Trusting the Computer in Computer Vision: A Privacy-Affirming Framework

Andrew Tzer-Yeu Chen, Morteza Biglari-Abhari, Kevin I-Kai Wang
Embedded Systems Research Group, Department of Electrical and Computer Engineering
University of Auckland, New Zealand
{andrew.chen, m.abhari, kevin.wang}@auckland.ac.nz

Abstract

The use of surveillance cameras continues to increase, ranging from conventional applications such as law enforcement to newer scenarios with looser requirements such as gathering business intelligence. Humans still play an integral part in using and interpreting the footage from these systems, but are also a significant factor in causing unintentional privacy breaches. As computer vision methods continue to improve, we argue in this position paper that system designers should reconsider the role of machines in surveillance, and how automation can be used to help protect privacy. We explore this by discussing the impact of the human-in-the-loop, the potential for using abstraction and distributed computing to further privacy goals, and an approach for determining when video footage should be hidden from human users. We propose that in an ideal surveillance scenario, a privacy-affirming framework causes collected camera footage to be processed by computers directly, and never shown to humans. This implicitly requires humans to establish trust, to believe that computer vision systems can generate sufficiently accurate results without human supervision, so that if information about people must be gathered, unintentional data collection is mitigated as much as possible.

1. Introduction

The use of surveillance cameras is an increasingly attractive option for collecting information about people, from identifying dangerous individuals in public spaces to measuring how shoppers move through shopping malls in market research studies. Traditionally, camera footage has to be processed manually, i.e. by a person watching either recorded or live video, interpreting that footage, and then conducting some form of data entry (e.g. counting the number of people) or making some decision (e.g. dispatching security guards to investigate an area). As the number of cameras and camera networks increases, the amount of data being collected will exceed our human capacity to process it, and therefore we will increasingly rely on computer vision algorithms to process video footage. However, this shift in control over the data has implications on how surveillance cameras impact the privacy of individuals, and offers new opportunities for reducing unnecessary breaches of privacy.

The philosopher William Parent defined privacy in 1983 as "the condition of not having undocumented personal information known or possess by others" [20]. He was not talking about a legal definition of privacy, but a notion that most people who value personal rights and freedom would be able to agree with [6] as a common, public, and collective right [25]. It can be argued that surveillance does not intrude on privacy if no undocumented information is gained; it therefore follows that if undocumented information has to be collected and made documented, privacy breaches are unavoidable. However, it can also be argued that this should not be considered in a binary nature; collecting a small amount of undocumented information is only a small privacy breach, and that may be preferable to a large privacy breach. The specific information becoming documented may also be important; most people would not mind low sensitivity information being documented, such as the colour of their clothing or the language(s) that they speak, but they might care more about high sensitivity information, such as their identity in conjunction with their recent purchases or their credit card details [12].

The majority of privacy-aware frameworks utilise some form of censorship so that a human user cannot see and inadvertently identify any individual in captured camera footage. This might range from superimposing a black box to block out the eyes of a person, to blocking image capture in certain locations such as bathrooms, to masking out individuals in each frame [4, 15, 23, 24, 26, 28]. However, this approach of obscuring part of the image misses the point of privacy. When discussing privacy in computer vision we should go beyond discussing it in terms of whether an individual is identifiable in any camera footage; we should ask what information we are receiving about a person in any footage, what information we genuinely need to collect about that person, and what other undocumented information we might inadvertently be collecting. Further, we need to know who is breaching the individual's privacy. Do our notions of privacy differ if that information is being collected, observed, and interpreted by humans, or by machines? Does reducing the presence of human interaction with the data make the collection of data less problematic?

In this position paper, we posit that humans need to design computer vision systems to only provide the information that is genuinely needed for a particular application. As the accuracy and speed of computer vision and image processing techniques improve, we suggest that privacy control should be taken out of the hands of humans, and that computers should be instructed to extract only the information necessary for human use. We propose that in an ideal surveillance scenario, no person should ever visually see any image footage, only processed outputs. We refer to this as a "privacy-affirming" approach (as opposed to "privacy-aware"). Fundamentally, that requires us to place our trust in computers to process data correctly, repeatably, and securely, so that no excess undocumented information falls into human hands.

