance videos in an irreversible fashion; (iv) real-time processing is required [54].

Special attention in the field of privacy protection is now being devoted to automatic *face de-identification in video surveillance systems* because of their privacy-intrusive characteristics [5]. The process of automatic face de-identification in videos includes face detection, face tracking and face masking. Currently, there are two main approaches to face detection [78]: the feature-based approach and the image-based approach. The feature-based approach uses low-level analyses (based on edges, colour, grey-level, motion), feature analyses (facial feature extraction, face detection based on anthropometric measures, statistical-based grouping of facial features in face-like constellations), and active shape models (snakes, deformable templates, point distributed models). The image-based approach detects faces via a learning procedure that classifies examples into face and non-face prototype classes. The main methods are linear subspace methods, neural networks, and statistical methods. A useful overview of the face-detecting methods in images and videos is given in [79].

In the time period 1998. – 2005. there were face-detector candidates for use in videos as follows: neural network based detector [80], Schneiderman-Kanade detector [58], Viola-Jones detector [59], local edge orientation histograms based (EOH) [81], and histograms of oriented gradients [82].

In [54], a detector based on the combination of background subtraction, bag-of-segments features and a Support Vector Machine (SVM) is described. The authors reported 92% accuracy for SVM classifier trained with 1,500 examples, in a test set consisting of 1,000 examples.

More recently, new methods have been proposed for face detection, pose estimation and landmark localization in the wild. Pose estimation and face landmark localization are important to preserve naturalness de-identified videos. In [83], a unified model for face detection, pose estimation and landmark localization using a mixture of trees with a shared pool of part templates is described. The

authors compared the results of face detection of proposed approach with OpenCV frontal and profile Viola-Jones detector, Boosted frontal and profile face detector, deformable part model (DPM) and commercial systems (Google Picasa's face detector, face.com). The proposed method significantly outperform popular detectors currently in use, and are on par with commercial systems trained with billions of examples, such as Google Picasa and face.com. In [84], the multiple registered image channels are computed using linear and non-linear transformations (e.g. gradient histograms, colour (including grayscale, RGB, HSV and CIE-LUV), gradient magnitude, Gabor filters, and Difference of Gaussian (DoG) filters) of the input image. In the next step, features are extracted from each channel using sums over local rectangular regions. These local sums and features, based on Haar-like wavelets, their various generalizations, and local histograms, are efficiently computed by using multiple sums and integral images. The proposed method combines the richness and diversity of information from image channels with the computational efficiency of the Viola and Jones detection. In [85, 86], in order to avoid the computational bottleneck of many modern detectors, i.e. the construction of an image pyramid, the authors proposed fast method for object detection based on approximation of multi-resolution image features, instead of their computing explicitly. Based on such an approach, the authors demonstrated on pedestrian detection tasks (INRIA, ETH, and TUD-Brussels databases) that speedup for $1 - 2$ orders of magnitude was achieved compared to state-of-the-art detection performance (6 fps on 640×480 image resolution).

Face tracking is the process of locating a moving human face (or multiple human faces) in a sequence of frames. In the case of multiple human faces, the process should be capable of discriminating and tracking individual faces in the given video. Tracking is based on features such as segmented regions, skin-colour models [87], local binary patterns (LBP) [88], a combination of LBP and skin-colour information [89], a combination of shape and texture information [90], and histogram-based Mean-Shift features [91]. Face tracking includes the prediction of a face location in the next image frame based on the motion model or the information obtained from the previous consecutive frames. Kalman filters and particle filters are normally used for predictions. On the basis of this prediction, the face tracking can be treated as a local search problem where the features are locally searched within a

search window instead of the entire image. In order to increase the tracking speed, an adaptive search window is used. Its size may grow with the square of the maximum velocity of the face.

The combination of face detection and tracking, i.e. the combination of the spatial and temporal correspondence between frames, can improve the effectiveness of the localization of faces. An example of such an approach is applying a bi-directional tracking algorithm that combines face detection, tracking and background subtraction [54]. The effectiveness of the face detection and tracking is very important because the face has to be detected and de-identified in each frame of the videos. If the face cannot be detected even in only one frame (and so is not de-identified), it leads to a major degradation in the privacy protection.

Each localized and traced face region in *each frame* has to be de-identified by some effective means. A possible method for this purpose is *masking*. Some approaches to face masking for privacy protection in video-surveillance systems follow techniques that are used in still-face images.

In [92], privacy filters with varying strength degrees, based on simple approaches such as masking, blurring, pixelation, warping and morphing, are applied on the FERET database to investigate the influence of the filters' strength parameters on the performance of PCA-, Linear Discriminant Analysis (LDA) -, LBP-based face recognition algorithms. The authors concluded that the morphing filter is the best choice among the tested privacy filters. In [93], a cartooning privacy filter, which converts raw images into abstracted frames where the privacy revealing details are removed, is described. Cartooning applied on pre-selected privacy sensitive regions of interest (ROIs) demonstrated an acceptable level of privacy protection while maintaining a good utility level.

An alternative approach to face de-identification, especially popular in the video-surveillance domain, is based on distortion applied to the face image by using transform-domain scrambling methods. For example, in [94, 95], the authors have proposed two scrambling methods for video coding standard H.264/AVC – one of the most commonly used formats for the recording, compression, and distribution of video content. Both methods scramble the quantized transform coefficient of each 4×4 block of the region of interest by pseudo-randomly flipping their sign, or by applying a random

permutation of the coefficients. These two methods are fully reversible – the authorized user, by using a secret encryption key, can reverse the scrambling process and recover the image of the face.

It is important to note that, the last few years have witnessed considerable attention towards real-time, privacy-protection video systems. Examples of systems in this category are Respectful Cameras [96], PrivacyCam [97], TrustCam [98], and the De-Identification Camera [35]. In the Respectful Cameras system, users who wish to be protected wear colour markers (hats or vests) that are tracked and the faces of such users are masked in real time. The tracker is based on a 9-dimensional colour space and the combination of a particle filter and a probabilistic AdaBoost algorithm. Because of the type of markers used, the system is well suited to dynamic scenes. An elliptical white cover is used to hide the faces of users.

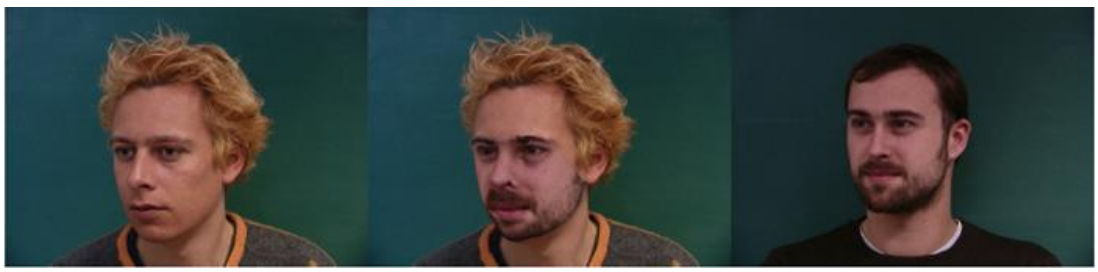
The DSP-based PrivacyCam [97] system implements the real-time Privacy through an Invertible Cryptographic Obscuration (PICO) process that consists of five basic steps: (i) capture of the image with a camera; (ii) detection of the region of interest (face detection, skin detection, motion detection); (iii) exchanging public key, generating session key, and storing the secured key along with the protected region information; (iv) selective encryption of the region (human face region) to be protected. The face is protected by scrambling the coefficients used for the JPEG image encoding.

The TrustCam prototype system [98] consists of a network of trustworthy cameras and a control station. Each camera is equipped with an individual Trusted Platform Module (TPM) that is used for the data encryption to hide the identity of individuals captured in a video.

The De-Identification Camera [35] is an example of real-time privacy protection at the sensor level. The de-identification pipeline in this case consists of the background segmentation (motion detection), person detection based on histograms of gradients (HOG) [82], tracking based on Mean-Shift, segmentation of an image based on a bounding box that forms the video tube for each person in real time, and a de-identification transform applied to the video tube. The real-time de-identification transform uses two types of “naive” procedures: the Gaussian blur of pixels inside a bounding box, and the binarization of the pixels inside the bounding box. Note that the De-Identification Camera performs de-identification of the whole human figure. Due to the scrambling of the coefficients, or

using “naive” de-identification techniques, all the above-described systems produce de-identified videos that do not preserve the naturalness of the original videos.

A more sophisticated privacy protection in videos is obtained by replacing a face with a generic face. The preliminary results of such an approach applied to video sequences are shown in [70]. Recently, in order to improve the naturalness and utility of a de-identified video, the adoption of de-identification methods for still images is proposed in [99]. Normally, the faces captured in a video sequence are of varied poses. Such variations may range from a full left profile to a full right profile (yaw angle from -90^0 to +90^0) and a pitch from -90^0 to +90^0, while the roll is usually more restricted. Following the idea from *k-Same-Select* [69], where images are grouped before de-identification to preserve the facial expression and the gender, the proposed approach groups the face images into a person-specific set of images according to their poses. Each person-specific set is represented by an active appearance model. A raw face image is matched with each of the active appearance models of a person-specific set of images. The model with the best matching based on shape and texture is chosen to represent the pose of the raw face image. Then, from the images in the selected person-specific set of images, one image is chosen to replace the texture of the raw image. The shape of the de-identified face image remains the same as that detected during the model fitting, but the texture is changed. Note that in order to enhance the privacy protection, instead of using the most similar appearance for the raw image, the appearance of an image that is far enough (q-far based on the Euclidean distance) is used [99]. The proposed de-identification method is irreversible. Figure 5. illustrates the above-described approach.



a) b) c)

Figure 5. Illustration of the q-far de-identification method [99]: a) original image; b) de-identified image q = 35; c) image used for the face swapping.

In [56], the authors give the general framework of de-identification by describing different scenarios of video capturing (casual videos, public surveillance and private surveillance videos), criteria for de-identification and methods of subverting the de-identification. They proposed a method of de-identification that consists of three modules: Detect-and-track, Segmentation and De-identification. The detect-and-track module combines a HOG-based person detector and a robust tracking algorithm. The tracking algorithm uses a patch-based recognition approach: the object is divided into multiple spatial patches and each of them is tracked by a voting mechanism based on the histogram of the corresponding image patch [55]. The system uses the bounding boxes of the person in every frame and forms a video tube across time. Each detected person in a video has his or her own video tube. The segmentation of the video tube is performed by using the so-called fixed-size voxels ($x \times y \times t$) in the spatial (x, y) and temporal (t) domains. The result of the segmentation is the classification of the voxels into two classes: foreground and background. The de-identification is performed on foreground voxels by applying the exponential blur of pixels in the voxel or line integral convolution. The implemented system was tested on standard databases like CAVIAR and BEHAVE.

5.3 De-identification in drone-based surveillance systems

Drones (RPAS - Remotely Piloted Aircraft Systems or UAV - Unmanned Aerial Vehicles) are aircraft without a human pilot on board, which are guided by a remote pilot. Drones normally carry video camera(s), but they can be equipped with high power zoom, thermal, night vision, Wi-Fi sensors, and microphones. They might also have the capability of recording and storing images or video footage and uploading the images/video to the internet.

Micro-drones (of a weight up to 2 kg) and small drones or mini-drones (of a weight up to 20/25 kg) are widely used in leisure time and in commercial applications, such as video surveillance and inspection, and photography, on account of their affordable prices (from a few hundred to more than twenty thousand Euro). Due to the drones' characteristics (a mobile view in 3D, the fact that they are often non-detectable, have the ability to observe a scene in detail and access different locations, and follow an object of interest), their video surveillance scenarios can be considerably different from those associated with "classic" CCTV surveillance systems. As a consequence, new issues for the risk

of privacy and data protection have arisen, especially when drones are used in illegal, unsafe or irresponsible ways. Typical examples of privacy violation are situations where a drone is very close to a room or bathroom window, or when it captures images of people in their gardens. Although privacy expectations are greatly reduced in public places, the non-governmental use of a drone to capture images and other information taken while an individual is in a public place could nonetheless constitute an invasion of privacy. Some national agencies for privacy and data protection, as well as bodies of the European Parliament, the USA, Australia, Canada and other countries are intensively working on documents related to the privacy and data protection implications of the (civil) use of drones [100-102].

