Privacy for Security Monitoring Systems

Kambiz Ghazinour
Advance Information Security and Privacy Lab,
Department of Computer Science
Kent State University, OH
kghazino@kent.edu

Emil Shirima
Advance Information Security and Privacy Lab,
Department of Computer Science
Kent State University, OH
eshirima@kent.edu

ABSTRACT

With cameras slowly becoming an increasingly daily part of our lives and constantly streaming information about people, different security and privacy concerns arise. Human analysis using cameras or surveillance footage has been an area of research for many years. Different methods have been introduced which showed success in not only detecting but tracking pedestrians. Once a human is detected and tracked, different motion analyses can be performed in helping understand and better model human behavior. A majority of these methods do not take user privacy or security into account, hence these security monitoring systems become a huge threat to the privacy of individuals. This threat becomes even more serious when the security cameras are installed in day-cares, schools, and retirement homes, which contain people who are more vulnerable. With our work, we present an initial thought for a framework that is able to understand human motion, but also take the individual's privacy and security into account by deploying different anonymization techniques.

1. BACKGROUND

Security cameras are widely used to monitor and protect the environment they are observing. This technology is now easily available in different devices and even doorbells are equipped with cameras. The new research has been focusing on transforming security cameras to smart monitoring systems that can detect, recognize and even trigger certain actions. For example, installing a smart monitoring system in a school can potentially detect an intruder or shooter and alarm the law-enforcement officers. While the idea behind installing security cameras are easily justifiable, there is a growing concern that these devices are recording and recognizing people and the data can be used for purposes that the individuals do not necessarily agree with [23] such as what time a person goes to work, or which store the individuals are visiting, etc.

Human detection, motion tracking, and understanding behavior have all been areas of great interest to researchers for the past decade especially with the ongoing rise in security concerns in public areas. There are many research works in which traditional methods used to help model human behavior [1, 2, 3, 4, 5]. In one form or another, these works try to extract local features from a given scene and create a model that will produce a correct classification of the given action. One major drawback to these methods is they are hand-crafted features. A standard hand-crafted feature extractor consists of: (i) feature extraction [1, 6, 7, 8, 9, 10] (ii) feature encoding [11, 12, 13] and (iii) feature classification [11, 14].

In recent years, with the success of convolutional neural networks (CNNs) [15] over the hand-crafted methodologies, deeper features are able to be extracted and represent both the spatial and temporal features for human action recognition [16, 17, 18, 19, 20].

Despite the success of these models, they mostly disregard user privacy and security. Face masking is the most common technique for user anonymization through cameras [21]. As shown in [21] complete blockage is commonly used which completely preserves user privacy but trades off feed usability. On the other hand, semi-anonymization [22] can be used which tries to find a balance between privacy preservation and feed usefulness.

2. PROPOSED FRAMEWORK

A majority of anonymization research is concentrated around data, video and image with very little work done on human behavior anonymization. A lot of information can be extracted by simply observing one's behavior. For example, by just looking at someone's motion and body rhythm, one can easily tell whether they are disabled or not. Couple this with other information such as face, height, and appearance, and the individual's identity can be easily revealed.

We propose a two-fold framework that is able to: (i) detect and analyze human behavior, (ii) anonymize identity and privacy revealing behaviors such as motion and physical appearance.

2.1.Human Behavior Detection/Analysis

The first part of our framework is able to detect normal human behavior. The KTH dataset is used for model training. For a given video feed, we are able to predict (with bounding boxes), the location of a respective action. The

USENIX Symposium on Usable Privacy and Security (SOUPS) 2018. August 12 -- 14, 2018, Baltimore, MD, USA.

bounded region is then fed as input to the anonymization model.

2.2.Human Behavior Anonymization

This network is responsible for detecting motion abnormalities and anonymize them if necessary. This task falls under two main categories:

(i) All motions categorized as being normal are anonymized in order to protect user privacy.
(ii) Motions which resemble the concealment or revelation of objects are not anonymized and are referred to as extreme abnormalities. Such motions tend to closely resemble weapon or criminal related incidents thus making them hard to truly distinguish.

REFERENCES

1. Laptev, Ivan. "On Space-Time Interest Points." *International Journal of Computer Vision*, vol. 64, no. 2-3, Sept. 2005, pp. 107–123., doi:10.1007/s11263-005-1838-7.

2. Roshtkhari, Mehrsan Javan, and Martin D. Levine. "An on-Line, Real-Time Learning Method for Detecting Anomalies in Videos Using Spatio-Temporal Compositions." *Computer Vision and Image Understanding*, vol. 117, no. 10, Oct. 2013, pp. 1436–1452., doi:10.1016/j.cviu.2013.06.007.