The following sections discuss the differences between human and machine processing of camera footage, applying the concept of abstraction to video footage for protecting privacy, how that abstraction can be applied through distributed computing, how computer vision system designers can use a simple test to ascertain when privacy-affirming protections are appropriate for specific applications, and what other options are available when the privacy-affirming approach is not appropriate.

2. The Human Element

A number of factors are driving the concern behind privacy breaches in camera-based surveillance. Firstly, the fact that footage is being recorded at all invariably means that previously undocumented information becomes documented. However, the act of data collection itself may not be a primary concern; rather, it is the fact that the documented data could be used to inform or influence decisions about the recorded individual, potentially by unauthorised people, that is of paramount concern [16, 17, 30]. If a surveillance camera system monitors a person walking into a fast food restaurant, the fact that this piece of information has become documented may not be a significant cause for concern. Rather, the privacy concern is that this piece of information could be used in the process of making decisions that may lead to negative outcomes for that person, for example, a health insurance company raising that person's premiums. The underlying concern is that collected information could be used for nefarious, embarrassing, or otherwise unfair means. Furthermore, the human element can be a weakness in the system; human operators of surveillance systems may use information that they gather for personal gain, discriminatory targeting, and voyeurism [18].

In this paper, we do not deal with whether surveillance systems should exist, i.e. we do not justify the main purpose(s) of monitoring and identifying people, even if it leads to negative outcomes for targets. Rather, we focus on the extraneous decisions that can be made following unintentional information capture. For example, a surveillance camera in a city centre that is intended to protect people from criminal activities may also collect footage that shows a person walking into an Alcoholic's Anonymous (AA) meeting. The privacy fear is that another human may watch the video footage and identify the individual, thereby breaching the target's privacy rights if they had wanted to keep that information private. The main purpose of this system is to catch criminals, not to track innocent people and their personal life choices, yet these types of side effects arise as a consequence of using surveillance cameras/video footage as a rich source of information. Whether we should have video surveillance to help prevent crime or not is outside the scope of this paper; rather, we want to understand how undesirable, unintentional collection of undocumented information about individuals can be mitigated. In other words, our paper posits the question: if a surveillance camera system must exist, how can we best minimise unintentional breaches of privacy?

The rise of computer-automated processing of video footage presents an opportunity for reducing the risk of unintended privacy breaches. In the previous scenario, if we program a computer to run activity recognition algorithms that detect if suspicious or criminal activity is occurring, then the computer will only arrive at conclusions related to that objective. We have control over the machine, and can program it so that it is oblivious to the fact that a person has walked into an AA meeting. On the other hand, with a human operator, even if instructions are given to only focus on suspicious or criminal activity, the human brain will inadvertently collect far more information. Even with no ill intention or malice, the operator may see an individual that they recognise walk into an AA meeting, and then the operator's opinions about that person may change, leading to changes in the operator's decision-making in relation to that person. This is clearly a breach of privacy, caused purely by the human brain's tendency to process sensory information regardless of our intentions. Some would argue that even without processing, the act of collecting excess data that is not necessary for achieving the primary objective is in itself a breach of privacy.

Most privacy-aware frameworks try to mitigate this by reducing the amount of information available to human operators. For example, we might have a system that uses face detection and blurs faces in recorded footage before showing it to a human operator [1]. This generally makes iden-