The problem of drone-based surveillance and its effects on privacy, from the ethical and legal aspects, have been elaborated in papers [103-105]. The common conclusion is that, based on current trends of technological development, law enforcement interests, political pressure and pressure from industry, and the lack of legal safeguards, it is clear that drones pose a looming threat to privacy and policy, and therefore regulatory responses are necessary. Regarding the ethical issue, it is assumed that the actions of drones are subject to ethical evaluation based on the actions of the person controlling the drone, the intentions of that person and the consequences produced by the drone. This raises privacy and ethical concerns, including issues of safety, discrimination, and the potential dehumanisation of the person or persons surveilled.

Additionally, in the absence of a comprehensive legislative framework, there is a need for a more flexible approach – one that proactively provides strong privacy protection and stimulates innovation in a win-win manner. In short, the subject of drones is one that is ripe for the attention of Privacy by Design. Until now, little has been done on the technical aspects of privacy protection for mini drone-based surveillance scenarios. In [106], the authors tested the five privacy filters: blurring, pixelation, masking, morphing [76] and warping [77] for privacy protection using their own video data set of typical drone-based sequences taken in a parking area. The dataset contains 38 video footages (16 to 24 seconds) captured in full HD resolution, captured by the mini-drone Phantom 2 Vision+. Privacy filters were applied, depending on the video surveillance scenarios, on the following manually

annotated privacy sensitive ROIs: body silhouette, facial region, accessories (bag, backpack), license plate and video capture information (video format, resolution, frame rate). For an assessment of the trade-off between privacy protection and the intelligibility of the de-identified videos, for each privacy filter and its different strength level, the authors used a crowdsourcing approach [107]. In our opinion, there are many problems related to de-identification for drone-based surveillance scenarios: automatically real-time and robust detection and localization of privacy sensitive ROIs, real-time adaptive adjustment of filter parameters due to changing the perspective view of the on-board camera, simultaneously using different types and sizes of privacy filters for different privacy sensitive ROIs, and a trade-off among intelligibility, privacy protection and naturalness. The above problems should be solved in the Privacy-by-Design approach. In [108], a simple false colouring method of an entire frame or ROI was applied for privacy protection in short clips captured by a surveillance mini-drone dataset. False colouring preserves privacy without compromising pleasantness and intelligibility, and it is applicable for a real-time system.

5.4 Fingerprint de-identification

Fingerprint still images as multimedia documents, at first glance, should not be a focus of interest in this paper for two reasons. First, in many situations fingerprint recognition is categorized as an *overt biometric* application, i.e. a person is cooperative and aware that he or she is being subjected to recognition [109]. Second, in the centre of our interest are multimedia documents mainly collected at a distance. However, there are two important reasons which have prevailed in the decision to include fingerprints. First, based on the Biometric Market Report [110], fingerprint-based biometric systems are the leading biometric technology in terms of market share. Consequently, with the widespread applications of fingerprint techniques in recognition systems, the privacy protection of the fingerprint becomes an extremely important issue. Second, according to the newest reports of ongoing research [111], it is possible to detect fingerprints by shining polarized light onto a person's hand at a distance of up to two meters and analyzing the reflection using two cameras configured to detect different polarizations. Based on the captured fingerprint image, it is possible to identify a person at a distance. This could be a privacy threat in the near future.

It is worth noting that fingerprints, besides identification information, carry additional private, sensitive information. Based on fingerprints, one can make an inference about gender [112], ethnicity [113], diseases such as Huntington's chorea and Parkinson's [114] and Alzheimer's [115], and others.

In traditional fingerprint-based recognition systems, fingerprint templates can be the subject of different types of attack: from a sensor (fake finger), through a feature extraction module, to a database with stored templates.

Fingerprint still images may be de-identified with the usual de-identification procedures such as black box, blurring, pixelation, replacement by a synthetic fingerprint [109] or by applying privacy filters based on image morphing and/or block scrambling. In addition, feature perturbation and noninvertible feature transforms [116], as well as watermarking techniques, are used for hiding biometric templates [117].

In order to protect the privacy of a fingerprint database for the authentication system, instead of an original fingerprint image, a binary thinned fingerprint image is used in the enrolment phase [118]. Additionally, the user identity is hidden in the thinned fingerprint image based on a data embedding key. Data are hidden by adding some boundary pixels in the thinned fingerprint. The template with a hidden identity is stored in an online database for user authentication. During the fingerprint matching process, first the added boundary pixels are removed and the original thinning fingerprint is recovered and then it is used for matching with the live thinning fingerprint. The same authors proposed a method for protecting fingerprint privacy based on a combination of two fingerprints captured from two different fingers of the same person [119]. From one fingerprint image, the minutia positions are extracted, while the orientation is taken from the other fingerprint. The reference points are extracted from both fingerprint images. Based on these extracted features, the combined minutia template is generated and stored in the database. The complete minutiae feature of a single fingerprint is protected, and an attacker is unable to reconstruct the complete minutiae feature of a single fingerprint. By using the reconstruction approach, it is possible to convert the combined minutiae template into a synthetic real-look fingerprint image [119]. A similar approach to fingerprint de-identification is

proposed in [120]. It is based on mixing two fingerprint images in order to generate a new cancellable fingerprint image, which looks like a plausible fingerprint (Figure 6).

a) b) c)

Figure 6. Mixing fingerprints: a) Original fingerprint; b) Transformation function - fingerprint from a different finger; c) a new mixed fingerprint image that obscures the identity of the original fingerprint [120].

Methods used for privacy enhancement based on different types of distortion of original biometric templates at the signal or feature level may also be applied to hide soft-biometric identifiers (gender, ethnicity) and/or medical information in fingerprint templates. In [121], the authors describe a relatively simple method of fingerprint de-identification for gender estimation. The proposed approach is based on image filtering in the frequency domain. The linear filtering process applies blurring by attenuating the high-frequency content. Certain frequency components are suppressed, while others are amplified. The de-identified fingerprint image is obtained by using the inverse of the Fourier transform. Experiments have shown that the gender estimation accuracy in de-identified fingerprint images for 100 users is reduced from the initial 88.7% (original fingerprints) to 50.5%.

To the best of our knowledge, apart from [119] and [121], there has been no research to evaluate the degree of protection of medical or other privacy sensitive information for such distorted fingerprints and its impact on the identification performance. In [119], the authors report that the recognition system based on a virtual fingerprint obtained by the combination of two different fingerprints achieved a relatively low error rate with FRR = 0.4% and FAR = 0.1%.

5.5 Iris de-identification

Iris represents an important biometric identifier and it enables an efficient approach to reliable, non-invasive identification of people due to its utmost cross-person variability, and minimal within-person variability across time [122, 123]. Most iris-recognition systems require users' cooperation to collect images of adequate quality. Due to the small size of the iris (about 10 mm in diameter) and the required typical resolution between 100 and 200 pixels across the iris diameter, the images are captured at a relatively close standoff (i.e. between 15 to 50 cm), where the standoff is the camera-to-subject-iris distance. A short overview of the main iris-recognition systems and their comparison is given in [124]. Most commercial iris-recognition systems operate at a standoff between 0.1 and 0.45 m, with a verification time of 2 to 7 seconds [124, 125]. However, the Iris at a Distance (IAD) system developed recently provides the capability to identify a person at a range of more than one metre in less than a second [126].

The recent iris-recognition technology is oriented to reducing the need for subject cooperation, reducing the time of image acquisition and increasing the distance between the sensor and the person [125, 127-132]. For example, in [131] the authors introduced the IAD prototype system, which is capable of acquiring an iris image at 30 metres standoff and perform iris recognition (Figure 7.).

a)

b)

Figure 7. a) The IAD prototype system; b) View of the eye at 30 meters by Iris Image Acquisition Camera [131].

Based on the characteristics of the current iris-recognition systems at a distance, and expected future advances in the field, it can be concluded that iris de-identification for privacy protection is a growing problem. An additional complexity to note is that most IAD systems combine face and iris image acquisition. Therefore, both biometric identifiers have to be simultaneously de-identified, i.e. a multimodal de-identification has to be applied.

To date, however, research into iris de-identification for privacy protection has been rather limited. A rare study related to de-identification of the eye areas, and thus the iris, is presented in [133]. The proposed system for the reversible de-identification of an eye region consists of two modules: an automatic eye-detection module and a privacy-enabling encoder module. The automatic eye-detection module in real time locates the human-eye region by a combination of colour-based and Haar-like/GentleBoost methods. The input to the privacy-enabling encoder module is the pixel location information of both eyes in the given input frame. Based on a JPEG XR encoder the macrobloks consisting of 16 × 16 pixels of located eye region are scrambled. The privacy-enabling JPEG XR encoder utilized three encryption techniques (Random Level Shift, Random Permutation, Random Sign Inversion) to transform the coefficients of frequency sub-bands on a macro-block basis. The de-identified images, due to scrambling, lose their original naturalness, but they prevent iris recognition. Also, depending of the dimensions of the scrambling block, the proposed scheme successfully prevents any correct face identification. Figure 8. depicts the organization of the eye region scrambling module.

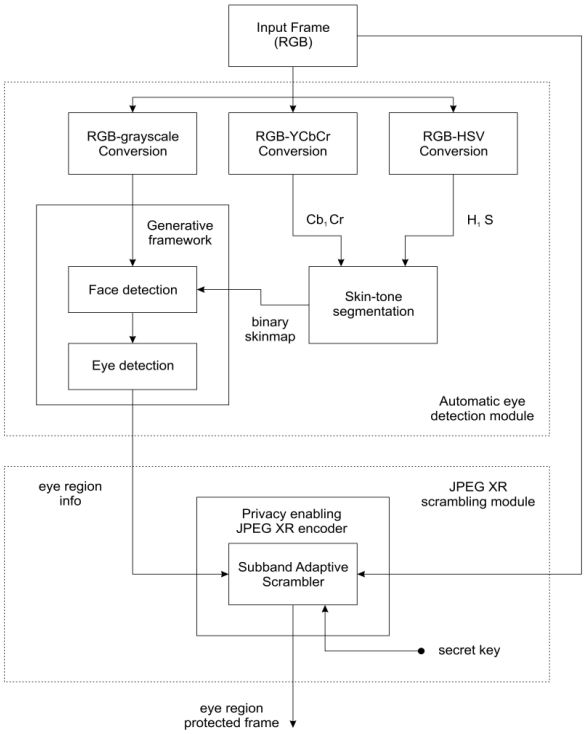


Figure 8. Eye region scrambling module [133].

5.6 Ear de-identification

Despite the fact that the face and iris, in addition to fingerprints, are the most used in biometric technologies for person recognition, they both have a number of drawbacks. Face-based biometrics can fail due to the changes in head pose, facial expressions, the growth of a beard, hair styles, the presence of obstacles (glasses, scarf, or collar), cosmetics, aging, and/or changing the illumination conditions in unconstrained environments. An iris is stable and consistent over time, but due to the relatively small dimension it requires a high-resolution camera and a long-distance Near-infrared (NIR) illuminator for image acquisition at a distance. Therefore, *a human ear* is offered as an alternative physiological biometric identifier for non-invasive person identification or verification at a distance. In [134, 135], comprehensive surveys on two-dimensional (2D) and three-dimensional (3D) ear recognition are presented. These studies have covered over 80 publications on ear detection and recognition in the period 2007–2012.

A 2D ear image can be easily acquired from a distance, even without the cooperation of the subject. This fact makes ear-based recognition systems also interesting for applications in intelligent video-

surveillance systems [136-138]. Until now, ear-recognition systems were successfully tested in controlled indoor conditions [134]. There are some unsolved problems in automatic ear recognition relating to the disruptive factors present in real-life scenes, like pose variations, scaling, varying lighting conditions, and hair occlusion, and these open up new research areas.

Despite the relatively long period of research in the field of automatic ear-based person recognition and its maturity, as far as we know, there are no existing commercial 2D or 3D ear-based biometric systems for automatic person identification or verification. This is the main reason for lack of research in the field of ear de-identification for privacy protection.

