CARDEA: A CONTEXT-AWARE AND INTERACTIVE VISUAL PRIVACY CONTROL FRAMEWORK

by

RUI ZHENG

A Thesis Submitted to
The Hong Kong University of Science and Technology
in Partial Fulfillment of the Requirements for
the Degree of Master of Philosophy
in Computer Science and Engineering

October 2016, Hong Kong

Authorization

I hereby declare that I am the sole author of the thesis.

I authorize the Hong Kong University of Science and Technology to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize the Hong Kong University of Science and Technology to reproduce the thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

RUI ZHENG

CARDEA: A CONTEXT-AWARE AND INTERACTIVE VISUAL PRIVACY CONTROL FRAMEWORK

by

RUI ZHENG

This is to certify that I have examined the above M.Phil. thesis

and have found that it is complete and satisfactory in all respects,

and that any and all revisions required by

the thesis examination committee have been made.

ASSISTANT PROF. PAN. HUI, THESIS SUPERVISOR

PROF. QIANG YANG, HEAD OF DEPARTMENT

Department of Computer Science and Engineering

21 October 2016

ACKNOWLEDGMENTS

thank god

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CARDEA: A CONTEXT-AWARE AND INTERACTIVE VISUAL PRIVACY CONTROL FRAMEWORK

by

RUI ZHENG

Department of Computer Science and Engineering

The Hong Kong University of Science and Technology

ABSTRACT

The growing popularity of mobile and wearable devices with builtin cameras, the bright prospect of camera related applications such as augmented reality and lifelogging system, the increased ease of taking and sharing photos, along with advances in computer vision techniques, have greatly facilitated peoples lives in many aspects, but inevitably raised peoples concerns about visual privacy at the same time.

Motivated by the finding that peoples privacy concerns are influenced by the context, in this thesis, we propose Cardea, a contextaware and interactive visual privacy control framework that enforces privacy policies according to peoples privacy preferences. The framework provides people with finegrained visual privacy control using: *i)* personal privacy profiles, with which people can define their contextdependent privacy preferences; *ii)* different visual indicators: face features and tags, for devices to automatically locates individuals who request privacy protection; *iii)* hand gestures, for people to temporarily update and flexibly inform cameras of their privacy preferences.

Benefited from recent progresses in face and object recognition, Cardea offers a way for context-dependent privacy control in a natural and flexible manner, which differs from tag and marker based systems. We design and implement the framework consisting of Android client app and cloud control server, with convolutional neural networks as core of the image processing module. Our evaluation results confirm such framework is practical and effective, showing promising future for contextaware visual privacy control on mobile and wearable devices.

CHAPTER 1

INTRODUCTION

The concern about visual privacy has been growing in last decade with increasing adoption of video surveillance systems for security reasons. The statistics shows there are 125 video surveillance cameras per thousand people in U.S. by 2014 [1]. Momentum of new technologies such as the Internet of Things (IOT) will keep driving global video surveillance market in following years, which will raise more privacy concerns.

Other than closed-circuit television (CCTV) surveillance systems for security reasons, handheld devices such as camera phones are also used extensively for the recording of meaningful life moments. Recently, coming with the explosion of products in augmented reality (e.g., Google Glass), robotics (e.g., iRobot Create platform), and gaming (e.g., Kinect), is more and more cameras being embedded in these platforms for the enhancement of life experiences. The trend of embedding cameras, especially in wearables, will keep growing, an example of which is smart contact lens [2]. However, the ubiquitous presence of cameras, the ease of taking photos and recording videos, along with "always on" and "non overt act" features threaten individuals to have private or anonymous social lives, raising people's concerns of visual privacy.

More specifically, photos and videos captured without getting permissions from bystanders, and then uploaded to social networking sites, can be accessed by everyone online, potentially leading to invasion of privacy. Malicious applications on the device may also inadvertently leak captured media data [3].