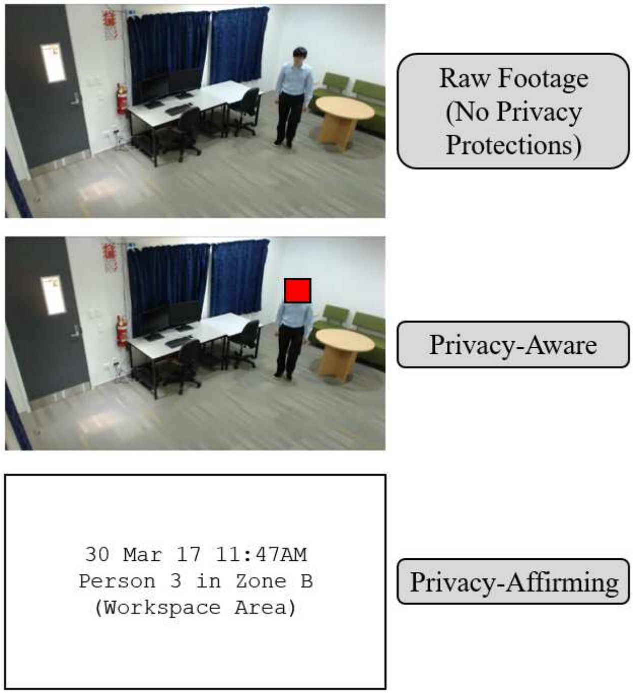

Figure 1. An example of the different views provided to human operators under different privacy-oriented frameworks

tifying the individual much more challenging since the face contains most of the discriminative features that allow us to distinguish one person from another. However, the human operator may still receive too much information; they could piece together other information like the clothing of the person, the time and place the person was observed at, and other contextual information that can lead to identification. Using post-processing to reduce the amount of information in the image can be computationally expensive, yet can still unintentionally reveal too much, especially when there are missed detections (false negatives) that cause videos to be momentarily uncensored [13, 27, 29].

Rather than taking all of the image data and using computer vision to remove a small amount of it before giving it to a human operator, perhaps the opposite privacy-affirming proposition is better: taking all of the information and using computer vision to extract only the necessary data and giving that to the human operator, as shown in Figure 1. Privacy-aware and privacy-affirming are somewhat analogous to black-listing and white-listing; rather than censoring some unwanted information, explicitly only delivering the desired information may be a better approach. Under the status quo, the privacy chain of accountability ends with a human, but perhaps machines can help absolve humans of that responsibility. We acknowledge that this may not be possible for all computer vision applications with current-day technology, but as computer vision continues to improve, we suggest that system designers should consider the merits of an abstractive approach to protecting privacy. In Section 5 we will discuss when privacy-affirming approaches may be suitable, and when we may have to rely on more traditional methods, but we will first discuss the principles behind privacy-affirming approaches in more detail and how these principles could be applied.

3. Hiding Information

Abstraction is one of the key concepts in computing; by moving to higher and higher levels of abstraction, developers can save time (and therefore money) by hiding unnecessary information and focusing on high-level processes. This concept could be applied to surveillance too. Rather than requiring human operators to interact in some way with the raw data (i.e. the video footage), we can hide that information from them and present higher-level information and statistics instead. Hiding unnecessary information from humans follows naturally from common privacy principles; in New Zealand, the Privacy Act lists twelve privacy principles, the first of which requires that personal information not be collected unless "the collection of the information is necessary for that purpose" [14]. Similar legislative principles or requirements exist in other jurisdictions [22]. Under the status quo, by collecting camera footage and presenting that directly to human users, we collect far more information that is necessary for the specific purposes we seek.

Images are incredibly rich in information in comparison to other forms of sensory data like audio. When humans look at an image, our brains abstract many of the details away from ourselves (also known as encoding), allowing us to focus on high-level features such as where objects of interest are, classifying those objects, identifying specific objects or people, and determining the context of the image [19]. Memory is also important, and our brains do not normally make an exact record of the past footage that has entered our eyes. We may be able to store images in short-term memory, we may remember high-level details for longer, and we may be able to reconstruct what the scene looked like with our imagination based on those details, but long-term memory does not store the vast majority of what we see [33]. It would be simply inefficient to do so, when the majority of the information that enters our eyes has little meaning or bearing on our future decisions, so our brains have learnt to discard that information. This presents a justification for how we can ask machines to abstract away information in images through the use of computer vision.

Naturally, system designers may feel uneasy at the idea of throwing away image information when it may be useful for debugging or as a backup in case of system failure. However, there are two types of machines that we already trust to discard unnecessary sensory information. Firstly, we can consider machines that are monitoring the environment, but discard data until they are activated. For example, personal assistant devices such as Amazon's Alexa utilise microphones to pick up human speech. They are actually constantly listening because they need to hear the activation/command word that enables the system. Developers trust that the system will hear that word correctly and that all other information that may be inadvertently collected while the device is listening can be discarded. At the same time, human users of these devices trust that audio not relevant to the operation of the device is discarded.

