

Finding suspects in multiple cameras for improved railway protection

Jan Willem Marck¹, Henri Bouma, Jan Baan, Julio de Oliveira Filho, Mark van den Brink

TNO, Oude Waalsdorperweg 63, 2597 AK The Hague, The Netherlands

ABSTRACT

The capability to find individuals using CCTV cameras is important for surveillance applications at large areas such as railway stations, airports and shopping centers. However, it is laborious to track and trace people over multiple cameras post incident. In this paper, we describe the live demonstration of our interactive re-identification system in a railway station. The system performs real-time track generation in multiple cameras and live re-identification and refinding of suspects which was live demonstrated in Poland. The system allows fast interactive retrieval of an individual by showing only similar candidates. An operator can find the origin or destination of a person more efficiently, especially over large time and space intervals.

Keywords: Surveillance, CCTV, security, tracking, left luggage, behavior analysis, forensic search, threat, railway.

1. INTRODUCTION

Railway stations, airports and shopping centers are subject to large impact threats such as explosives in left-luggage. Frequently, these locations are also stage for small criminal activities, such as graffiti, vandalism, theft from passengers, physical and verbal abuse against passengers and staff, sexual harassment, and trespassing [8][9]. The negative impact of these events can be minimized, and the suspect can be held accountable if he can be found quickly after the incident. Localizing suspects after incidents requires time from the operator and careful attention. Especially if the operator loses track of the suspect – e.g. he enters a shop and leaves through another exit – and he has to find the individual at another time and location. Support for localizing the suspect quicker and more efficiently is therefore of economic value for security operators in the rail station, airport or retail setting.

This paper describes a demonstration of our tracking system at a railway station in Poland. Our tracking technology supports an operator in finding suspects interactively after an event occurs, with proven operator-efficiency and speed improvement [5]. The integration and demonstration of this system was done in the project PROTECTRAIL [<http://www.PROTECTRAIL.eu>], which is a Railway-Industry Partnership for Integrated Security of Rail Transport. PROTECTRAIL is an integration project funded by the European Union under the 7th framework programme (FP7). The aim of the project is to provide an integrated and economically/technically viable set of railway security solutions, and to make these solutions interoperable by designing a modular architectural framework in which each solution can be easily “plugged”. The consortium has about 30 partners including rail operators, technology providers and research institutes. Among other tasks in this project, TNO has provided the security operator with a support system for tracking individuals. This system was live demonstrated to a group of stakeholders in October 2013 in Zmigrod, Poland.

The outline of the paper is as follows. The method is described in Section 2. The experiments and results are shown in Section 3. Finally, the conclusions are presented in Section 4.

2. METHOD

2.1 System overview and interface

The TNO system for fast interactive retrieval of an individual is integrated in the PROTECTRAIL system, which includes multiple security operator control centers (SOCC's), an event broker and other integrative infrastructure. In the PROTECTRAIL System many different security systems provided different events (e.g. an alarm from motion detector).

¹ jan_willem.marck@tno.nl; phone +31 888 66 1024; <http://www.tno.nl>

An event broker allowed services to subscribe to specific events as is common in Service Oriented Architecture. The security operator in the SOCC coordinates, based on the events, the actions of security personnel. Figure 1 shows the TNO system and its relevant environment in the PROTECTRAIL system. The TNO system interfaced with a number of IP cameras using RTSP/ONVIF standards [21][25] and it generated track information which, through event brokering, was visualized on the SOCC. The event brokering was implemented using WSDL [11] for the infrastructure and OASIS CAP [3] to structure the XML alert message formats. Using both formats resulted in an easy and efficient plug-and-play integration with the rest of the PROTECTRAIL System. This resulted in a system which ran parallel to the SOCC systems and when needed provided accurate and up-to-date location information of the individual of interest to the SOCC.

