VIDEO-BASED LOG GENERATION FOR SECURITY SYSTEMS IN INDOOR SURVEILLANCE SCENARIOS.

David Münch, Stefan Becker, Hilke Kieritz,
Wolfgang Hübner, and Michael Arens

david.muench@iosb.fraunhofer.de
Fraunhofer IOSB, 76275 Ettlingen (Germany)

Abstract

In the EU FP7-SEC-2012-1-313034 project SAWSOC (Situation AWare Security Operations Center) the objective is to achieve "the convergence of physical and logical security technologies, particularly improving correlation techniques across existing technology silos (video surveillance, access control, network monitoring, etc.)".

In this paper two use cases developed in SAWSOC are presented in the perspective of video-based log generation. The first is a critical infrastructure where we log visual observable occurrences in a critical server room. The second is a soccer stadium environment where we log the patrol path of guards and ensure the correct handling of each checkpoint of their patrol path.

Our approach consists of several generic computer vision modules and spatio-temporal data fusion using scene dependent knowledge. Each component of its own does not allow to make any statements about the current situation in the observed area. Instead, the sum of all components has to be considered.

Keywords: SAWSOC, visual log-file generation, security system, indoor visual surveillance.

1 INTRODUCTION

Security services in the physical and in the logical domain have been undergone much progress in the last decades. Services in the physical domain are e.g. video surveillance or identity and credential management software systems. Whereas monitoring network traffic and logging (computer) systems' events are to be considered in the logical domain. Both, the physical and the logical services are stable and mature. Unfortunately, they have been developed independently of each other and do not make use of each other. All of these services are monitored in Security Operation Centers (SOC). Up to now, SOCs have been created for diversified and tailored needs, with independent approaches for physical and logical security purposes.

1.1 The SAWSOC project

In the EU FP7-SEC-2012-1-313034 project SAWSOC (Situation AWare Security Operations Center) the objective is to overcome these disadvantages mentioned above. The main objective is to "identify, implement, and validate techniques for achieving the convergence of physical and logical security technologies, particularly improving correlation techniques across existing technology silos (video surveillance, access control, biometrics, security information and event management (SIEM), network monitoring, Business Activity Monitoring and Business Process Monitoring, etc.)". The SAWSOC project started on 1st November 2013 and will ended on 30th April 2016. The consortium consists of different companies and research institutions from Italy, Finland, Germany, Ireland, Israel, Poland, and the UK.

1.2 Use cases

Three prototypical use cases have been defined. The first one is a critical infrastructure for air traffic management in Italy. There, SAWSOC will protect the air traffic control infrastructure from malicious attacks. Additionally, inside jobs from companies' employees will be considered, too. The second use case is a critical infrastructure for energy production and distribution in Israel. There, SAWSOC will be applied on top of Supervisory Control and Data Acquisition (SCADA) systems. Last, a soccer stadium in Krakow, Poland, will be protected.

1.3 Summary and contribution

As a partner of the SAWSOC project, we provide video surveillance modules. They are more than basic image processing methods; they are context sensitive video analysis comprising of different levels from image processing, to scene understanding, and up to using background knowledge.

In this paper we present the advanced video surveillance modules for the SAWSOC project; in detail these are person detection and tracking, change detection, and skin detection as basis for information fusion applied in the scenario of infrastructure manipulation detection and in the scenario of the soccer stadium.

1.4 Related work

The survey papers [1-5] address the detection of meaningful events and situations in video surveillance. We can distinguish between two different strategies in extracting semantic information from raw videos. One kind are direct approaches. They use massive training data and methods from machine learning. They work as black box approaches: As input data the videos are fed in and the recognized events and situations are the results. The detailed comprehension of the underlying decision function is difficult. Nevertheless, direct approaches are powerful if there is enough training data. The other kind of approaches are hierarchical approaches, where the problem is partitioned into different layers in which different subproblems are addressed. The underlying formal model in hierarchical approaches goes from probabilistic graphical models to formal grammar and finally to description-based approaches such as higher order formal logic. Probabilistic graphical models are powerful but inference is complex, formal grammars are less flexible but easy to model. In contrast to these complex approaches, it is often handy to use established and easier usable methods, such as modelling background knowledge using a basic deterministic finite automaton.

Fig. 1. Example views from the first sequence of the infrastructure manipulation use case. Each camera has a different view, which allows the detection of different events in each view.