Secondly, we can consider machines where delivering raw data would be unhelpful to human users. For example, complex algorithms are used in driver-assistance systems to process lidar and camera data to identify hazards, but we do not show raw outputs to the human user (i.e. the driver). Instead, we trust the car to process the sensory information and draw usable conclusions, such as indicating to the driver that there is a hazard that requires the driver to slow down or stop the car. We would not expect the human driver to be able to interpret raw data while also driving, so the car abstracts the low-level sensor data away from the human and only presents high-level summaries (in this case, an alert). Clearly, abstraction is already an acceptable concept when it comes to dealing with machines which discard unnecessary information and hide other information from human users.

Combining these two examples presents a scenario suitable for use in surveillance camera systems. For example, we could use surveillance cameras that are always watching but do not record the footage until trigger conditions occur, such as a target individual being identified or a violent action being detected. Those cameras also could hide the raw footage away from humans, but simply present a report. For example, if we are trying to track a suspect across a camera network, the report could just list the times and positions where that particular person was detected, or show this on a map. This prevents the human operator from accessing more information than strictly necessary for the system objective to be achieved, fulfilling the promise of the principle of least privilege [9]. This can be considered to be the least intrusive form of camera surveillance, assuming that surveillance has to occur.

4. Distributed Computing

In traditional surveillance systems, cameras are simply sensor devices; they capture images at a fixed location and transmit that footage to a central location for processing and storage. However, the centralised approach presents several challenges: as the number of cameras increases in the network, the amount of computation required by any central computer or human operator increases, and the network can become saturated as the amount of data exceeds the network capacity. Additionally, there are security concerns as a central computer is a single point of failure, so by compromising that single computer (or operator) an attacker has control and access to the entire network's data. As hardware capabilities continue to improve, and computer vision algorithms become further optimised for computational efficiency, we have seen the rise of smart cameras; essentially cameras with some on-board processing capacity [7]. These smart cameras can be used in a distributed manner, alleviating processing load on a central computer and improving security by eliminating the single point of failure. Camera surveillance research has already started to incorporate different types of smart cameras, balancing computational capacity with power and energy limitations [32, 35]. These advantages can be complementary to the objective of preserving the privacy of observed individuals.

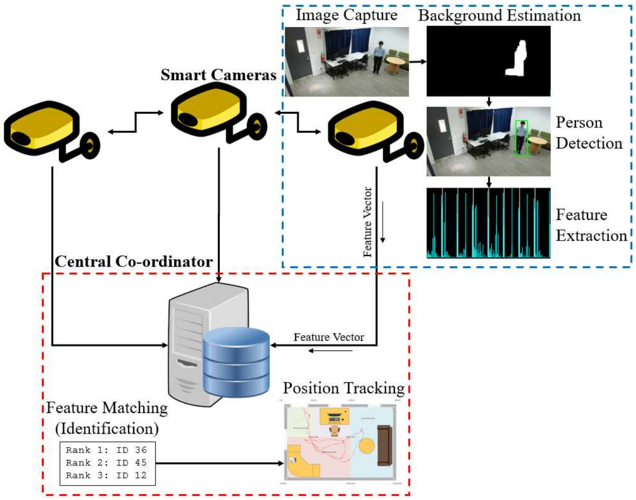

Figure 2. A distributed system architecture, with tasks divided between the smart cameras (blue) and the central co-ordinator (red)

Figure 2 shows a potential distributed computing layout, where smart cameras do some of the processing at the point of image capture. Smart cameras generally have constrained computational resources, so it may be inappropriate to execute a complete algorithm there, and some central processing capability may be helpful for reducing the computational load on the smart cameras. Additionally, the presence of some central co-ordinator can simplify the communication protocols significantly, avoiding complex data sharing procedures. The role of the smart camera is therefore to process the image initially, extract some top-level features, and then pass those to the central co-ordinator for further processing. Importantly, the human operator only interacts with the central co-ordinator and never directly with any of the smart cameras, therefore preventing them from ever seeing the raw captured image. Additionally, the cameras have no notion of identity at the point of image capture, so if an individual camera is compromised, there is no assigned identity information available.