Benefited from research breakthroughs from deep learning community, current vision perception systems are advancing fast in their capabilities of understanding image and video contents [4]. Nowadays, recognition technologies can link images to specific people [5, 6, 7], places [8], and general objects [9], making what previously unsearchable now searchable [10], thus reveal far more private information than expected.

Both legal and technical measures have been proposed to resolve visual privacy concerns. For instance, Google Glass is banned at places such as banks, hospitals, and bars [11]. However, prohibition of cameras usage does not resolve the issue fundamentally, instead it may intrude people's rights to capture happy moments. As a result, there are growing needs to design technical solutions to protect individuals' visual privacy in a world with pervasive cameras. Technical solutions that have been proposed so far are still limited, in the way that they are mostly based on static policies, thus users can not flexibly express their individualized privacy preferences based on surrounding contexts when they are captured. Moreover, previous works require users to wear visual markers such as hats [12] for the detection of interested persons, or clip tags such as QR codes [13, 14] for the fetching of privacy polices. Despite technical feasibilities of these approaches, the extra need of setting up markers/tags and the resulting aesthetically unpleasant appearance will hinder users' willingness to adopt these solutions.

Therefore, the motivation of this thesis is to seek a more natural, userfriendly, flexible, and fine-grained mechanism for people to express, modify, and control their individualized privacy preferences. Under this guideline, we propose Cardea, a context-aware and interactive visual privacy control framework, which lets individuals control their visual privacies through: *i)* personal privacy profiles, with which people can define their contextdependent privacy preferences; *ii)* different visual indicators: face features and tags, for devices to automatically locates individuals who request privacy protection; *iii)* hand gestures, for people to temporarily update and flexibly inform cameras of their privacy preferences. When using Cardea, the device will automatically compute context factors, compare them with peoples privacy profiles, and finally enforce privacy policies conforming to peoples privacy preferences. To our knowledge, this is one of the pioneering works, if not the first, that leverages deep learning models to enable visual privacy control in a context-specific and interactive manner.

The rest of the thesis is organized as follows: We first review and discuss related works on visual privacy control in Chapter 2. Following that we introduce convolutional neural networks, the core of Cardea's image processing module, and their applications on related computer vision problems. We then give details about the design, implementation and evaluation of Cardea in Chapter 4. Finally, we share our thoughts on possible future work and conclude the thesis in Chapter 5.

CHAPTER 2

RELATED WORKS

2.1 User Studies

year	privacy survey	problem setting		technical solution		enforcement time		privacy object	
		video surveillance	mobile cameras	computer vision	cryptography	in-situ / run time	access / distribution	user	bystander
I-Pic: A Platform for Privacy-Compliant Image Capture [15]									
2016	✓		✓	✓	✓	✓			✓
Recent work which allows people to broadcast their privacy preferences and appearance information to nearby devices through BLE. These preferences are based on social context.									
What You Mark is What Apps See [16]									
2016			✓	✓		✓	✓	✓	
Propose a system that give users control to mark secure regions for third-party applications. It is implemented within Android camera subsystem.									
Sensitive Lifelogs: A Privacy Analysis of Photos from Wearable Cameras [17]									
2015	✓		✓						
Analyze the photos collected in [18], seeking to understand what makes a photo private and what participants said about their images.									
Screenavoider: Protecting Computer Screens from Ubiquitous Cameras [19]									
2014			✓	✓		✓	✓	✓	
Present a framework that controls the collection and disclosure of lifelogging datasets which contain computer screens and possible sensitive contents.									
PlaceAvoider: Steering First-Person Cameras away from Sensitive Spaces [20]									
2014			✓	✓		✓	✓	✓	
Introduce a prototype for owners of first-person cameras to 'blacklist' sensitive places (like bathrooms and bedrooms).									
Privacy.Tag: Privacy Concern Expressed and Respected [13]									
2014			✓	✓	✓		✓		✓
Propose using QR code as privacy tag to link an individual with his photo sharing preferences. These preferences are based on web domains.									
Privacy Behaviors of Lifeloggers using Wearable Cameras [18]									
2014	✓		✓			✓	✓	✓	✓
Conducted an *in situ* user study on privacy behaviors of 36 participants who wore lifelogging devices for a week.									
Courteous Glass [21]									
2014			✓	✓		✓			✓