6. De-identification of behavioural biometric identifiers

6.1 Voice de-identification

Such biometric identifiers as the face, iris and ear refer to *the visual identity* of a person. However, in addition to a visual identity, a person has *an audio identity*. This is based on the speech signal, which carries privacy-sensitive information such as gender, age, emotional state, health status, level of education, origin and the identity of the speaker. The human voice is a unique pattern that identifies an individual – there are no two individuals that sound identical [139]. Voice is a significant modality that can be used effectively by humans and machines for the recognition of individuals. Applications and services such as audio-video surveillance, speech-based services, life-log systems and telephone-based services enable person identification based on voice, and therefore flag the importance of privacy protection.

The human voice is usually classified as a behavioural identifier in the field of biometrics [140, 141], but it is a hybrid of physiological and behavioural identifiers. A voice pattern is determined by physiological properties, such as vocal folds, vocal tract shapes, and the characteristics of the excitation source (lungs, trachea), but it also conveys behavioural characteristics: rhythm, intonation, vocabulary, particular accent and pronunciation pattern, and talking style.

Voice-recognition systems (i.e. voice-based person identification or verification systems or speaker recognition systems; see Section 2) are classified as text-dependent (or fixed-text) and text-independent (or free-text) systems [139]. Text-dependent systems are suitable for cooperative users

and require the speaker to say a certain phrase. Text-independent systems have no such request and any speech can be captured and analysed in order to verify or identify a user. The text-independent voice recognition systems that require low-level or even no user cooperation are particularly interesting from the privacy point of view. In [142], a text-independent privacy-preserving speaker verification system based on a password matching principle is proposed. The process of authentication is based on a client-server model, where the speaker verification system has the role of server, and the user executes a client program on a network-enabled computation device (e.g. computer or smart-phone). The authentication system does not observe the (raw) speech input provided by the user. Instead, speech input is represented in the form of 2496-dimensional supervectors (64 × 39, where 64 is the number of components of the Gaussian Mixture Model (GMM) and 39 is the dimension of the Mel Frequency Cepstral Coefficients (MFCC)-based feature vector) on which the cryptographic hash function is applied. The speech samples, needed for matching in the verification phase, are stored in the same form in the internal storage of the system. So, the speech samples are irreversibly obfuscated from the system and this one-way transformation preserves the privacy of a user's speech utterances.

Voice de-identification is based on the principles of voice transformation (VT). Voice transformation refers to modifications of the non-linguistic characteristics of a given utterance without affecting its textual content. The non-linguistic information of speech signals, such as voice quality and voice individuality, may be controlled by VT [143], which is based on three types of voice modifications [144]: source, filter and their combination. Source modifications include time-scale, pitch and energy modifications, while filter modifications refer to a modification that changes the magnitude response of the vocal tract system. Voice conversion [145-147] is a special form of VT where the characteristics of a source speaker's voice are mapped to those of a specific (target) speaker. Voice conversion may be text-dependent or text-independent. In the first case, during the learning phase a parallel corpora (training material of source and target speaker uttering the same text) is required. This is the main limitation of using such an approach for voice de-identification in real-world applications.

Text-independent voice conversion [148-150] does not require parallel corpora in the learning phase and it is more realistic for speaker-privacy protection.

One of the earliest proposed voice-conversion methods that can be used for de-identification is described in [146]. The authors present a text-dependent voice-conversion method based on vector quantization and spectral mapping. The method produces a mapping codebook that shows correspondences between the codebook of the source and target speaker. The voice-conversion method consists of two sets: a learning step and a conversion-synthesis step. During the learning step, based on the parallel corpora, the mapping codebooks for several acoustic parameters that describe a mapping between the vector spaces of two speakers are generated. The synthesized speech from using the mapping codebooks is generated in the conversion-synthesis step. The evaluation of the proposed method (for male-to-female and male-to-male conversion) is performed subjectively.

Voice de-identification for the privacy protection of life-log video [151, 152] is based on voice distortion by altering the pitch by the Pitch-Scale Synchronous Overlap and Add (PitchScale SOLA) method. The distortion is accomplished in two steps, i.e. by time stretching the audio signal, and then re-sampling it to obtain the original length.

In [153], the authors propose a transformation of the speaker's voice that enables the secure transmission of information via voice without revealing the identity of the speaker to unauthorized listeners. Owing to the transmitted key, which allows the authorized listeners to perform back-transformation, the voice de-identification is reversible. The authors use a strategy for de-identifying these results in the speech of various speakers to be transformed to the same synthetic (target) voice. They use the GMM-mapping based VT to convert a relatively small set of source speakers (24 males) to a syntactic voice. The proposed VT system has training and a testing or transformation phase. During the training phase a parallel corpora of utterances is used. The authors tested different voice-transformation strategies (standard GMM-mapping-based voice transformation, de-duration voice transformation, double voice transformation, and transterpolated voice transformation). The best results for de-identification are obtained with transterpolated voice transformation (100% de-identification rate for the GMM-based the voice-identification system, and 87.5% for Phonetic voice-identification system). In [154], the same authors present voice de-identification via voice transformation, similar to [153], but de-identification with larger groups of speakers is easier and it

can keep the de-identified voices distinguishable from each other, which contributes to its naturalness. They reported a 97.7% de-identification rate for male and 99% for female speakers.

A novel scheme for voice de-identification, where a set of pre-calculated voice transformations based on GMM mapping is used to de-identify the speech of a new speaker, is presented in [155]. The scheme enables the online de-identification of speakers whose speech has not been used in the training phase to build a voice transformation. The scheme uses automatic voice identification within the set that is used to build pre-calculated voice transformations to select the appropriate transform, which is then used to de-identify the speech of the new user. The approach avoids the need for a parallel corpus, even for training of the initial set of transformations based on GMM mapping, and it was inspired by an approach that is used for face de-identification (e.g., *k-Same*). The preliminary experiments showed that the proposed scheme produces de-identified speech, which has satisfactory levels of naturalness and intelligibility, and a similar de-identification rate in comparison with previous VT systems [153, 154].

In [156], an approach to voice de-identification based on a combination of diphone recognition and speech synthesis is proposed. De-identification is performed in two steps. First, the input speech is recognized with a diphone-based recognition system and converted into phonetic transcription. In the second step, phonetic transcription is used by a speech synthesis subsystem to produce a new speech. With this approach, the acoustic models of the recognition and synthesis subsystems are completely independent and a high level of protection of speaker identity is ensured. Two different techniques for speech synthesis are used: one is Hidden Markov Model (HMM)-based and one is based on the diphone Time-Domain Pitch Synchronous Overlap and Add (TD-PSOLA) technique. Since every user's speech utterance is converted into the speech of the same speaker (whose data were used during the training phase of the synthesis subsystem), the described process of de-identification is irreversible. The system is applicable in different scenarios where users either want to conceal their identity or are reluctant to transmit their natural speech through the communication channel. The proposed voice de-identification system runs in real time and is language dependent and text independent. The obtained de-identified speech was evaluated for intelligibility and evaluated in

speaker recognition experiments by a state-of-art speaker recognition system (i-vector/Probabilistic LDA). The experiments showed that the speaker recognition system was unable to recognize the true speaker identities from the de-identified speech with a performance better than chance, while the de-identified speech was intelligible in most cases.

6.2. Gait and gesture de-identification

Gait is defined as a manner of walking and represents a behavioural biometric characteristic [157, 158]. Gait, as a body gesture, which is usually a motion without meaning, conveys information that can be used for person identification or for diagnostics. Besides the dynamics of individual walking, gait includes information about individual appearance, such as silhouette, leg length, height, even age, and gender [159, 160]. By introducing visual surveillance systems in people's daily lives, and owing to the development of computer-vision techniques, it is possible to recognize non-cooperating individuals at a distance based on their walking characteristics.

In general, there are two basic approaches to gait recognition: sensor-based and video-based. In sensor-based recognition the individuals are cooperative and have tactile and wearable sensors. This approach is normally used in medicine for diagnosis of patients' health status. In a video-based gait-recognition approach, optical cameras are used to obtain the videos of the walking individual(s) [161]. There are two common categories of automatic video-based gait recognition [162]: model-based and appearance-based (or model-free) approaches. Model-based approaches [160, 163, 164] rely on the identification of specific gait parameters in the gait sequence and extract the motion of the human body by means of fitting their models to the input gait sequence. Such models are view and scale invariant, but require high-quality gait sequences. Model-free approaches [165-167] do not require structural models of human motion. They establish a correspondence between successive frames in the video sequence based upon a prediction or estimation of the features related to position, velocity, shape, texture, and colour. One of the most popular approaches to gait recognition is silhouette-based gait recognition [165, 167, 168]. In [165], the authors proposed a simple baseline method for person identification based on the body silhouette and the gait, which provides a lower bound against which to evaluate more complicated procedures.

A gait recognition process typically consists of the following phases: capturing the walking sequence, background subtraction, feature extraction, and recognition where the extracted gait features are compared with gait features that are stored in a database.

The performance of gait recognition systems is evaluated on the HumanID challenge database using different ranks (rank-1 and rank-5) [169]. The experiments have shown that for the baseline algorithm, for twelve experiments, the average recognition rate for rank 1 was 40.95%, while it was 64.54% for rank 5. The different gait-recognition algorithms, based on the HMM, LDA, and Gabor filter approaches, achieved rank-1 recognition rates from 42% to 60%, and for rank 5 it was from 65% to 78%.

Based on the state of the art for gait recognition systems, their characteristics and performances, we can conclude that gait-based technologies can be used for biometric-based person verification in controlled environments. It is technically unfeasible for large-scale surveillance systems to record all the gait parameters of individuals in public places, as well as to identify them by searching in a database [162].

Very few studies have been directly geared towards *gait de-identification*. The study in [170] presents an automated video-surveillance system designed to ensure the efficient and selective storage of data, to provide a means for enhancing privacy protection, and to secure visual data against malicious attacks. The approach to the privacy enhancement of captured video sequences is based on two main steps: the first step is performed by the salient motion detector, which finds ROIs (corresponding mainly to moving individuals), and the second step applies to those regions with a procedure of information concealment based on a scrambling technique described in [95]. The DCT-based scrambling is applied to each ROI, represented by a rough binary mask, which covers a silhouette of the moving individual, so the gait information is obscured. Image regions corresponding to the involved individuals in the scene are distorted, while the scene still remains comprehensible. Owing to the reversible scrambling procedure, the authorized user can get a clear video sequence and reveal all the privacy details by using the embedding and scrambling keys. The de-identified videos, due to the scrambling procedure, do not preserve the naturalness of the original videos.

In [55, 56], *gait de-identification* based on two de-identification transformations, i.e.(for the definition of voxel see Section 5.2), and line integral convolution (LIC) is proposed. These two kinds of smooth temporal blurring of the space-time boundaries of an individual aim to remove any gait information.

Gestures are defined as the movement of a body part (fingers, hands, arms, head, or face) or a whole body that is made with or without the intension meaning something [171, 172]. For example, the expressive and meaningful motion of fingers or hands conveys meaningful information to another human, or it can be used for interacting with a real or virtual environment (virtual reality, augmented reality).

The fact that gestures vary between individuals can be exploited for person recognition [173, 174]. From the gesture-recognition point of view there is a problem because gestures vary for the same individuals at different instances. The approaches to the tracking, analysis and recognition of gestures in video [175] enable the effective interaction with the environment, but can also be used for people verification or identification.

To date, there have only been a few attempts to develop biometric verification systems based on *hand-gesture recognition* [173, 174], [176, 177].

As far as we know, there has been no research into the problem of *hand gesture de-identification*. The problem of gesture de-identification in video surveillance is similar to the problem of gait de-identification and can be solved by approaches similar to those used for gait.

7. De-identification of soft biometric identifiers

Soft biometric identifiers are physical, behavioural or adhered human characteristics of the person that provide some information about the person, but lack the distinctiveness and permanence to sufficiently differentiate any two persons [42]. However, soft biometric identifiers, as ancillary information, can be combined by biometric identifiers to improve the overall recognition, particularly when recognition system is designed to work in accordance with the less constrained scenarios including recognition at a distance [178].

There are four main modalities of using soft biometric identifiers for:

i) person identification or verification based on the measured soft biometric identifiers [178],

ii) person identification or verification based on verbal descriptions of soft biometric identifiers [43],

iii) person identification or verification in a biometric system based on the fusion of soft biometric identifiers and physiological and/or behavioural biometric identifiers in order to ensure better accuracy of the recognition process [42],

iv) retrieval of large biometric databases [179, 180].