For example, in a market research application we may want to track individuals as they move through a large store to find out which paths they take so that we can design more efficient store layouts. We could use a person tracking system that combines several computer vision solutions into a simple image processing pipeline. A number of modules are executed sequentially one after another, separated so that they can be swapped out for newer and better methods as necessary. With the use of smart cameras, one potential system architecture (as shown in Figure 2) is for the background subtraction, person detection, and feature extraction modules to be executed on the camera itself. Feature vectors (including the person location) could then be sent to the central co-ordinator for more computationally expensive matching against a large database, and rectification between multiple camera views to centrally determine the person track. With this architecture, the smart cameras never send the raw image to the central co-ordinator, only feature vectors, which is the only information necessary from the raw image for completing the person re-identification and tracking objective of the system. The image therefore only exists in the memory of the smart camera, and can be immediately deleted after processing while still retaining the utility of the overall system. Additional security precautions such as hardware-enforced memory access control [31] can be used to restrict access to the image by foreign processes, and images are never transmitted across any networking interface, preventing hackers from stealing those images before deletion. This has the added benefit of significantly reducing network bandwidth requirements, which is particularly important in wireless systems [8, 35]. The distributed nature of the system architecture allows for a natural separation of responsibilities in a way that also eliminates the unintentional collection of undocumented information.

This is not to say that distributed computing is a necessity for privacy-affirming frameworks. While distributed architectures may more naturally fit with privacy-affirming principles, centralised systems can still be designed to reduce the likelihood of leaking the raw image to a human user. This may be important in less modularised computer vision methods, such as state-of-the-art Convolutional Neural Networks (CNN) that may require the entire raw image as an input. If significant processing resources are required, it may not be possible to split a complicated CNN into separate modules for distributed processing without expensive co-ordination. In these types of scenarios, the system can be designed so that the raw image is processed and then discarded, with only processed outputs/statistics shown to the user, rather than overlaying that on top of the raw image as is currently often the case.

Jana et al [10] describe a system that intercepts video frames before they are processed (via image library APIs such as OpenCV) called Darkly, which vastly reduces the amount of information made available for processing. This is based on user-selected levels of privacy, and *privacy transforms* which use various pre-processing algorithms, such as thresholding and clustering, depending on the API call(s). Darkly is perhaps a seminal example of a system that moves towards privacy-affirming principles, abstracting away data and not only hiding it from the human user, but also potentially untrustworthy computers, in a similar way to our proposed distributed computing architecture. The Darkly paper reports that for the relatively simple applications tested, these principles can be applied successfully without significant loss of accuracy in the overall application. The challenge comes in applying these principles to more complex tasks with unspecified or broad application needs as computer vision techniques continue to evolve. Further steps such as deleting the raw image data immediately after privacy transforms have been applied, or further abstracting away results for display to human users, could also improve Darkly's ability to affirm privacy.

5. A Test for Privacy-Oriented Systems

In 2005, a team from IBM proposed a framework that used smart cameras to extract information at the point of image capture, and produce outputs at different levels of abstraction [29]. The primary limitation of their framework is that it promotes the use of permissions, providing access to data at different levels of the abstraction hierarchy. Other papers have also supported the use of access controls or usage controls to limit access to raw data to only trustworthy individuals [3, 21]. Unfortunately, this still allows for potential abuse, either by the human operators themselves who may find ways to circumvent access controls, or by external hackers who can use social engineering or network-based exploits to gain access to user accounts. Instead, we propose removing permissions entirely (whenever it is appropriate to do so) by making it so that no human, regardless of their role or rank, can have access to potentially privacy-infringing material.

We propose a simple test that system designers can use when developing surveillance camera systems, in order to help decide whether a privacy-affirming approach is appropriate for a given application. There are two main factors to consider: whether the footage is required for visual evidence, and whether the footage requires human processing as part of achieving the main objective of the surveillance system. We show these factors in Table 1 and discuss potential scenarios for each part of the rubric.