3. Boiman, Oren, and Michal Irani. "Detecting Irregularities in Images and in Video." *International Journal of Computer Vision*, vol. 74, no. 1, 3 Aug. 2007, pp. 17–31., doi:10.1007/s11263-006-0009-9.

4. Mahadevan, Vijay, et al. "Anomaly Detection in Crowded Scenes." 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 5 Aug. 2010, doi:10.1109/cvpr.2010.5539872.

5. Zhong, Hua, et al. "Detecting Unusual Activity in Video." Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004., doi:10.1109/cvpr.2004.1315249.

6. Dalal, N., and B. Triggs. "Histograms of Oriented Gradients for Human Detection." 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 25 July 2005, doi:10.1109/cvpr.2005.177.

7. Laptev, Ivan, et al. "Learning Realistic Human Actions from Movies." *2008 IEEE Conference on Computer Vision and Pattern Recognition*, 5 Aug. 2008, doi:10.1109/cvpr.2008.4587756.

8. Klaeser, A., et al. "A Spatio-Temporal Descriptor Based on 3D-Gradients." *Procedings of the British Machine Vision Conference 2008*, Sept. 2008, doi:10.5244/c.22.99.

9. Scovanner, Paul, et al. "A 3-Dimensional Sift Descriptor and Its Application to Action Recognition." *Proceedings of the 15th International Conference on Multimedia - MULTIMEDIA '07*, Sept. 2007, doi:10.1145/1291233.1291311.

10. Willems, Geert, et al. "An Efficient Dense and Scale-Invariant Spatio-Temporal Interest Point Detector." *Lecture Notes in Computer Science Computer Vision – ECCV 2008*, 2008, pp. 650–663., doi:10.1007/978-3-540-88688-4_48.

11. Kuehne, H., et al. "HMDB: A Large Video Database for Human Motion Recognition." *2011 International Conference on Computer Vision*, Nov. 2011, doi:10.1109/iccv.2011.6126543.

12. Perronnin, Florent, et al. "Improving the Fisher Kernel for Large-Scale Image Classification." *Computer Vision – ECCV 2010 Lecture Notes in Computer Science*, 2010, pp. 143–156., doi:10.1007/978-3-642-15561-1_11.

13. Jegou, Herve, et al. "Aggregating Local Descriptors into a Compact Image Representation." *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 13 Dec. 2011, doi:10.1109/cvpr.2010.5540039.

14. Wu, Jianxin, et al. "Towards Good Practices for Action Video Encoding." *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 25 Sept. 2014, doi:10.1109/cvpr.2014.330.

15. Krizhevsky, Alex, et al. "ImageNet Classification with Deep Convolutional Neural Networks." *Communications of the ACM*, vol. 60, no. 6, 2017, pp. 84–90., doi:10.1145/3065386.

16. Wang, Liangliang, et al. "Three-Stream CNNs for Action Recognition." *Pattern Recognition Letters*, vol. 92, 1 June 2017, pp. 33–40., doi:10.1016/j.patrec.2017.04.004.

17. Ji, Shuiwang, et al. "3D Convolutional Neural Networks for Human Action Recognition." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 1, 2013, pp. 221–231., doi:10.1109/tpami.2012.59.

18. Feichtenhofer, Christoph, et al. "Convolutional Two-Stream Network Fusion for Video Action Recognition." *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, doi:10.1109/cvpr.2016.213.

19. K. Simonyan, A. Zisserman, “Two-stream convolutional networks for action recognition in videos.” Adv. Neural Inf. Process. Syst. 1 (2014) 568–576.

20. Tran, Du, et al. “Learning Spatiotemporal Features with 3D Convolutional Networks.” *2015 IEEE International Conference on Computer Vision (ICCV)*, 2015, doi:10.1109/iccv.2015.510.

21. Wang, Junjue, et al. “A Scalable and Privacy-Aware IoT Service for Live Video Analytics.” *Proceedings of the 8th ACM on Multimedia Systems Conference - MMSys'17*, 2017, doi:10.1145/3083187.3083192.

22. Muraki, Tomoya, et al. “Anonymizing Face Images by Using Similarity-Based Metric.” *2013 International Conference on Availability, Reliability and Security*, 7 Nov. 2013, doi:10.1109/ares.2013.68.

23. Marc Weber Tobias, "Is Your Smart Security Camera Protecting Your Home Or Spying On You?", Forbes.com, Aug 22, 2016, available at: https://www.forbes.com/sites/marcwebertobias/2016/08/22/is-your-smart-security-camera-protecting-your-home-or-spying-on-you/#130d050e56dd