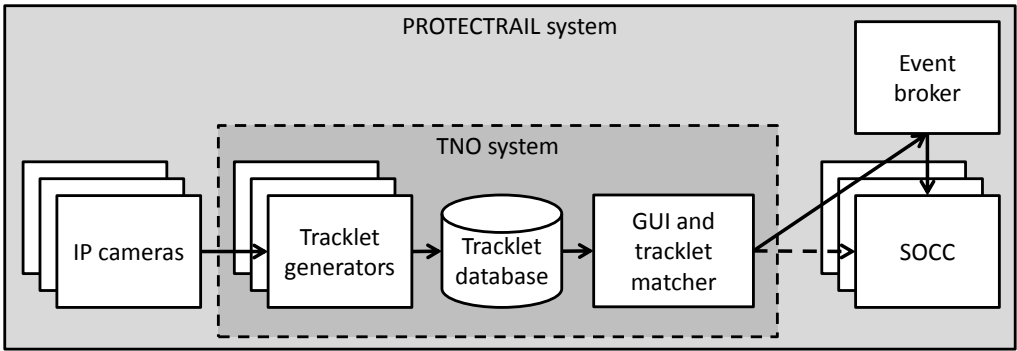


Figure 1: Overview of the TNO system and related components in the PROTECTRAIL system.

2.2 Real-time tracking

The TNO person tracking system performs real-time track generation of every detected pedestrian in every camera [5], as shown in Figure 2. Currently, one PC is used for every camera to perform the person detection and track generation. A scalable system would require many PC's, or sufficient computational power inside the camera, or a more flexible assignment of computational resources that is adapting to the task of the operator. The generated track contains information about position, time, bounding boxes, one small image to visualize the candidate and an appearance-based histogram model [6]. The track information is stored in a database that is used for re-identification of the individual.



Figure 2: All people are tracked in every camera. The tracks are stored together with a small image of the candidate and an appearance-based representation that is use to match similar people.

2.3 Re-identification engine

Person re-identification is an active research field [14][16][18], with several recent developments [1][2][24]. The re-identification engine can be used to combine observations of multiple operators or multiple (explosive) detectors to obtain a merged and fused result [7], and it can be used to search people and track them. The re-identification engine in our system [6] is used to retrieve candidates that are similar to the person of interest (e.g., the suspect). This allows a user to answer questions more quickly that are related to the origin and destination of the track. For every track, an appearance-based histogram model is stored. The model of the suspect is compared to all other candidates by the re-identification engine, and only the most similar candidates are shown to the operator. Experiments on the international VIPER benchmark dataset [16][17] showed that the system allows the human to retrieve individuals five times faster than without this functionality [6].

2.4 Graphical user interface

The graphical user interface (GUI) is shown in Figure 5. It includes a map, a spot view, camera selection and a candidate view. The candidate view has a horizontal axis for time and a vertical axis for the cameras. On historic data, a complete day can be loaded with a fixed number of candidates, and an operator can move forward and backward in time. During live demonstration, the number of displayed candidates is extended and novel incoming candidates added on the screen. A scalable GUI would use the vertical axis for presenting candidates that are ranked from similar (top) to less similar (bottom), without showing fixed cameras on the vertical axis. Experiments with human operators showed that the system leads to a significant reduction of missed people by using the interactive functionality, especially for large time intervals [5]. The tracking system was integrated and validated in a crowded [13][19] shopping mall in the Netherlands, in collaboration with five SMEs [4]. The system performed consistently and robustly. This performance translated well to the new PROTECTRAIL test environment in Zmigrod, Poland by adapting the interfaces to connect to other PROTECTRAIL components and without modification of the core technology.

3. EXPERIMENTS AND RESULTS

3.1 Experimental setup at railway station in Poland

The PROTECTRAIL demonstration was organized in October 2013 in Zmigrod, Poland. This site consists of a small train station and a test railway track, as shown in Figure 3. The aim of the PROTECTRAIL demonstration was to show in real conditions new innovative and also working technologies with added value to operators of the technologies and especially the integration there-of. One of the demonstration scenarios focused on the consequences of left luggage. Part of the demonstration scenario was the TNO system, which was used to find the current location and origin of a suspect that left luggage in a crowd.



Figure 3: The PROTECTRAIL demonstration was given in Zmigrod in October 2013 at a small station. A photo of the railway station is shown left and the map of the test railway track is shown right.

The scenario comprises the following steps.

1. Left luggage was reported by a traveler to the SOCC.
2. The SOCC investigated the call and sends a mitigation team and simultaneously starts the TNO system to locate as soon as possible the person who left the luggage.
3. The mitigation team deals with the explosive bag.
4. The TNO system locates the individual, who has reappeared on the scene. Security personnel is informed.
5. The individual is captured for further investigations.