Regardless of the above-described modalities of using soft biometric identifiers, it is obvious that soft biometric identifiers, such as silhouette, gender, race, moles, tattoos, birthmarks, and scars, carry privacy-intrusive information about individuals, and have to be de-identified in a multimedia document.

7.1 Body silhouette de-identification

The *body silhouette* is an important soft biometric identifier and it can help the recognition process (on its own or in combination with other biometric identifiers, e.g. gait). In addition to recognition, body silhouettes are used for people re-identification, i.e. tracking people across multiple cameras with non-overlapping fields of view in surveillance applications [181].

To the best of our knowledge there are only a few papers on *body silhouette de-identification*. In [55, 56], the authors showed that the masking of a silhouette is relatively easy, through the use of dilatation or Gaussian blurring. The Gaussian blurring of the silhouette is also used for the de-identification of individuals in activity videos (Figure 9.) [182]. In [56], it has been shown that a combination of line integral convolution (LIC) and the exponential blurring of pixels of a voxel gives the best results for silhouette de-identification.

a) b)

Figure 9. De-identification of individuals in activity videos depicting: a) walking; b) jumping in place actions after 2D Gaussian filtering [182].

An approach to reversible body de-identification in video is based on distortion applied to the ROI which contains the silhouette of an individual by using transform-domain scrambling methods proposed in [94, 95] (see Section 5.2). Figure 10. illustrates the result of the body de-identification by the scrambling method described in [94].

Figure 10. Result of the body silhouette de-identification by the scrambling method described in [94].

An interesting approach to silhouette de-identification is described in [183], it involves replacing a person with another one from a dataset gallery.

7.2 Gender, age, race and ethnicity de-identification

In literature, there are many papers related to the automatic recognition of *gender, age, race and ethnicity*, but relatively little is done on their de-identification in multimedia content. Information about *gender*, *age*, *race* and *ethnicity* is usually obtained from facial images [184-188] and/or a speaker utterance [189], gait and silhouette [158], and silhouetted face profiles [190]. In [56], the authors have mentioned that the masking of race and gender is a difficult problem. However, they agreed that it is possible to mask skin colour (which is closely related to race) using different colour transformations at the price of destroying the naturalness of the de-identified videos.

7.3 Scars, marks and tattoos de-identification

Scars, marks and tattoos (SMT) are imprints on skin that provide more discriminative information than age, height, gender, and race to identify a person [191]. In [192], the authors have showed that

facial marks, such as freckles, moles, scars and pockmarks can improve automatic face recognition and retrieval performance. For example, the experimental face-recognition system based on a combination of Active Appearance Models (AAMs) to locate and segment a facial image on eyes, nose, and mouth regions, and Laplacian-of-Gaussian (LoG) and morphological operators to detect facial marks, has improved the rank-1 identification accuracy of a state-of-the-art face recognition system from 92.96% to 93.90% on the FERET database and from 91.88% to 93.14% on the Mugshot database.

A methodology for detecting SMT found in unconstrained imagery normally encountered in forensics scenarios is described in [193]. As far as we know, there are no published papers related to de-identification of scars and marks.

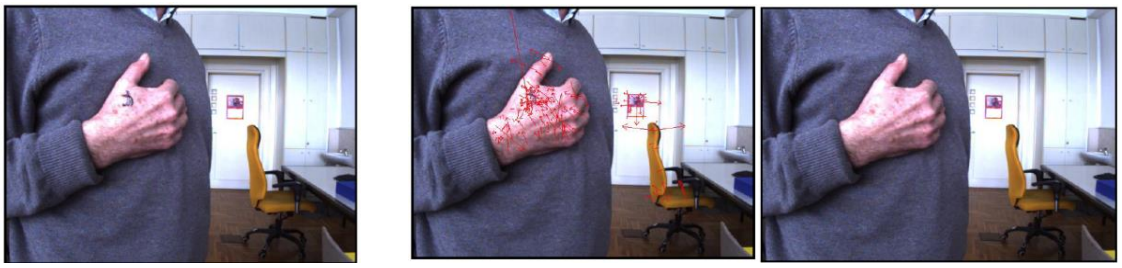
Tattoos are not only popular in particular groups, such as motorcyclists, sailors, and members of criminal gangs, they have become very popular in the wider population. In fact, about 24 percent of people aged from 18 to 50 in the USA have at least one tattoo, and this number is increasing [194].

Tattoos are primarily used for Content-based Image Retrieval (CBIR) in law-enforcement applications [195, 196], but based on the visual appearance of tattoos and their location on a body [194], they can be used for person recognition, as well as for suspect and victim identification in forensics.

The main features used for tattoo recognition are Scale Invariant Feature Transform (SIFT) features [191], [193], active contours and so-called *glocal* features – local features that contain global information regarding colour and edge orientation [197].

There are no published papers related to SMT de-identification, except [198]. The experimental system for tattoo localization and de-identification for privacy protection [198] was intended to be used for still images, but it was also tested for videos. The system consists of the following modules: skin and ROI detection, feature extraction, tattoo database, matching, tattoo detection, skin swapping, and quality evaluation. An image or a sequence of frames obtained by a colour camera is an input to the skin and ROI detection module. Uncovered body parts like the head, neck, hands, legs or torso are detected in two phases. In the first phase, skin-colour cluster boundaries are obtained by a pixel-based method through a series of decision rules in the RGB colour space. In the second phase, geometrical

constraints are used to eliminate skin-like colour regions that do not belong to the uncovered body-part areas. The SIFT features are extracted from a ROI in the feature-extraction module. The SIFT features are matched with template SIFT features from the tattoo database. Experimentally, 24 tattoos with at least two tattoos from each of the eight classes of tattoos labelled in the ANSI/NIST-ITL 1-2000 standard are used. Each tattoo in the tattoo database has an average of 56 template SIFT features, so the tattoo database consists of 1338 SIFT features. The de-identification process is performed in the skin-swapping module in such a way that the original tattoo's region is replaced by pixels from a surrounding, non-tattoo region. After replacement, a median filter is applied to the de-identified area. With this procedure, the authors try to hide the tattoo location and its visual appearance, and preserve the naturalness of the de-identified image (Figure 11.).



a) b) c)

Figure 11. Tattoo de-identification; a) An example of a still image obtained by a colour camera; b) Extracted SIFT features; c) De-identified tattoo still frame [198].

The experiments have shown that tattoo localization based on SIFT features gave satisfactory results in well-controlled conditions, such as lighting, high tattoo resolution, and no motion blur. For tattoos with a low-quality visual appearance, the SIFT features have to be combined with some region segmentation based on a combination of colour, gradient and/or texture methods. For surveillance applications, by using skin- and tattoo-area tracking based on a spatial and temporal correspondence between the frames, tattoo detection, localization and de-identification can be improved.

8. Discussion

In spite of the huge efforts of various academic research groups, institutions and companies, research in the field of de-identification and multimodal de-identification in multimedia content is still in its infancy. Relatively little has been done in the field of de-identification of non-biometric identifiers, except in the field of text and license plate de-identification. To avoid "pair-wise constraint" identification [54] and the classification of individuals in categories (which can be treated as privacy invasive), additional efforts have to be made in the field of dressing style and hairstyle de-identification (initial and pioneering efforts have been made only to conceal hairstyles and hair colour [199]). The problem of selectively concealing or removing the context-sensitive information or objects from the environment which can be used to reveal the identity of a person is still open. This could be solved in the near future by using a knowledge-based approach for modelling a specific environment and situation to detect additional ROIs and to obscure them.

At first glance, it looks as though the problem of license plate de-identification has been solved in web services like Google Street View and EveryScape, but the main problem is plate detection in video footages. According to some recent reports [57], undetected license plates amount to between 4% and 6%. The percentage is too high, so it is difficult to claim that de-identification of license plates has been successful. It is important to stress that computer vision techniques for the detection and tracing of object(s) of interest have to be improved and made reliable and robust. Privacy protection in web services like Google Street View and EveryScape is a typical example of the open problem of multimodal de-identification where multiple ROIs (e.g. faces, body silhouettes, license plates) have to be detected and obscured. For example, Google reports that its completely automatic system is able to blur 89% of faces, which means that many faces remain unblurred in Google Street View video footages [57].

In the field of face de-identification in still images, irreversible naive methods such as "blurring" and "pixelation" of the image region occupied by the face may protect the identity from the human observer, but these naive methods may be subverted by a machine applying so-called parrot recognition.

The irreversible de-identification methods referred to as *k-Same*, *k-Same-Select* and *Model-based k-Same* algorithms for face de-identification guarantee theoretical probable privacy (1/*k*), where *k* is the number of the closest face images to the raw face image in the person specific set. Reversible privacy-preserving methods for photo sharing applications, and morphing-based visual privacy protection, as well as scrambling-based methods, are convenient for privacy protection in social networks and photo hosting platforms. The state-of-the-art of face de-identification methods in still images enables a balance between privacy and naturalness, and simultaneously offers preservation of data utility (e.g. facial expression, age).

De-identification of the face in video surveillance systems is far from a complete solution. The problem lies not in the de-identification of ROIs, but in computer vision algorithms for the detection and localization of face(s) in video sequences. Despite recently intensive research in computer vision, numerous problems still remain to be solved in automatic face detection. These include issues such as the detection of the face under different illumination conditions, bad lighting conditions, different head positions, the presence of structural components (e.g., glasses, sunglasses, beards, moustaches), and occlusions. The unsolved problems are the detection of faces in crowd scenes and real-time de-identification. Privacy might be compromised in video sequences if the face detection algorithm fails in a single frame, so one of the directions of research is the development of robust and effective algorithms for privacy protection that can efficiently cope with situations when computer vision algorithms fail [38], [200].

De-identification in drone-based surveillance systems deserves special attention due to specific problems which are, in a computer vision sense, very close to Moving-Camera-Moving Object (MCMO) problems and different scenarios in comparison with "classic" CCTV surveillance. There are open problems in the detection of several ROIs (face, body silhouette, accessories, different positions and sizes) in dynamic scenes. Due to the complex problem of de-identification in drone-based surveillance systems, it is expected that the Privacy-by-Design approach has to be applied together with strict laws regarding the use of drones.

The de-identification of fingerprint still images is important in two respects: (i) privacy protection of the fingerprint as a biometric template in authentication systems; (ii) hidden privacy sensitive information (e.g. gender, ethnicity, health status) which can be revealed from the fingerprint pattern. Regarding the first aspect, there are already standards and architectures for biometric template protection. For the de-identification of other privacy sensitive information, different de-identification methods based on privacy filters or generating syntactic fingerprints can be used. The same methods employed for fingerprint template protection can be used for iris template protection in authentication systems. In the near future, we can expect surveillance systems capable of acquiring an iris image at a distance of more than 30 meters and performing the identification of an individual. There is therefore a need for iris de-identification. Pioneering research work in this direction based on scrambling an eye region has been conducted. The Iris at a Distance systems are also capable of acquiring a face, which leads to multimodal de-identification.

Due to the development of relatively low-cost, high-resolution, video cameras and telescopic equipment, we can expect ear-based recognition and tracking in semi- or non-controlled outdoors conditions. This will lead to the need for research and development of ear de-identification methods in order to protect the privacy of individuals. Most ear-recognition systems use the combination of a profile face and ear detection. Therefore, in the near future, ear de-identification will be a multimodal de-identification problem – the face and the ear have to be de-identified simultaneously.

There are several challenges in the field of online voice or speaker de-identification, such as de-identification in an environment with background noise, voice de-identification in situations where there are multiple individuals speaking simultaneously, which leads to crosstalk and overlapped speech. Additional efforts have to be made to develop more sophisticated voice de-identification systems with "personalized" multi-target voices and the preservation of the emotional expression of a speaker.

Approaches to gait and gesture de-identification are mainly based on scrambling techniques and the temporal blurring of the space-time boundaries of an individual. The main problem with gait and gesture de-identification in a video-surveillance system (which may be feasible in the near future) is

how to obscure the characteristics of an individual's movement and/or walking patterns, and at the same time preserve the usability and naturalness of the de-identified video. As far as we know, there are no published research reports on gesture de-identification.