2 VIDEO SURVEILLANCE MODULES

As physical manipulation of infrastructure is assumed to be done by people, the core component is a person detector and tracker. This component is based on offline learning of a general person model and online learning of the individual person's model [7]. A second component is a modified version of an abandoned object detection method to detect physical changes in the environment caused by a present person. We use the approach of Becker [8] with an IMM filter setup as presented in [9]. Finally, basic skin detection or door-status detection methods support the visual observable events. The video surveillance modules are described in detail in [10].

3 USE CASES

In this section two use cases from the SAWSOC project are addressed in the perspective of video-based log generation.

3.1 Infrastructure manipulation

In this section, a prototypical infrastructure manipulation setting – one of SAWSOC's use cases – is described. When running a critical infrastructure, special attention is needed to log the panned maintenance or to detect suspicious events. To provide data for this use case, we recorded four different scenes with four different persons using four synchronized cameras in a critical server room environment.

Each person is performing all or some of the subsequent five basic events:

1. Enter server room.

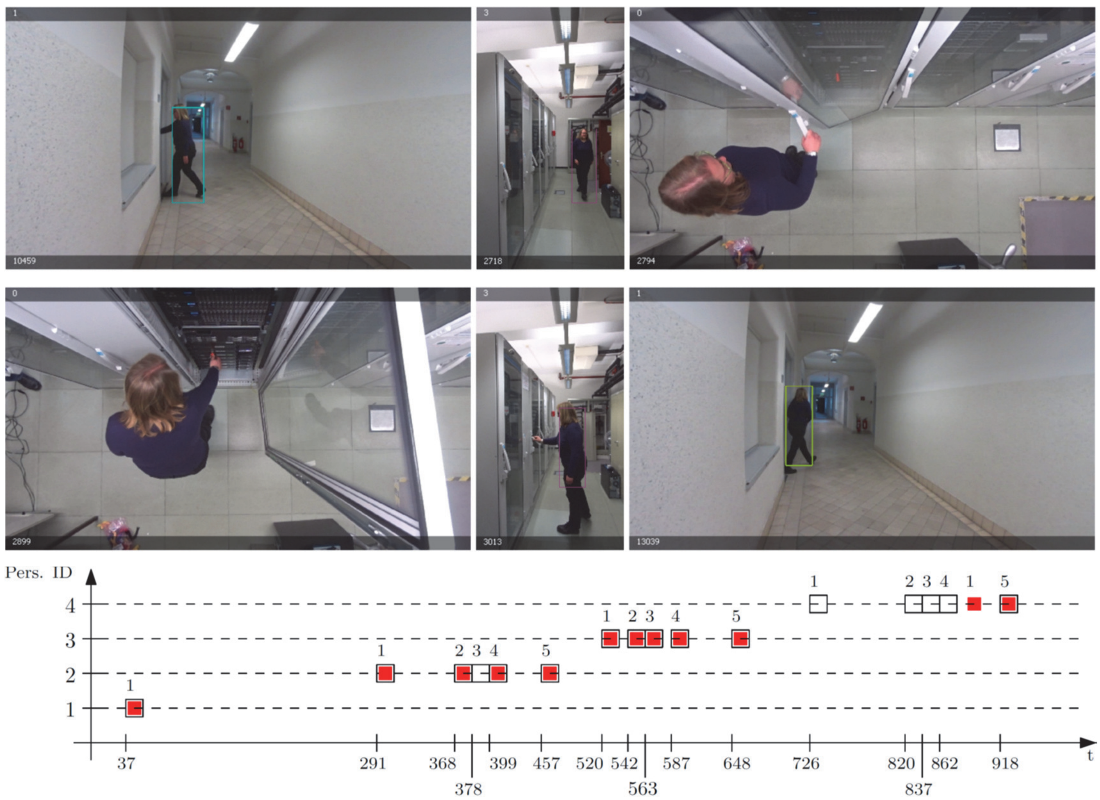

Fig. 2. Qualitative (top) and quantitative (bottom) evaluation in the first scene. The images show the events "enter the server room", "open server door", "server manipulation", "close server door", and "exit server room". Unfilled square stands for groundtruth, a filled square represents a detection. From [6].

2. Open server door.
3. Server manipulation.
4. Close server door.
5. Exit server room.

In Figure 1 some snapshots of the first sequence of our recordings are shown. The four cameras have different views, one is in the aisle in front of the server room, one is on the ceiling, one is in front of the servers, and one is parallel according to the servers. The precise evaluation of this use case can be found in [6]. The deterministic finite automaton (see Figure 5, [6]) is applied on top of the video surveillance events. In Figure 2 the results of the first scenes are shown. It can be seen that each time the person enters the server room again, it gets assigned a new person ID. Because no person identification is done, it is valid that the tracker assigns a new person ID for every new person track. At time 378 the server manipulation event was missed. This was due to only opening the server door very little, so that the cameras could see only the server door there. Later, at time 726 ff., the person was missed by the person detector, thus, all modules building upon this information will fail, too. In conclusion it can be said, that the sum of video surveillance events and moderate knowledge encoded in a deterministic finite automaton allows generating more complex events.