The suspect (an actor) was continuously surrounded by a crowd of 50 students (Figure 4). Both the crowd and suspect are wearing normal and natural clothing. The demonstrations were performed under normal weather condition, which included light variations, such as strong shearing sunlight and cloudy conditions.



Figure 4: The suspect (indicated by an arrow) is surrounded by a small crowd of 50 volunteers.

3.2 Results at railway station in Poland

The TNO system initiated where the luggage was found (Figure 5). In the demonstration scenario a human reports the suspicious luggage, although automatic left-luggage detection is becoming feasible [15][28]. Figure 5 shows a vertical line in the candidate view indicating the 'current' time. This screen shot was generated with recorded video data. The luggage was left by one of the persons in the crowd. This crowd can be recognized as a group of candidates that is present before the 'current' time. We can also see the motion of the group from camera 5 (bottom) towards camera 1 (top), and then a small group leaving camera 1 before the current time. Interaction with the time line is used to find the person who left the luggage with a few mouse-clicks.



Figure 5: Screenshot of the graphical user interface (GUI), which includes a map (top left), spot view (top middle), camera selection (top right) and candidate view (bottom). The candidate view has a horizontal axis for time and a vertical axis for the cameras. The spot view shows left luggage.

Once the person is identified that left the luggage, the re-identification engine is used to show only similar candidates (Figure 6). This allows the operator to answer the question: “where did the suspect go to?”. The suspect left the scene by entering the train. However, the train is on the test track and after a few minutes the train arrived again at the station. When the suspect reappears on the scene, he is detected by the TNO system. At this moment, security personnel is informed and the suspect is caught (Figure 6).

The main advantage of the tool is that it allows the operator to easily find the origin or destination of the person. This functionality is especially useful to find the same person over a large time interval. Figure 7 shows that the suspect already inspected the scene in the morning. The quick retrieval is possible because only similar candidates are shown. When all candidates are shown (Figure 8), it is much harder to find the person of interest. An additional benefit is that in the candidate view it is easily visible when and where people are. In the white areas, the cameras do not register people. This is very useful and intuitive information for the operator in case of manual or semi-automatic search.