In spite of the fact that soft biometric identifiers do not offer enough distinctive information to differentiate any two individuals, certain types of these identifiers (e.g. SMT, body silhouette, gender, age, race, birthmarks) carry private, sensitive and intrusive information on individuals, and therefore should be hidden or removed from multimedia content.

De-identification of soft biometric identifiers, such as the body silhouette, is based on naive privacy filters, reversible filters based on scrambling methods, or replacing a person with another one from a dataset gallery. The precondition for successful body silhouette de-identification is foreground (i.e. body silhouette) detection in videos. But, due to complex environments, non-stationary background motion, illumination variation, and camera vibration, detection is still far from perfect. In addition, the problem of masking the temporal variation of a body silhouette in such a way as to preserve the naturalness of de-identified videos remains unresolved.

The masking of soft biometric identifiers such as race, ethnicity and gender in video surveillance applications, is a difficult problem. Experts agree that it is possible to mask these identifiers, but at the cost of destroying the naturalness of the de-identified videos.

Preliminary research has been carried out in the field of tattoo de-identification in still images, but there are many unsolved problems: the localization of tattoos in the images of complex scenes, the localization of tattoos with a low-quality visual appearance and images taken under different angles of view.

Due to recent advances in multi-sensor acquisition and recording devices and remote surveillance systems, there is a need for the research and development of multimodal de-identification methods that simultaneously hide, remove or substitute different types of personal identifiers from multimedia content. A solution to the problem of multimodal de-identification still remains a major challenge.

Important aspects of de-identification are metrics in measuring privacy protection in multimedia content, the utility or intelligibility and naturalness or/and pleasantness of the de-identified data, as

well as the evaluation protocol [201]. There is not yet a common framework for the evaluation and assessment of these components in de-identified multimedia contents. Researchers are primarily focusing on the evaluation of privacy protection, intelligibility, pleasantness and the trade-off between privacy protection and utility/intelligibility for privacy filters applied on face regions in images and video sequences (FERET database, PEViD-HD and PEViD-UHD datasets [76, 92]). The evaluation of privacy protection and the trade-off between privacy protection and utility/intelligibility are usually performed by objective methods (PCA-, LDA- and LBP-based automatic face recognition) and subjective evaluation [95], [202] based on crowdsourcing [107], or by experts (video-analytics technology and privacy protection solution developers, or law enforcement personnel). Ongoing research activities regarding privacy protection and its evaluation in surveillance systems are presented in MediaEval workshops, established as an independent benchmarking initiative in 2010 (http://www.multimediaeval.org/).

The assessment of the de-identification of behavioural biometric identifiers is mainly devoted to privacy protection and to the intelligibility of de-identified speech [203].

Due to the social, legal and political importance of privacy protection, de-identification also requires a platform for studies of the legal, ethical and social aspects of de- and re-identification in multimedia content and social network sites, as well as the strong cooperation of experts in the technical and social sciences.

9. Conclusion

Privacy is one of the most important social and political issues in any free society. In our networked society, which is characterized by technologies and services such as internet, wireless communication, social networks, biometrics, multimedia, big data, data-mining, and audio and video surveillance, and drone-based surveillance, the problem of the privacy protection of individuals has become a major challenge for experts from law, political, ethical and technical domains. De-identification – a process of concealing, removing or substituting personal identifiers in multimedia content – is a method for protecting privacy. In this paper, we try to give an up-to-date review of de-identification methods for privacy protection in multimedia content. Based on proposed taxonomy of personal identifiers present

in multimedia documents we have presented de-identification of non-biometric, physiological, behavioural biometric identifiers, and soft-biometric identifiers. Regarding the trends in the surveillance technology, we have announced some new directions in the de-identification research: de-identification of iris and fingerprints captured at distance, gait and gesture de-identification, and multimodal de-identification which combines non-biometric, physiological, behavioural and soft-biometric identifiers. We have pointed out the problems of detecting and removing or hiding social and environmental privacy sensitive context in multimedia contents, as well as open problems of metrics and protocols for evaluation and assessment of privacy protection, intelligibility, and naturalness or/and pleasantness in de-identified multimedia contents.

This paper covers mainly the technical aspects of de-identification. But, due to the social, legal and political importance of privacy protection, we are aware that real solutions for de-identification, which are acceptable to both users and the law enforcement organisations in a networked society, will have to be based on the collective effort of experts from the fields of law, ethics, sociology and psychology as well as technical experts.

Acknowledgement

This work has been supported by the Croatian Science Foundation under project 6733 De-identification for Privacy Protection in Surveillance Systems (DePPSS). It is also the result of activities in COST Action IC1206 "De-identification for Privacy Protection in Multimedia Content".

References

[1] K. Abas, C. Porto, K. Obraczka, Wireless Smart Camera Networks for the Surveillance of Public Spaces, IEEE Computer, vol. 47, no. 5, (2014) 37 - 44.

[2] D. T. Raty, Survey on Contemporary Remote Surveillance Systems for Public Safety, IEEE Trans. on Systems, Man, and Cybernetics - Part C, vol. 40, no. 5, (2010) 493 - 515.

[3] D. J. Solove, Nothing to Hide, Yale University Press, New Haven & London, 2011.

[4] A. Cavallaro, Privacy in Video Surveillance, IEEE Signal Processing Magazine, vol. 24, no. 2, (2007) 168 - 169.

[5] A. Senior, Privacy Protection in a Video Surveillance System, in: A. Senior (Ed.), Protecting Privacy in Video Surveillance, Springer, Dordrecht, 2009, pp. 35 - 47.

[6] J. Angwin, Dragnet Nation, St. Martin's Press, New York, 2015.

[7] plato, http://plato.stanford.edu, (2009) (accessed 12.06.14).

[8] S. D. Warren, L. D. Brandeis, The Right to Privacy, Harvard Law Review, vol. IV, no. 5. (1890) http://readingnewengland.org/app/books/righttoprivacy/ (accessed 02.06. 14).

[9] A. F. Westin, Social and Political Dimensions of Privacy, The Society for the Psychological Study of Social Issues, (2003) 431 - 453.

[10] D. J. Solove, Understanding Privacy, Harvard University Press, Cambridge, 2008.

[11] P. Campisi, Security and Privacy in Biometrics: Towards a Holistic Approach, in: P. Campisi (Ed.), Privacy and Security in Biometrics, Springer, 2013, pp. 1 - 24.

[12] Peck v United Kingdom, Reference (2003) 36 EHRR 41; [2003] EMLR 287 Court European Court of Human Rights, Date of Judgment 28 Jan 2003.
http://www.worldlii.org/eu/cases/ECHR/2003/44.html (accessed 04.11.15).

[13] M. Krause, The Expanding Surveillance State: Why Colorado Should Scrap the Plan to Map Every Driver's Face and Should Ban Facial Recognition in Public Places, Issue Paper Number 8 -, 2001, (2001).
https://www.i2i.org/articles/8-2001.PDF (accessed 04.11.15).

[14] S. Rane, Standardization of Biometric Template Protection, IEEE MultiMedia, vol. 21, no. 4, (2014) 94 - 99.

[15] M. Savvides, B.V.K. Vijaya Kumar, P.K. Khosla, Cancelable Biometric Filters for Face Recognition, Proc. of the Int. Conf. on Pattern Recognition (ICPR), vol. 3, (2004) 922 - 925.

[16] N. K. Ratha, S. Chikkerur, J. H. Connell, R. M. Bolle, Generating Cancelable Fingerprint Templates, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 29, no. 4, (2007) 561 - 572.

[17] F. Hao, R. Anderson, J. Daugman, Combining Cryptography with Biometrics Effectively, IEEE Trans. on Computers, vol. 55, no. 9, (2006) 1081-1088.

[18] W. Xu, Q. He, Y. Li, T. Li, Cancelable Voiceprint Templates Based on Knowledge Signatures, Proc. of the Int. Symposium on Electronic Commerce and Security, (2008) 412 - 415.

[19] C. Rathgeb, A. Uhl, A Survey on Biometric Cryptosystems and Cancelable Biometrics, EURASIP Journal on Information Security 2011:3, (2011) 1 - 25.
http://jis.eurasipjournals.com/content/2011/1/3 (accessed 10.01.15)

[20] V. M. Patel, N. K. Ratha, R. Chellappa, Cancelable Biometrics: A Review, IEEE Signal Processing Magazine, vol. 32, no. 5, (2015) 54 - 65.

[21] T. E. Boult, PICO: Privacy through Invertible Cryptographic Obscuration, IEEE/NSF Proc. of the Workshop on Computer Vision for Interactive and Intelligent Environments, (2005) 27 - 38.

[22] A. J. Bharucha, A. J. London, D. Barnard, H. Wactlar, M. A. Dew, C. F. Reynolds III, Ethical Considerations in the Conduct of Electronic Surveillance Research, Journal of Law, Medicine & Ethics, (2006) 1 - 10.

[23] R. Gellman, Fair Information Practices: A Basic History, (2015) 1 - 33.
http://www.bobgellman.com/rg-docs/rg-FIPShistory.pdf (accessed 05.06.15).
[24] Directive 95/46/EC of the European Parliament and of the Council of 24 October, (1995) /http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTML (accessed 10.01.15).
[25] N. Robinson, H. Graux, M. Botterman, L. Valeri, Review of EU Data Protection Directive: Summary, (2009).
https://ico.org.uk/media/about-the-ico/documents/1042347/review-of-eu-dp-directive-summary.pdf (accessed 02.07.13).
[26] D. J. Solove, Privacy and Power: Computer Databases and Metaphors for Information Privacy, 53 Stan. L. Rev. 1393, (2001) 1393 - 1461.
[27] S. Hinde, Privacy Legislation: A Comparison of the US and European Approaches, Computers & Security, vol. 22, no. 5, (2003) 378 - 387.
[28] B. L. Movius, N. Krup, U.S. and EU Privacy Policy: Comparison of Regulatory Approaches, Int. Journal of Communication 3, (2009) 169 - 187.
[29] F. Boehm, A Comparison Between US and EU Data Protection Legislation for Law Enforcement Purposes, European Parliament, (2015) 1 - 81.
http://www.europarl.europa.eu/RegData/etudes/STUD/2015/536459/IPOL_STU%282015%29536459_EN.pdf (accessed 17.11.15).
[30] Reform of EU Data Protection Rules, (2016).
http://ec.europa.eu/justice/data-protection/reform/index_en.htm (accessed 04.05.16).
[31] Common Criteria for Information Technology Security Evaluation, (1999).
https://www.niap-ccevs.org/Documents_and_Guidance/cc_docs/cc_users_guide.pdf (accessed 08.07.13).
[32] G. W. van Blarkom, J. J. Borking, J. G. E. Olk, (Eds.), Handbook of Privacy and Privacy-Enhancing Technologies, College Bescherming Persoonsgegevens, The Hague, 2003.
[33] P. Langendörfer, M. Maaser, K. Piotrowski, S. Peter, Privacy Enhancing Techniques: A Survey and Classification, (2008) 1 - 18.
http://www.ics.uci.edu/~steffenp/files/langendoerfer2008privacy.pdf (accessed 29.11.15).
[34] A. Cavoukian, Privacy by Design, (2010) 1 - 2.
https://www.privacybydesign.ca (accessed 07.11.15).
[35] M. Mrityunjay, P. J. Narayanan, The De-Identification Camera, Proc. of the 3rd National Conf. on Computer Vision, Pattern Recognition, Image Processing and Graphics, (2011) 192 - 195.
[36] G. S. Nelson, Practical Implications of Sharing Data: A Primer on Data Privacy, Anonymization, and De-Identification, (2015) 1 - 23.

http://thotwave.com/wp-content/uploads/2015/09/data_sharing_privacy_anonymization_and_de-identification_rev_13.pdf (accessed 25.06. 15).

[37] IC1206 COST Action, Memorandum of Understanding (MoU) (2013) http://w3.cost.eu/fileadmin/domain_files/ICT/Action_IC1206/mou/IC1206-e.pdf (accessed 03.12.15).

[38] J. R. Padilla-Lopez, A. A. Chaaraoui, F. Florez-Revuelta, Visual Privacy Protection Methods: A Survey, Expert Systems with Applications, 42 (9) (2015) 4177 - 4195.