Archive-critical scenarios are ones where the raw video footage is actually a requirement for meeting the principal objective or purpose of the surveillance system. For example, a law enforcement camera system may require the raw footage for the purposes of evidence in court trials. Even though we already trust other types of machines

	Archive-Critical	Archive-Free
Human Processing	Access Control or Permissions	Privacy-Aware
Machine Processing	Privacy-Aware *and* Permissions	Privacy-Affirming

Table 1. A rubric for making system design decisions for privacy in surveillance camera systems

to produce processed results for evidential purposes (e.g. DNA tests, GPS-based electronic monitoring of offenders, phone call metadata), the high fidelity of visual information means that it is considered to be extremely valuable, and any intentional degradation of that fidelity could be considered destruction of evidence. This is a reflection of our reliance on visual information, and in a sense we expect raw footage for court trials because that is what we have always had in the past. As computer vision technologies improve, it is possible that new privacy legislation could be introduced that forces a shift in these expectations. However, under the status quo, a permissions-based system as proposed in [29] may be the only way to mitigate unintentional privacy breaches for archive-critical scenarios, especially if a human is required for processing the data. If the majority of the processing can be done by a machine (i.e. computer), then we can use privacy-aware techniques to produce anonymised footage for general use [23], while keeping the raw footage accessible via permissions-based systems for use in archive-critical applications [29].

Archive-free scenarios are ones where the raw video footage can be discarded, and the main objective is to extract high-level processed data and statistics. For example, in an urban planning application where we want to measure the flow of people through a public space, retaining access to recorded footage is not necessary if we can obtain processed statistics reporting the number of people at different times. If some human processing is required (because computer vision is not accurate or fast enough), then privacy-aware post-processing should be applied before the footage is provided for human interpretation. If the computer vision approach is accurate and fast enough, then we should use the proposed privacy-affirming approach, eliminating any exposure of the raw footage to humans. Even under existing legislation, such as the “Right to Privacy” in the European Union [2], it could be argued that the privacy-affirming approach is required wherever possible in order to avoid collecting information superfluous to the primary objective.

Currently, the vast majority of surveillance camera applications still require some human processing, and most large-scale networks are implemented with public safety or law enforcement objectives so they are archive-critical. It may be that privacy-affirming approaches are not suitable for most current-day scenarios. However, two trends may tip this balance in favour of privacy-affirming principles. Firstly, the number of archive-free applications will continue to grow as the cost of camera systems fall and more businesses become aware of new opportunities enabled by smart cameras, such as market research. Secondly, our expectations of what evidence is genuinely required for archive-critical applications may change, especially if privacy-conscious activists and governments promote higher standards of protection. In the past, privacy-affirming frameworks were not technologically possible, but that barrier is quickly eroding, and since they can lead to better privacy outcomes for individuals and for society as a whole, they should be carefully considered for adoption as a baseline standard.

For privacy-affirming methods to become more popular, trust must be established with the computer in computer vision. This trust can only be earned over time, through repeated demonstrations of the validity and accuracy of computer vision applications, as we have seen with other technologies such as assembly line automation and aeroplanes. Currently, machine processing may not be accurate or fast enough for many applications, and because of that, humans may not be able to trust machines with their camera footage. High accuracy will need to be proven beyond minimum thresholds, and a certification and auditing regime may be necessary to build public confidence in these systems. Human operators will have to get used to the idea of making decisions informed only by processed algorithm outputs rather than seeing the raw footage themselves. Public education and marketing campaigns will be needed to convince the public that certain camera systems are privacy-affirming, and that these may protect their privacy better than privacy-aware systems (or systems with no privacy protections at all). There are (justified) concerns that integrating computer vision with surveillance may enable more information to be extracted than possible by humans [11], so privacy-conscious surveillance camera network owners and operators will need to become much more transparent about how their systems work, what data is being collected, how that data is being used, and what data is not being collected. Only then will we be able to empirically measure the impacts of a privacy-affirming framework, and whether it reduces unintentional privacy breaches in reality. Ultimately, whether privacy is respected by surveillance camera systems is dependent on the humans involved in creating that system - privacy-aware and privacy-affirming frameworks can never comprehensively protect against individuals, corporations, or states with malicious intentions.