year	privacy survey	problem setting		technical solution		enforcement time		privacy object	
		video surveillance	mobile cameras	computer vision	cryptography	in-situ / run time	access / distribution	user	bystander
Wearable camera integrated with a FIR (far-infrared) imagers that turns off recording when new persons or specific gestures are detected.									
World-Driven Access Control for Continuous Sensing [14]									
2014			✓	✓		✓	✓	✓	
Propose a general framework that allows objects to explicitly specify its access policies. Policy triggers can be visual indicators or anything that can be detected in other research works.									
In Situ with Bystanders of Augmented Reality Glasses: Perspectives on Recording and Privacy-Mediating Technologies [22]									
2014	✓		✓			✓	✓	✓	✓
Investigate the privacy perspectives of individuals when they are bystanders around AR devices. Conducted 12 field sessions in cafés and interviewed 31 bystanders regarding their reactions to a co-located AR device.									
A Scanner Darkly: Protecting User Privacy From Perceptual Applications [23]									
2013		✓	✓	✓		✓	✓	✓	
Perceptual applications can only access transformed objects such as sketches, faces, etc.									
Enabling Fine-Grained Permissions for Augmented Reality Applications With Recognizers [24]									
2013			✓	✓		✓		✓	
Third party AR applications can only access high-level objects such as Skeleton, Hand Position, etc.									
PriSurv: Privacy Protected Video Surveillance System Using Adaptive Visual Abstaction [25]									
2008		✓		✓			✓	✓	✓
Propose a privacy control mechanism for surveillance videos based on closeness between content objects and content viewers. RFID tags are used to improve the detection of people in videos.									
Respectful Cameras: Detecting Visual Markers in Real-Time to Address Privacy Concerns [12]									
2007		✓		✓			✓	✓	✓
A video surveillance system that allows people who wish to remain anonymous wear colored markers such as hats or vests, and their faces will be blurred.									
Privacy Management for Portable Recording Devices [26]									
2004			✓		✓		✓	✓	
Propose an approach that closed closed devices can encrypt data together during recording utilizing short range wireless communication to exchange public keys and negotiate encryption key. Only by obtaining all of permissions from people who encrypt the recording can one decrypts it.									

Table 2.1: Mobile cameras include cameras in smartphones, AR, VR, lifelogging and other wearable devices. Computer vision methods may be assisted with extra sensors such as RFID tags [25], FIR imagers [21]. There are other factors like implementation layer (app level or os level) that are not listed due to space limitation.

2.2 Proposed Solutions

2.3 Challenges

CHAPTER 3

CONVOLUTIONAL NEURAL NETWORKS

CHAPTER 4

CARDEA

4.1 System Design

4.2 Implementation

4.3 Evaluation

CHAPTER 5

CONCLUSION AND FUTURE WORK

Bibliography

[1] *Number of surveillance cameras per thousand people by country*. URL: http://www.statista.com/statistics/484956/number-of-surveillance-cameras-per-thousand-people-by-country/.

[2] Heather Kelly. *iPhone photography is cool, eyeball photography is cooler*. URL: http://money.cnn.com/2016/05/12/technology/eyeball-camera-contact-sony.

[3] Tara Seals. *Popular Android Camera App Leaks Sensitive Data*. URL: http://www.infosecurity-magazine.com/news/popular-android-camera-app-leaks/.

[4] Jiwon Kim. *Awesome Deep Vision*. URL: http://github.com/kjw0612/awesome-deep-vision.

[5] Yaniv Taigman et al. “Deepface: Closing the gap to human-level performance in face verification”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2014, pp. 1701–1708.