[39] V. Bhagwan, T. Grandison, C. Maltzahn, Recommendation-based De-Identification, A Practical Systems Approach Towards De-identification of Unstructured Text in Healthcare, (2012) 155 - 162. http://www.almaden.ibm.com/cs/people/tgrandison/SPE2012-ReDid.pdf (accessed 15.07.14).

[40] HIPAA, (2014), http://www.hhs.gov/ocr/hipaa (accessed 15.12.14).

[41] A. Dantcheva, C. Velardo, A. D'Angelo, J - L. Dugelay, Bag of Soft Biometrics for Person Identification, Multimedia Tools and Applications, vol. 51, no. 2, (2011) 739 - 777.

[42] A. K. Jain, S. C. Dass, K. Nandakumar, Soft Biometric Traits for Personal Recognition Systems, Proc. of the Int. Conf. on Biometric Authentication, (2004) 731 - 738.

[43] D. A. Reid, M. S. Nixon S. V. Stevenage, Soft Biometrics; Human Identification Using Comparative Descriptors, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 36, no. 6, (2014) 1216 - 1228.

[44] I. Neamatullah, M. M. Douglass, L. H. Lehman, A. Reisner, M. Viallarroel, W.J. Long, et al., Automated De-identification of Free-text Medical Records, BMC Medical Informatics and Decision Making, vol. 8, no. 32 (2008) 1 - 73.

[45] L. Sweeney, Replacing Personally-identifying Information in Medical Records, the Scrub system, Proc. of the AMIA Annual Fall Symposium, (1996) 333 - 337.

[46] L. Sweeney, Computational Disclosure Control: A Primer on Data Privacy Protection, Ph.D. Thesis, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, (2001) 216 pages.

[47] R. Fraser, D. Willison, Tools for De-Identification of Personal Health Information, Canada Health Infoway, (2009) 1 - 40.
http://www.ehealthinformation.ca/wp-content/uploads/2014/08/deid.pdf (accessed 05.05.16).

[48] Ö. Uzuner, Y. Luo, P. Szolovits, Evaluating the State-of-the-Art in Automatic De-identification, Journal of the American Medical Informatics Association, vol. 14, no. 5 (2007) 550 - 563.

[49] L. S. Garfinkel, NISTIR 8053 De-Identification of Personal Information, (2015) 1 - 54. http://dx.doi.org/10.6028/NIST.IR.8053 (accessed 17.12.15).

[50] N. Kumar, A. C. Berg, P. N. Belhumeur, S. K. Nayar, Attribute and Simile Classifiers for Face Verification, Proc. of the 12th IEEE Int. Conf. on Computer Vision (ICCV), (2009) 365 - 372.

[51] M. Feng, Z. Kun, S. Nong, A Classified Method of Human Hair for Hair Sketching, Proc. of the Congress on Image and Signal Processing (CISP), vol. 4, (2008) 109 - 114.

[52] M. Yang, K. Yu, Real-time Clothing Recognition in Surveillance Videos, Proc. of the 18th IEEE Int. Conf. on Image Processing (ICIP), (2011) 2937 - 2940.

[53] M. Rahman, H. S. Kim, S. Ishikawa, Solving a Dress Problem for a Human Model Recognition, Proc. of the Society of Instrument & Control Engineers (SICE), (2001) 210 - 213.

[54] D. Chen, Y. Chang, R. Yan, J. Yang, Protecting Personal Identification in Video, in: A. Senior (Ed.), Protecting Privacy in Video Surveillance, Springer, Dordrecht, 2009, pp. 115 - 128.

[55] P. Agrawal, De-identification for Privacy Protection in Surveillance Videos, Master of Science Thesis, Center for Visual Information Technology International Institute of Information Technology Hyderabad, (2010) 49 pages.

[56] P. Agrawal, P. J. Narayanan, Person De-Identification in Videos, IEEE Trans. on Circuits and Systems for Video Technology, vol. 21, no. 3, (2011) 299 - 310.

[57] A. Frome, G. Cheung, A. Abdulkader, M. Zennaro, B. Wu, A. Bissacco, H. Neven, L.Vincent, Large-scale Privacy Protection in Google Street View, Proc. of the IEEE 12th Int. Conf. on Computer Vision (ICCV), (2009) 2373 - 2380.

[58] H. Schneiderman, T. Kanade, A Statistical Method for 3D Object Detection Applied to Faces and Cars, Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), vol. I, (2001) 746 - 751.

[59] P. Viola, M. J. Jones, Robust Real-Time Face Detection, Int. Journal of Computer Vision 57(2), (2004) 137 - 154.

[60] L. Du, H. Ling, Preservative License Plate De-identification for Privacy Protection, Proc. of the Int. Conf. on Document Analysis and Recognition (ICDAR), (2011) 468 - 472.

[61] S. Z. Li, A. K. Jain (Eds.), Handbook of Face Recognition, Springer, New York, 2005.

[62] M. Boyle, C. Edwards, S. Greenberg, The Effects of Filtered Video on Awareness and Privacy, Proc. of the ACM Conf. on Computer Supported Cooperative Work, Philadelphia, (2000) 1 - 10.

[63] C. Neustaedter, S. Greenberg, M. Boyle, Blur Filtration Fails to Preserve Privacy for Home - Based Video Conferencing, ACM Trans. on Computer Human Interaction, vol. 13, issue 1, (2006) 1 - 36.

[64] cmu-pie, http://www.computervisiononline.com/dataset/cmu-pie-database (accessed 02. 11.12).

[65] S. Ribaric, N. Pavesic, An Overview of Face De-identification in Still Images and Videos, Proc. of the 11th IEEE Int. Conf. and Workshops on Automatic Face and Gesture Recognition (FG), (2015) 1 - 6.

[66] R. Gross, L. Sweeney, F. de la Torre, S. Baker, Model-Based Face De-Identification, Proc. of the Conf. on Computer Vision and Pattern Recognition Workshop (CVPRW), (2006) 161 - 169.

[67] P. J. Phillips, Privacy Operating Characteristic for Privacy Protection in Surveillance Applications, in: T. Kanade, A. K. Jain, and N. K. Ratha (Eds.), Audio- and Video-Based Biometric

Person Authentication, Lecture Notes in Computer Science (LNCS), vol. 3546, Springer, 2005, pp. 869 - 878.

[68] E. Newton, L. Sweeney, B. Malin, Preserving Privacy by De-identifying Facial Images, IEEE Trans. on Knowledge and Data Engineering, vol. 17, no. 2, (2005) 232 - 243.

[69] R. Gross, E. Airoldi, B. Malin, L. Sweeney, Integrating Utility into Face De-identification, in: G. Danezis and D. Martin (Eds.), PET - Privacy Enhancing Technologies 2005, Lecture Notes in Computer Science (LNCS), vol. 3856, Springer, 2006, pp. 227 - 242.

[70] R. Gross, L. Sweeney, J. Cohn, F. de la Torre, S. Baker, Face De-Identification, in: A. Senior (Ed.), Protecting Privacy in Video Surveillance, Springer, Dordrecht, 2009, pp. 129 - 146.

[71] L. Sweeney, k-Anonymity: A Model for Protecting Privacy, International Journal on Uncertainty, Fuzziness and Knowledge-Based Systems, vol. 10, no. 5, (2002) 557 - 570.

[72] T. Cootes, G. Edwards, C. Taylor, Active Appearance Models, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 23, no. 6 (2001) 681 - 685.

[73] L. Meng, Z. Sun, A. Ariyaeeinia, K. L. Bennett, Retaining Expressions on De-identified Faces, Proc. of the Special Session on Biometrics, Forensics, De-identification and Privacy Protection (BiForD), (2014) 27 - 32.

[74] L. Meng, Z. Sun, Face De-identification with Perfect Privacy Protection, ibid, (2014) 9 - 14.

[75] L. Yuan, P. Korshunov, T. Ebrahimi, Privacy-Preserving Photo Sharing Based on a Secure JPEG, Proc. of the 3rd Int. Workshop on Security and Privacy in Big Data Security, (2015) 185 - 190.

[76] P. Korshunov, T. Ebrahimi, Using Face Morphing to Protect Privacy, Proc. of the IEEE Int. Conf. on Advanced Video and Signal-based Surveillance, (2013) 208 - 213.

[77] P. Korshunov, T. Ebrahimi, Using Warping for Privacy Protection in Video Surveillance, Proc. of the 18th Int. Conf. on Digital Signal Processing, (2013) 1 - 6

[78] E. Hjelmas, B. K. Low, Face Detection: A Survey, Computer Vision and Image Understanding 83, (2001) 236 - 274.

[79] M.-H. Yang, D. J. Kriegman, N. Ahuja, Detecting Faces in Images: A Survey, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 24, no. 1, (2002) 34 - 58.

[80] H. A. Rowley, S. Baluja, T. Kanade, Neural Network-Based Face Detection, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 20, no. 1, (1998) 23 - 38.

[81] K. Levi, Y. Weiss, Learning Object Detection from a Small Number of Examples: the Importance of Good Features, Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR), vol. 2, (2004) II-53 - II-60.

[82] N. Dalal, B. Triggs, Histograms of Oriented Gradients for Human Detection, Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), (2005) 886 - 893.

[83] X. Zhu, D. Ramanan, Face Detection, Pose Estimation, and Landmark Localization in the Wild, Proc. of the Conf. on Computer Vision and Pattern Recognition (CVPR), (2012) 2879 - 2886.

[84] P. Dollár, Z. Tu, P. Perona, Integral Channel Features, Proc. of the British Machine Vision Conf. (BMVC), (2009) 1- 11.

[85] P. Dollár, S. Belongie, P. Perona, The Fastest Pedestrian Detector in the West, Proc. of the British Machine Vision Conf. (BMVC), (2010) 1 - 11.

[86] P. Dollár, R. Appel, S. Belongie, P. Per, Fast Feature Pyramids for Object Detection, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 36, no. 8, (2014) 1532 - 1545.

[87] J. Yang, A. Waibel, A Real-Time Face Tracker, Proc. of the 3rd IEEE Workshop on Applications of Computer Vision (WACV), (1996) 142 - 147.

[88] L. Xu, J. Li, K. Wang, Real-time and Multi-View Face Tracking on Mobile Platform, Proc. of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), (2011) 1485 - 1488.

[89] W. Chuan-xu, L. Zuo-yong, A New Face Tracking Algorithm Based on Local Binary Pattern and Skin Color Information, Proc. of the Int. Symposium on Computer Science and Computational Technology, (2008) 657 - 660.

[90] W-P. Choi, K-M. Lam, An Effective Shape-Texture Weighted Algorithm for Multi-view Face Tracking in Videos, Proc. of the Congress on Image and Signal Processing (CISP), (2008) 156 - 160.

[91] D. Comaniciu, V. Ramesh, P. Meer, Real-Time Tracking of Non-Rigid Objects using Mean Shift, Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), vol. 2, (2000) 142 - 149.

[92] P. Korshunov, T. Ebrahimi, Towards Optimal Distortion-based Visual Privacy Filter, Proc. of the IEEE Int. Conf. on Image Processing (ICIP), (2014) 6051 - 6055.

[93] A. Erdely, T. Barat, P. Valet, T. Winkler, B. Rinner, Adaptive Cartooning for Privacy Protection in Camera Networks, Proc. of the 11th IEEE Int. Conf. on Advanced Video and Signal Based Surveillance (AVSS), (2014) 26 - 29.

[94] F. Dufaux, T. Ebrahimi, Scrambling for Privacy Protection in Video Surveillance Systems, IEEE Trans. on Circuits and Systems for Video Technology, vol. 18, no. 8, (2008) 1168 - 1174.

[95] F. Dufaux, T. Ebrahimi, A Framework for the Validation of Privacy Protection Solutions in Video Surveillance, Proc. of the IEEE Int. Conf. on Multimedia and Expo (ICME), (2010) 66 - 71.

[96] J. Schiff, M. Meingast, D. K. Mulligan, S. Sastry, K. Goldberg, Respectful Cameras: Detecting Visual Markers in Real-time to Address Privacy Concerns, in: A. Senior (Ed.), Protecting Privacy in Video Surveillance, Springer, Dordrecht, 2009, pp. 65 – 89.