6. Conclusions

In this paper, we propose a privacy-affirming approach for managing privacy concerns in surveillance camera systems. By only presenting the required data and hiding everything else, this protects privacy much more completely and consistently than censorship-based privacy-aware approaches. Abstraction is already accepted in many of our computing and sensor systems, and combining that with distributed computing leads to a natural separation of responsibilities that mitigates the human element in privacy breaches. We describe a simple test for determining the appropriateness of this framework, and acknowledge that it may not be suitable for all scenarios. Additionally, privacy-affirming approaches should be combined with other privacy-conscious steps, such as encryption of transmitted data, event-based triggering to avoid unnecessary recording, deleting unneeded data by default after a certain time period, and using relative IDs where possible to preserve anonymity, as well as other general security precautions such as those presented in [34]. Even though privacy-affirmed data is safer than privacy-aware data, it should still be treated as privileged and not immune from exploitation. The privacy-affirming approach forms part of our research project on computationally efficient indoor person tracking systems, and will be a guiding principle in the design and implementation of our experimental platform.

In the late 1990s, futurist David Brin suggested two choices for managing privacy in surveillance camera systems: trusting human authorities to act with the interest of the citizenry at heart, or to democratise access by allowing the human public to "watch the watchers" in order for everyone to trust everyone else [5]. Perhaps there is a third option - to trust machines to process and extract exactly the right amount of information needed to serve human objectives, reducing the power imbalance between the observers and the observed, and thus preserving privacy from unintended and unnecessary vulnerability.

Acknowledgements

The authors would like to thank Benjamin Tan, Charlotte Austin, Sarah and Anton Wilson-Smith, and the anonymous reviewers for their valuable feedback and comments.

References

[1] M. S. Barhm, N. Qwasmi, F. Z. Qureshi, and K. el Khatib. Negotiating privacy preferences in video surveillance systems. In *International Conference on Industrial Engineering and Other Applications of Applied Intelligent Systems (IEA/AIE)*, pages 511–521, 2011. 2

[2] F. E. Bignami. Privacy and law enforcement in the european union: the data retention directive. *Chicago Journal of International Law*, 8(1):233–256, 2007. 6

[3] P. Birnstill and A. Pretschner. Enforcing privacy through usage-controlled video surveillance. *IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pages 318–323, 2013. 5

[4] M. Blazevic, K. Brkic, and T. Hrkac. Towards reversible de-identification in video sequences using 3d avatars and steganography. In *Croatian Computer Vision Workshop (CCVW)*, pages 9–14, 2015. 1

[5] D. Brin. *The Transparent Society: Will Technology Force Us To Choose Between Privacy And Freedom?* Basic Books, 1999. 7

[6] J. DeCew. Privacy. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, 2015. 1

[7] A. Hornberg. *Handbook of Machine Vision*, page 703. Wiley, 2006. 4

[8] W. Hu, B. Amos, Z. Chen, K. Ha, W. Richter, P. Pillai, B. Gilbert, J. Harkes, and M. Satyanarayanan. The case for offload shaping. In *International Workshop on Mobile Computing Systems and Applications (HotMobile)*, pages 51–56, 2015. 5

[9] S. Jana, D. Molnar, A. Moshchuk, A. Dunn, B. Livshits, H. J. Wang, and E. Ofek. Enabling fine-grained permissions for augmented reality applications with recognizers. In *USENIX Conference on Security (SEC)*, pages 415–430, 2013. 4

[10] S. Jana, A. Narayanan, and V. Shmatikov. A Scanner Darkly: Protecting User Privacy from Perceptual Applications. In *IEEE Symposium on Security & Privacy*, May 2013. 5

[11] H. Koch, T. Matzner, and J. Krumm. Privacy enhancing of smart cctv and its ethical and legal problems. *European Journal of Law and Technology*, 4(2), 2013. 6

[12] L. Lee, J. Lee, S. Egelman, and D. Wagner. Information disclosure concerns in the age of wearable computing. In *NDSS Workshop on Usable Security (USEC)*, 2016. 1

[13] S. Moncrieff, S. Venkatesh, and G. A. W. West. Dynamic privacy in public surveillance. *Computer*, 42(9):22–28, 2009. 3

[14] New Zealand Government. Privacy Act 1993, 2013. 3

[15] E. M. Newton, L. Sweeney, and B. Malin. Preserving privacy by de-identifying face images. *IEEE Transactions on Knowledge and Data Engineering*, 17(2):232–243, 2005. 1