[6] Yi Sun et al. “Deepid3: Face recognition with very deep neural networks”. In: *arXiv preprint arXiv:1502.00873* (2015).

[7] Florian Schroff, Dmitry Kalenichenko, and James Philbin. “Facenet: A unified embedding for face recognition and clustering”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015, pp. 815–823.

[8] Tobias Weyand, Ilya Kostrikov, and James Philbin. “Planet-photo geolocation with convolutional neural networks”. In: *arXiv preprint arXiv:1602.05314* (2016).

[9] Olga Russakovsky et al. “Imagenet large scale visual recognition challenge”. In: *International Journal of Computer Vision* 115.3 (2015), pp. 211–252.

[10] Alessandro Acquisti, Ralph Gross, and Fred Stutzman. “Face recognition and privacy in the age of augmented reality”. In: *Journal of Privacy and Confidentiality* 6.2 (2014), p. 1.

[11] Alison Spiegel. *Lost Lake Cafe, Seattle Restaurant, Kicks Out Patron For Wearing Google Glass*. URL: http://www.huffingtonpost.com/2013/11/27/lost-lake-cafe-google-glass_n_4350039.html.

[12] Jeremy Schiff et al. “Respectful cameras: Detecting visual markers in real-time to address privacy concerns”. In: *Protecting Privacy in Video Surveillance*. Springer, 2009, pp. 65–89.

[13] Cheng Bo et al. “Privacy. tag: Privacy concern expressed and respected”. In: *Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems*. ACM. 2014, pp. 163–176.

[14] Franziska Roesner et al. “World-driven access control for continuous sensing”. In: *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*. ACM. 2014, pp. 1169–1181.

[15] Paarijaat Aditya et al. “I-pic: A platform for privacy-compliant image capture”. In: *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services (New York, NY, USA, 2016), MobiSys*. Vol. 16. 2016.

[16] Nisarg Raval et al. “What You Mark is What Apps See”. In: *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services (New York, NY, USA, 2016), MobiSys*. Vol. 16. 2016.

[17] Roberto Hoyle et al. “Sensitive lifelogs: A privacy analysis of photos from wearable cameras”. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM. 2015, pp. 1645–1648.

[18] Roberto Hoyle et al. “Privacy behaviors of lifeloggers using wearable cameras”. In: *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM. 2014, pp. 571–582.

[19] Mohammed Korayem et al. “Screenavoider: Protecting computer screens from ubiquitous cameras”. In: *arXiv preprint arXiv:1412.0008* (2014).

[20] Robert Templeman et al. “PlaceAvoider: Steering First-Person Cameras away from Sensitive Spaces.” In: *NDSS*. 2014.

[21] Jaeyeon Jung and Matthai Philipose. “Courteous glass”. In: *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*. ACM. 2014, pp. 1307–1312.

[22] Tamara Denning, Zakariya Dehlawi, and Tadayoshi Kohno. “In situ with bystanders of augmented reality glasses: Perspectives on recording and privacy-mediating technologies”. In: *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*. ACM. 2014, pp. 2377–2386.

[23] Suman Jana, Arvind Narayanan, and Vitaly Shmatikov. "A Scanner Darkly: Protecting user privacy from perceptual applications". In: *Security and Privacy (SP), 2013 IEEE Symposium on*. IEEE. 2013, pp. 349–363.

[24] Suman Jana et al. "Enabling fine-grained permissions for augmented reality applications with recognizers". In: *Presented as part of the 22nd USENIX Security Symposium (USENIX Security 13)*. 2013, pp. 415–430.

[25] Kenta Chinomi et al. "PriSurv: privacy protected video surveillance system using adaptive visual abstraction". In: *International Conference on Multimedia Modeling*. Springer. 2008, pp. 144–154.

[26] J Alex Halderman, Brent Waters, and Edward W Felten. "Privacy management for portable recording devices". In: *Proceedings of the 2004 ACM workshop on Privacy in the electronic society*. ACM. 2004, pp. 16–24.