[97] A. Chattopadhyay, T.E. Boult, PrivacyCam: a Privacy Preserving Camera Using uCLinux on the Blackfin DSP, Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), (2007) 1 - 8.

[98] T. Winkler, B. Rinner, TrustCAM: Security and Privacy-Protection for an Embedded Smart Camera based on Trusted Computing, Proc. of the 7th IEEE Int. Conf. on Advanced Video and Signal Based Surveillance (AVSS), (2010) 593 - 600.

[99] B. Samarzija, S. Ribaric, An Approach to the De-Identification of Faces in Different Poses, Proc. of the Special Session on Biometrics, Forensics, De-identification and Privacy Protection (BiForD), (2014) 21 - 26.

[100] O. Marzocchi, Privacy and Data Protection Implications of the Civil Use of Drones, PE 519.221, (2015) 1 - 34.
http://www.europarl.europa.eu/RegData/etudes/IDAN/2015/519221/IPOL_IDA%282015%29519221_EN.pdf (accessed 13.11.15).

[101] R. Clarke, The Regulation of the Impact of Civilian Drones on Behavioural Privacy, Computer Law & Security Review 30, (2014) 286 - 305.

[102] A. Cavoukian, Surveillance, Then and Now: Securing Privacy in Public Spaces, Technical Report Information and Privacy Commissioner of Ontario, (2013) 1- 64.

[103] R. L. Wilson, Ethical Issues with use of Drone Aircraft, Proc. of the IEEE Int. Symposium on Ethics in Science, Technology and Engineering, (2014) 1- 4.

[104] J. Villasenor, Observations from above: Unmanned Aircraft Systems and Privacy, Harvard Journal of Law & Public Policy, vol. 36, (2013) 458 - 517.

[105] A. Cavoukian, Privacy and Drones: Unmanned Aerial Vehicles, (2012) 1 - 30
https://www.ipc.on.ca/images/Resources/pbd-drones.pdf (accessed 24.11.15).

[106] M. Bonetto, P. Korshunov, G. Ramponi, T. Ebrahimi, Privacy in Mini-drone Based Video Surveillance, Proc. of the 11th IEEE Int. Conf. and Workshops on Automatic Face and Gesture Recognition (FG), vol. 4, (2015) 1 - 6.

[107] P. Korshunov, S. Cai, and T. Ebrahimi, Crowdsourcing Approach for Evaluation of Privacy Filters in Video Surveillance, Proc. of the ACM Multimedia 2012 Workshop on Crowdsourcing for Multimedia (CrowdMM), (2012) 35 - 40.

[108] S. Çiftçi, P. Korshunov, A. O. Akyüz, T. Ebrahimi, MediaEval 2015 Drone Protect Task: Privacy Protection in Surveillance Systems Using False Coloring, Proc. of the MediaEval Workshop, (2015) 1 - 2.

[109] D. Maltoni, D. Malo, A. K. Jain, S. Prabhakar, (Eds.), Handbook of Fingerprint Recognition, Springer, New York, 2003.

[110] MarketsandMarkets, Next Generation Biometric Market-Forecasts & Analysis 2014 - 2020, (2014), www.marketsandmarkets.com (accessed 15.11.15).

[111] technologyreview, (2015)
https://www.technologyreview.com/s/422400/fingerprints-go-the-distance (accessed 07.12.15).

[112] A. Badawi, M. Mahfouz, R. Tadross, R. Jantz, Fingerprint - Based Gender Classification, Proc. of the Int. Conf. on Image Processing, Computer Vision, and Pattern Recognition (IPCV), (2006) 41 - 46.

[113] D. Dessimoz, J. Richiardi, C. Champod, A. Drygajlo, Multimodal Biometrics for Identity,

State-of-the- Art, Research Report PFS 341 - 08.05, (Version 2.0), (2009)156 pages.

[114] A. Barbeau, J-G. Trudeau, C. Coiteux, Fingerprint Patterns in Huntington's Chorea and Parkinson's Disease, Canad. Med. Ass. J., vol. 92, (1965) 514 - 515.

[115] H. J. Weinreb, Fingerprint Patterns in Alzheimer's Disease, Arch Neurol. 42(1) (1985) 50 - 54.

[116] N. K. Ratha, J. H. Connell, R. M. Bolle, Enhancing Security and Privacy in Biometrics-based Authentication Systems, IBM Systems Journal, vol. 40, no. 3, (2001) 614 - 634.

[117] A. K. Jain, U. Uludag, Hiding Biometric Data, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 25, no. 11, (2003) 1494 - 1498.

[118] L. Sheng, A. C. Kot, Privacy Protection of Fingerprint Database Using Lossless Data Hiding, Proc. of the IEEE Int. Conf. on Multimedia and Expo (ICME), (2010) 1293 - 1298.

[119] L. Sheng, A. C. Kot, Fingerprint Combination for Privacy Protection, IEEE Trans. on Information Forensics and Security, vol. 8, no. 2, (2013) 350 - 360.

[120] A. Ross, De-identifying Biometric Images for Enhancing Privacy and Security, (2014) 1 - 27. http://biometrics.nist.gov/cs.../08_tuesday_ross_VC-MIXING_IBPC2014.pdf (accessed 17.12.15).

[121] L. Lugini, E. Marasco, B., Cukic, J. Dawson, Removing Gender Signature from Fingerprints, Proc. of the Special Session on Biometrics, Forensics, De-identifications and Privacy Protection (BiForD), (2014) 63 - 67.

[122] J. Daugman, How Iris Recognition Works, IEEE Trans. on Circuits and Systems for Video Technology, vol. 14 , issue 1, (2004) 21 - 31.

[123] R. P. Wildes, Iris Recognition: An Emerging Biometric Technology, Proc. of the IEEE, vol. 85, no. 9, (1997) 1348 - 1363.

[124] A. M. George, C.A.D. Durai, A Survey on Prominent Iris Recognition Systems, Proc. of the Int. Conf. on Information Communication and Embedded Systems (ICICES), (2013) 191 - 195.

[125] J. R. Matey, O. Naroditsky, K. Hanna, R. Kolczynski, D. J. Loiacono, S. Mangru, M. Tinker, T. M. Zappia, W.Y. Zhao, Iris on the Move: Acquisition of Images for Iris Recognition in Less Constrained Environments, Proc. of the IEEE, vol. 94, no. 11, (2006) 1936 - 1947.

[126] morpho (2014).
http://www.morpho.com/en/media/20140311_iris-distance-power-behind-iris (accessed 23.07.14).

[127] C. Fancourt, L. Bogoni, K. Hanna, Y. Gua, R. Wildes, N. Takahashi, U. Jain, Iris Recognition at a Distance, AVBPA 2005, Lecture Notes in Computer Science (LNCS), vol. 3546, Springer, 2005, pp. 1 - 13.

[128] W. Dong, Z. Sun, T. Tan, X. Qiu, Self-adaptive Iris Image Acquisition System, Proc. of the SPIE, Biometric Technology for Human Identification, vol. 6944, (2008) 6 - 14.

[129] F. W. Wheeler, A. G. Amitha Perera, G. Abramovich, B. Yu, P. H. Tu, Stand-off Iris Recognition System, Proc. of the 2nd IEEE Int. Conf. on Biometrics: Theory, Applications and Systems (BTAS), (2008) 1 - 7.

[130] F. Bashir, P. Casaverde, D. Usher, M. Friedman, Eagle-Eyes: A System for Iris Recognition at a Distance, Proc. of the IEEE Conf. on Technologies for Homeland Security, (2008) 426 - 431.

[131] J. A. De Villar, R. W. Ives, J. R. Matey, Design and Implementation of a Long Range Iris Recognition System, Proc. of the Conf. Record of the 44th Asilomar Conf. on Signals, Systems and Computers (ASILOMAR), (2010) 1770 - 1773.

[132] R. Abiantun, M. Savvides, P. K. Khosla, Automatic Eye-level Height System for Face and Iris Recognition Systems, Proc. of the 4th IEEE Workshop on Automatic Identification Advanced Technologies (2005) 155 - 159.

[133] D. Lee, K. N. Plataniotis, A Novel Eye Region Based Privacy Protection Scheme, Proc. of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), (2012) 1845 - 1848.

[134] A. Abaza, A. Ross, C. Hebert, M. A. F. Harrison, M. S. Nixon, A Survey on Ear Biometrics, ACM Comput. Surv. 45, 2, Article 22, (2013) 1 - 35.

[135] A. Pflug, C. Busch, Ear Biometrics: A Survey of Detection, Feature Extraction and Recognition Methods, IET Biometrics, vol. 1, issue 2, (2012) 114 - 129.

[136] L. Yuan, Z-C. Mu, Ear Detection Based on Skin-Color and Contour Information, Proc. of the 6th Int. Conf. on Machine Learning and Cybernetics, (2007) 2213 - 2217.

[137] A. Kumar, M. Hanmandlu, M. Kuldeep, H. M. Gupta, Automatic Ear Detection for Online Biometric Applications, Proc. of the 3rd National Conf. on Computer Vision, Pattern Recognition, Image Processing and Graphics, (2011) 146 - 149.

[138] A. Abaza, C. Hebert, M. A. F. Harrison, Fast Learning Ear Detection for Real-time Surveillance, Proc. of the 4th IEEE Int. Conf. on Biometrics: Theory Applications and Systems (BTAS), (2010) 1 - 6.

[139] T. Kinnunen, H. Li, An Overview of Text-independent Speaker Recognition: From Features to Supervectors, Speech Communication, vol. 52, issue 1, (2010) 12 - 40.

[140] R. M. Bolle, J. H. Connell, S. Pankanti, N. K. Ratha, A. W. Senior, Guide to Biometrics, Springer, New York, 2004.

[141] J. P. Campbell, Speaker Recognition, in: A. K. Jain, R. M. Bolle, S. Pankanti (Eds.), Biometrics, Personal Identification in Networked Society, Kluwer, Dordrech, 1999. pp. 165 - 189.

[142] M. A. Pathak, B. Raj, Privacy-preserving Speaker Verification as Password Matching, Proc. of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), (2012) 1849 - 1852.

[143] Y. Stylianou, Voice Transformation: A Survey, Proc. of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), (2009) 3585 - 3588.

[144] L. B. Muda, M., I. Elamvazuthi, Voice Recognition Algorithms using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques, Journal of Computing, vol. 2, issue 3, (2010) 138 - 143.

[145] D. Sundermann, Voice Conversion: State-of-the-Art and Future Work, Fortschritte der Akustik, no. 31, issue 2, (2005) 1 - 2.

[146] M. Abe, S. Nakamura, K. Shikano, H. Kuwabara, Voice Conversion Through Vector Quantization, Proc. of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), (1988) 655 - 658.

[147] G. Upperman, M. Hutchinson, B. Van Osdol, J. Chen, Methods for Voice Conversion, (2014) 1-40.
http://ftpmirror.your.org/pub/misc/cd3wd/1006/Methods_for_Voice_Conversion_electr_physics_cnx_x10252_.pdf (accessed 14.12.14).

[148] D. Sundermann, A. Bonafonte, H. Ney, H. Hoge, A First Step Towards Text-Independent Voice Conversion, Proc. of the Int. Conf. on Spoken Language Processing (ICSLP), (2004) 1 - 4.

[149] D. Sundermann, H. Hoge, A. Bonafonte, H. Ney, J. Hirschberg, Text-Independent Cross-Language Voice Conversion, Proc. of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), vol.1, (2006) 1 - 4.

[150] A. Mouchtaris, J. Van Spiegel, P. Mueller, Non-Parallel Training for Voice Conversion by Maximum Likelihood Constrained Adoption, Proc. of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), vol. 1, (2004) I-1 - I-4.

[151] J. S. Chaudhari, Privacy Protection for Life-log System, University of Kentucky Master's Theses. Paper 491. (2007)
http://uknowledge.uky.edu/gradschool_theses/491 (accessed 06.12.14).

[152] J. S. Chaudhari, S.-C. S. Cheung, M. V. Venkatesh, Privacy Protection for Life-log System, IEEE Workshop on Signal Processing Applications for Public Security and Forensics (SAFE), (2007) 1 - 5.

[153] Q. Jin, A. R. Toth, T. Schultz, A. W. Black, Voice Converging: Speaker De-identification by Voice Transformation, Proc. of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), (2009) 3909 - 3912.