[16] D. H. Nguyen, A. Bedford, A. G. Bretana, and G. R. Hayes. Situating the concern for information privacy through an empirical study of responses to video recording. In *SIGCHI Conference on Human Factors in Computing Systems (CHI)*, pages 3207–3216, 2011. 2

[17] D. H. Nguyen, A. Kobsa, and G. R. Hayes. An empirical investigation of concerns of everyday tracking and recording technologies. In *International Conference on Ubiquitous Computing (UbiComp)*, pages 182–191, 2008. 2

[18] C. Norris and G. Armstrong. *The maximum surveillance society: the rise of CCTV*. Berg, 1999. 2

[19] B. A. Olshausen and D. J. Field. Vision and the coding of natural images: The human brain may hold the secrets to the best image-compression algorithms. *American Scientist*, 88(3):238–245, 2000. 3

[20] W. A. Parent. Privacy, morality, and the law. *Philosophy & Public Affairs*, 12(4):269–288, 1983. 1

[21] Q. M. Rajpoot and C. D. Jensen. Security and privacy in video surveillance: Requirements and challenges. In *IFIP International Information Security Conference*, pages 169–184, 2014. 5

[22] Q. M. Rajpoot and C. D. Jensen. *Video Surveillance: Privacy Issues and Legal Compliance*. IGI Global, 2015. 3

[23] H. A. Rashwan, A. Solanas, D. Puig, and A. Martínez-Ballesté. Understanding trust in privacy-aware video surveillance systems. *International Journal of Information Security*, 15(3):225–234, 2016. 1, 6

[24] N. Raval, L. P. Cox, A. Srivastava, A. Machanavajjhala, and K. Lebeck. Markit: privacy markers for protecting visual secrets. In *International Conference on Ubiquitous Computing (UbiComp)*, pages 1289–1295, 2014. 1

[25] P. M. Regan. *Legislating Privacy: Technology, Social Values, and Public Policy*. University of North Carolina Press, 1995. 1

[26] F. Roesner, D. Molnar, A. Moshchuk, T. Kohno, and H. J. Wang. World-driven access control for continuous sensing. In *ACM SIGSAC Conference on Computer and Communications Security (CCS)*, pages 1169–1181, 2014. 1

[27] M. Saini, P. K. Atrey, S. Mehrotra, and M. Kankanhalli. W3-privacy: understanding what, when, and where inference channels in multi-camera surveillance video. *Multimedia Tools and Applications*, 68(1):135–158, 2014. 3

[28] J. Schiff, M. Meingast, D. K. Mulligan, S. Sastry, and K. Goldberg. *Respectful Cameras: Detecting Visual Markers in Real-Time to Address Privacy Concerns*, pages 65–89. Springer London, 2009. 1

[29] A. Senior, S. Pankanti, A. Hampapur, L. Brown, Y.-L. Tian, A. Ekin, J. Connell, C. F. Shu, and M. Lu. Enabling video privacy through computer vision. *IEEE Security & Privacy*, 3:50–57, 2005. 3, 5, 6

[30] H. J. Smith and S. J. Milberg. Information privacy: Measuring individuals' concerns about organizational practices. *MIS Quarterly*, 20(2):167–196, June 1996. 2

[31] B. Tan, M. Biglari-Abhari, and Z. Salcic. A system-level security approach for heterogeneous mpsocs. In *Conference on Design and Architectures for Signal and Image Processing (DASIP)*, pages 74–81, 2016. 5

[32] H. Tan, J. Ni, S. Zhang, J. Wu, S. Chan, and Y. Hung. On the hardware/software design and implementation of a high definition multiview video surveillance system. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 3(2):248–262, 2013. 4

[33] J. J. Todd and R. Marois. Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature*, 428:751–754, 2004. 3

[34] T. Winkler and B. Rinner. Security and privacy protection in visual sensor networks: A survey. *ACM Computing Surveys*, 47(1):2:1–2:42, 2014. 7

[35] T. Zhang, A. Chowdhery, P. Bahl, K. Jamieson, and S. Banerjee. The design and implementation of a wireless video surveillance system. In *21st Annual International Conference on Mobile Computing and Networking (MobiCom)*, pages 426–438, 2015. 4, 5