3.2 Security guard patrol

In this section the SAWSOC use case in the soccer stadium is addressed. To ensure security, guards have to walk designated patrol paths in the inside of the stadium. In this use case we log the patrol path of a guard and ensure the correct handling of each checkpoint of the patrol path. The soccer stadium's camera coverage is huge, but not every area is covered in detail. Thus we use the person detector and tracker mentioned above to find persons in the different cameras' images and then use additional scene information to detect the guards in the checkpoint regions. The positions of each guard during its patrol path are logged and deviations are reported. Background knowledge about the ground plane is added to improve the runtime of the person detector and tracker.

The soccer stadium dataset was recorded between 11/2015 and 02/2016 in the Cracovia stadium with cooperation of the Cracovia Football Club and Comarch (T. Grabowski, J. Szczebak). The dataset consists of video recordings with fourteen cameras covering security guards patrol path through the hallway inside the stadium. The entire dataset covers six different scenes, in which the security guards act differently. Figure 3 depicts a security guard (orange) together with an ordinary person.

Fig. 3. A security guard and an ordinary person are walking along the hallside inside the soccer stadium. The guard is detected as guard and its position gets logged.

4 CONCLUSION

In this paper, we summarized in short the SAWSOC project and its main goals. Then, we addressed the video surveillance modules as generic and specific building block. Finally, all the modules were put together to show their strength in two use cases: In the infrastructure manipulation use case and in the security guard use case. We used existing technologies to build more complex systems and the results show that the whole is more than the sum of its parts. Future work after the end of the SAWSOC project is the integration of the video surveillance modules into the products enabled by SAWSOC.

ACKNOWLEDGMENTS

This work is supported in part by the European Unions Seventh Framework Programme for research, technological development and demonstration under grant agreement FP7-SEC-2012-1-313034 (SAWSOC-Project).

REFERENCES

[1] Turaga, P., R. Chellappa, V. S. Subrahmanian, and O. Udrea (2008). Machine Recognition of Human Activities: A Survey. IEEE Transactions on Circuits and Systems for Video Technology.

[2] Lavee, G., E. Rivlin, and M. Rudzsky (2009). Understanding Video Events: A Survey of Methods for Automatic Interpretation of Semantic Occurrences in Video. IEEE Transactions on System, Man, and Cybernetics, Part C: Applications and Reviews.

[3] Aggarwal, J. K. and M. S. Ryoo (2011). Human Activity Analysis: A Review. ACM Computing Surveys.

[4] Vishwakarma, S. and A. Agrawal (2012). A survey on activity recognition and behavior understanding in video surveillance. The Visual Computer.

[5] Ye, Juan, Simon Dobson, and Susan McKeever (2012). Situation identification techniques in pervasive computing: A review. Pervasive Mob. Comput.

[6] Münch, D., B. Hilsenbeck, H. Kieritz, S. Becker, A.-K. Grosselfinger, W. Hübner, M. Arens. (2016) Detection of infrastructure manipulation with knowledge-based video surveillance. Proc. of the Security and Defence Conference SPIE 2016, Edinburgh, UK.

[7] Kieritz, H., W. Hübner, and M. Arens (2013). *Learning transmodal person detectors from single spectral training sets.* SPIE 8901A Optics and Photonics for Counterterrorism, Crime Fighting and Defence.

[8] Becker, S., D. Münch, H. Kieritz, W. Hübner, and M. Arens (2015). Detecting Abandoned Objects using Interacting Multiple Models. Proc. SPIE Volume 9652 Optics and Photonics for Counterterrorism, Crime Fighting, and Defence.

[9] Becker, S., H. Kieritz, W. Hübner, and M. Arens (2016). On the benefit of state separation for tracking in image space with an Interacting Multiple Model Filter. Prof. of 7th International Conference on Image and Signal Processing (ICISP), Trois-Rivières, QC, Canada.

[10] Münch, D., S. Becker, H. Kieritz, W. Hübner, M. Arens (2015). Knowledge-based Situational Analysis of Unusial Events in Public Places. Proc. of the 10th Future Security Research Conference, Berlin, Germany.