[154] Q. Jin, A. R. Toth, T. Schultz, A. W. Black, Speaker De-identification via Voice Transformation, Proc. of the IEEE Workshop on Automatic Speech Recognition & Understanding (ASRU), (2009) 529 - 533.

[155] M. Pobar, I. Ipsic, Online Speaker De-identification Using Voice Transformation, Proc. of the Special Session on Biometrics, Forensics, De-identification and Privacy Protection (BiForD), (2014) 33 - 36.

[156] T. Justin, V. Struc, S. Dobrisek, B. Vesnicer, I. Ipsic, F. Mihelic, Speaker De-identification using Diphone Recognition and Speech Synthesis, Proc. of the 11th IEEE Int. Conf. and Workshops on Automatic Face and Gesture Recognition (FG), vol. 4, (2015), 1 - 7.

[157] D. D. Zhang, Automated Biometrics, Technology and Systems, Kluwer Academic Publishers, New York, 2000.

[158] M. S. Nixon, J. N. Carter, D. Cunado, P. S. Huang, S.V. Stevenage, Automatic Gait Recognition, in: A. K. Jain, R. Bolle, and S. Pankanti, (Eds.), Biometrics, Personal Identification in Networked Society, Kluwer Academic Publishers, New York, 1999. pp. 231 - 249.

[159] J.-H. Yoo, D. Hwang, M. S. Nixon, Gender Classification in Human Gait Using Support Vector Machine, ACIVS 2005, Lecture Notes in Computer Science (LNCS), vol. 3708, Springer, 2005, pp. 138 - 145.

[160] L. Lee, W.E.L. Grimson, Gait Analysis for Recognition and Classification, Proc. of the IEEE Int. Conf. on Automatic Face and Gesture Recognition (FG), (2002) 148 - 155.

[161] Z. Zhang, M. Hu, Y. Wang, A Survey of Advances in Biometric Gait Recognition, CCBR 2011, Lecture Notes in Computer Science (LNCS), vol. 7098, Springer, 2011, pp. 150 - 158.

[162] N. V. Boulgouris, D. Hatzinakos, K. N. Plataniotis, Gait Recognition: A Challenging Signal Processing Technology for Biometric Identification, IEEE Signal Processing Magazine, vol. 11, (2005) 78 - 90.

[163] D. K. Wagg, M. S. Nixon, On Automated Model-based Extraction and Analysis of Gait, Proc. of the IEEE Int. Conf. on Automatic Face and Gesture Recognition (FG), (2004) 11 - 16.

[164] D. Cunado, M. S. Nixon, J. N. Carter, Automatic Extraction and Description of Human Gait Models for Recognition Purposes, Computer Vision and Image Understanding, vol. 90, issue 1, (2003) 1 - 41.

[165] R. T. Collins, R. Gross, J. Shi, Silhouette-based Human Identification from Body Shape and Gait, Proc. of the 5th IEEE Int. Conf. on Automatic Face and Gesture Recognition (FG), (2002) 351 - 356.

[166] D. Tao, X. Li, X. Wu, S. J. Maybank, General Tensor Discriminant Analysis and Gabor Features for Gait Recognition, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 29, no. 10, (2007) 1700 - 1715.

[167] G.V. Veres, L. Gordon, J.N. Carter, M.S. Nixon, What Image Information is Important in Silhouette-based Gait Recognition?, Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR), (2004) II-776 - II-782.

[168] L. Wang, T. Tan, H. Ning, W. Hu, Silhouette Analysis-Based Gait Recognition for Human Identification, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 25, no. 12, (2003) 1505 - 1518.

[169] S. Sarkar, P. J. Phillips, Z. Liu, I. R. Vega, P. Grother, K. W. Bowyer, The HumanID Gait Challenge Problem: Data Sets, Performance, and Analysis, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 27, no. 2, (2005) 162 - 177.

[170] N. Baaziz, N. Lolo, O. Padilla, F. Petngang, Security and Privacy Protection for Automated Video Surveillance, Proc. of the IEEE Int. Symposium on Signal Processing and Information Technology, (2007) 17 - 22.

[171] S. Mitra, T. Acharya, Gesture Recognition: A Survey, IEEE Trans. on Systems, Man and Cybernetics - PartC: Applications and Reviews, vol.37, no.3, (2007) 311 - 324.

[172] M. B. Abdallah, M. Kallel, M. S. Bouhlel, An Overview of Gesture Recognition, Proc. of the 6th Int. Conf. on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT), (2012) 20 - 24.

[173] N.D. Lentsoane, K. Kith, B.J. Van Wyk, M. A. Van Wyk, Identity Verification System Using Hand Gesture Information, Proc. of the 17th Int. Symposium of the Pattern Recognition Society of South Africa, (2006) 1 - 6.
http://www.prasa.org/proceedings/2006/prasa06-13.pdf (accessed 12.04. 16).

[174] N. D. Lentsoane, Identity Verification System Using Hand Gesture Information. Magister Technologiae: Electronic Engineering, Department Of Electrical Engineering. Faculty of Engineering, Tshwane University of Technology, (2007) 202 pages.

[175] S. Sclaroff, M. Betke, G. Kollios, Alon, Jonathan, V. Athitsos, Rui Li, J. Magee, Tai-Peng Tian, Tracking, Analysis, and Recognition of Human Gestures in Video, Proc. of the 8th Int. Conf. on Document Analysis and Recognition (ICDAR), vol.2, (2005) 806 - 810.

[176] S. Fong, Y. Zhuang, I. Fister, A Biometric Authentication Model Using Hand Gesture Images, BioMedical Engineering OnLine 12:111, (2013) 1 - 18.
https://biomedical-engineering-online.biomedcentral.com/articles/10.1186/1475-925X-12-111
(accessed 24.06.14).

[177] S. Yang, P. Premaratne, P. Vial, Hand Gesture Recognition: An Overview, Proc. of the IEEE Int. Conf. on Broadband Network & Multimedia Technology (BNMT), (2013) 63 - 69.

[178] P. Tome, J. Fierrez, R. Vera-Rodriguez, M. S. Nixon, Soft Biometrics and Their Application in Person Recognition at a Distance, IEEE Trans. on Information Forensics and Security, vol. 9, no. 1, (2014) 464 - 475.

[179] J. L. Waymann, Large-scale Civilian Biometric Systems Issues and Feasibility, Proc. of the Card Tech / Secur. Tech ID, (1997).

[180] U. Park, A. K. Jain, Face Matching and Retrieval Using Soft Biometrics, IEEE Trans. on Information Forensics and Security, vol. 5, no. 3, (2010) 406 - 415.

[181] D.-N. T. Congl, C. Achard, L. Khoudour, People Re-identification by Classification of Silhouettes Based on Sparse Representation, Proc. of the 2nd Int. Conf. on Image Processing Theory Tools and Applications (IPTA), (2010) 60 – 65.

[182] M. Ivasic-Kos, A. Iosifidis, A. Tefas, I. Pitas, Person De-identification in Activity Videos, Proc. of the Special Session on Biometrics, Forensics, De-identification and Privacy Protection (BiForD), (2014) 63 - 68.

[183] A. Nodari, M. Vanetti, I. Gallo, Digital Privacy: Replacing Pedestrians from Google Street View Images, Proc. of the 21st Int. Conf. on Pattern Recognition (ICPR), (2012) 2889 - 2893.

[184] M. Yang, K. Yu, Adapting Gender and Age Recognition System for Mobile Platforms, Proc. of the 3rd Chinese Conf. on Intelligent Visual Surveillance (IVS), (2011) 93 - 96.

[185] H. Lin, H. Lu, L. Zhang, A New Automatic Recognition System of Gender, Age and Ethnicity, The Sixth World Congress on Intelligent Control and Automation (WCICA), vol. 2, (2006) 9988 - 9991.

[186] G. Guo, G. Mu, Y. Fu, C. Dyer, T., Huang, A Study on Automatic Age Estimation Using a Large Database, Proc. of the 12th IEEE Int. Conf. on Computer Vision (ICCV), (2009) 1986 - 1991.

[187] D.-Y. Chen, K.-Y. Lin, Robust Gender Recognition for Real-Time Surveillance System, Proc. of the IEEE Int. Conf. on Multimedia and Expo (ICME), (2010) 191 - 196.

[188] G. Muhammad, M. Hussain, F. Alenezy, A. M. Mirza, G. Bebis, H. Aboalsamh, Race Recognition Using Local Descriptors, Proc. of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), (2012) 1525 - 1528.

[189] O.T.-C. Chen, J. J. Gu, P.-T. Lu, J.-Y. Ke, Emotion-Inspired Age and Gender Recognition Systems, Proc. of the 55th IEEE Int. Midwest Symposium on Circuits and Systems (MWSCAS), (2012) 662 - 665.

[190] U. Tariq, Y. Hu, T. S. Huang, Gender and Ethnicity Identification from Silhouetted Face Profiles, Proc. of the 16th IEEE Int. Conf. on Image Processing (ICIP), (2009) 2441 - 2444.

[191] J.-E. Lee, A. K. Jain, R. Jin, Scars, Marks and Tattoos (SMT): Soft Biometric for Suspect and Victim Identification, Proc. of the Biometrics Symposium (BSYM), (2008) 1 - 8.

[192] A. K. Jain, U. Park, Facial Marks: Soft Biometric for Face Recognition, Proc. of the 16th IEEE Int. Conf. on Image Processing (ICIP), (2009) 37 - 40.

[193] B. Heflin, W. Scheirer, T.E. Boult, Detecting and Classifying Scars, Marks, and Tattoos Found in the Wild, Proc. of the 5th IEEE Int. Conf. on Biometrics: Theory, Applications and Systems (BTAS), (2012) 31 - 38.

[194] A. E. Laumann, A. J. Derick, Tattoos and Body Piercing in the United States: A National Dataset, of the American Academy of Dermatology, vol. 55, issue 3, (2006) 413 - 421.

[195] D. Manger, Large-Scale Tattoo Image Retrieval, Proc. of the 9th Conf. on Computer and Robot Vision (CRV), (2012) 454 - 459.

[196] J.-E. Lee, R. Jin, A. K. Jain, Image Retrieval in Forensics: Tattoo Image Database Application, IEEE MultiMedia, vol. 19, no.1, (2011) 40 - 49.

[197] S. T. Acton, A. Rossi, Matching and Retrieval of Tattoo Images: Active Contour CBIR and Glocal Image Features, Proc. of the IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI), (2008) 21 - 24.

[198] D. Marcetic, S. Ribaric, V. Struc, N. Pavesic, An Experimental Tattoo De-identification System for Privacy Protection in Still Images, Proc. of the Special Session on Biometrics, Forensics, De-identification and Privacy Protection (BiForD), (2014) 57 - 62.

[199] J. Prinosil, A. Krupka, K. Riha, M. K. Dutta, A. Singh, Automatic Hair Color De-identification, Proc. of the Int. Conf. on Green Computing and Internet of Things (ICGCIoT), (2015) 732 - 736.

[200] M. Saini, P. K. Atrey, S. Mehrotra, M. Kankanhalli, Adaptive Transformation for Robust Privacy Protection in Video Surveillance, Advances in Multimedia Hindawi Publishing Corporation, vol. 2012, Article ID 639649, (2012) 1 - 14.

[201] T. Winkler, B. Rinner, Security and Privacy Protection in Visual Sensor Networks: A Survey, ACM Computing Surveys 47 (1), Article no. 2 (2014) 1 - 39.

[202] H. Sohn, D. Lee, W. De Neve, K. N. Plataniotis, Y. M. Ro, An Objective and Subjective Evaluation of Content-based Privacy Protection of Face Images in Video Surveillance Systems Using JPEG XR, in: F. Flammini, R. Setola, and G. Franceschetti (Eds.), Effective Surveillance for Homeland Security: Balancing Technology and Social Issues. CRC Press / Taylor & Francis, 2013, pp. 111-140.

[203] T. Justin, F. Mihelič, S. Dobrišek, Intelligibility Assessment of the De-identified Speech Obtained Using Phoneme Recognition and Speech Synthesis Systems, Lecture Notes in Artificial Intelligence (LNAI), vol. 8655, Springer, 2014, pp. 529 - 536.