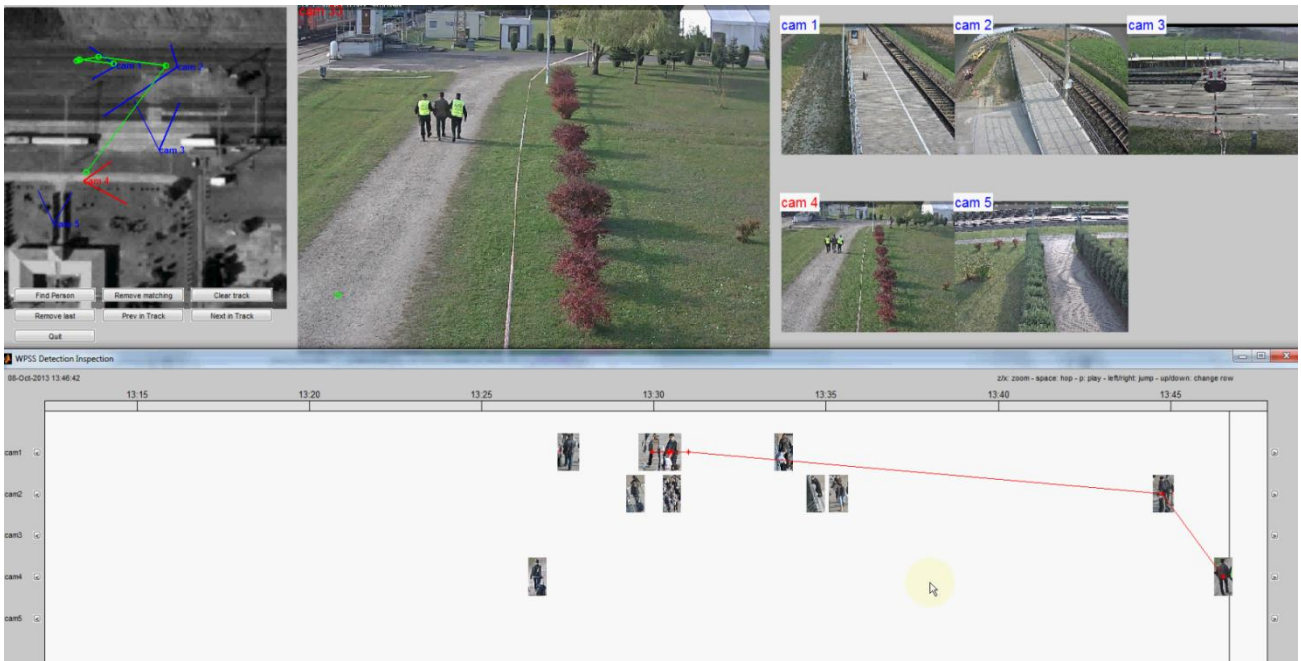


Figure 6: Many candidates are suppressed and only similar candidates are shown to find the current location of the suspect that left the luggage. The spot view shows that the suspect is caught after reappearing on the scene.

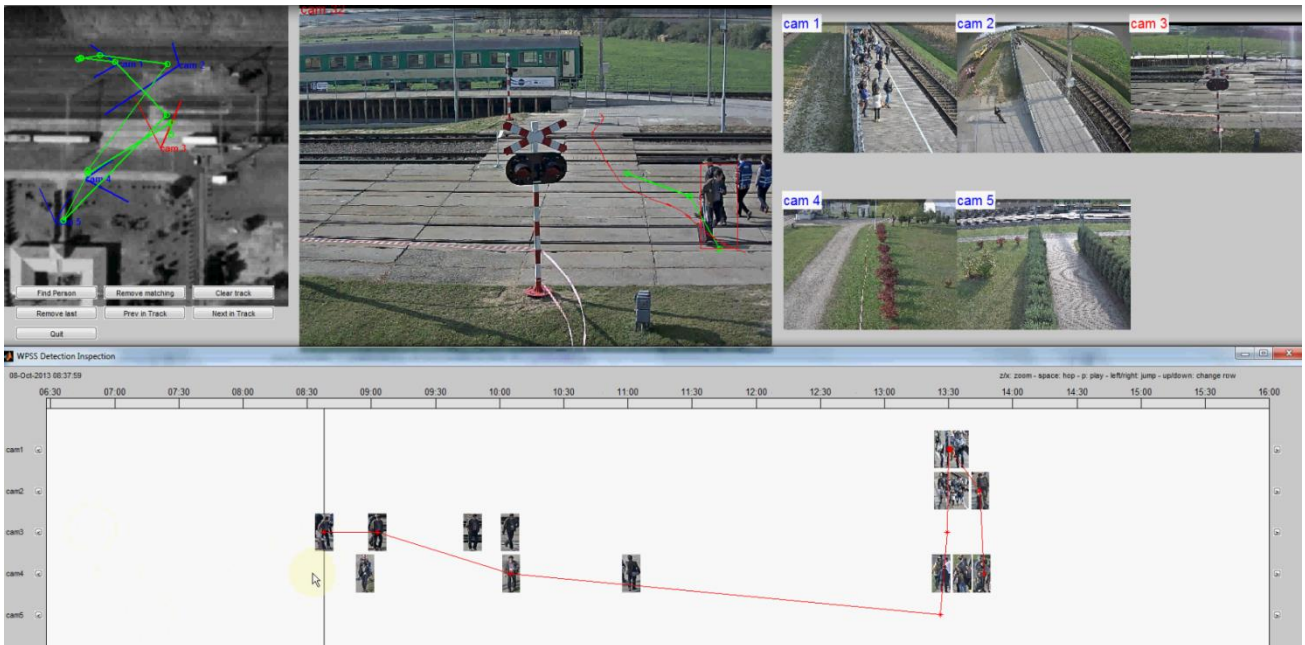


Figure 7: The main advantage of the tool is that it allows the operator to easily answer questions like “where did the suspect come from?” or “where did he go to?”. The functionality is especially useful to find the same person over a large time interval. The figure shows that the suspect already inspected the scene in the morning.

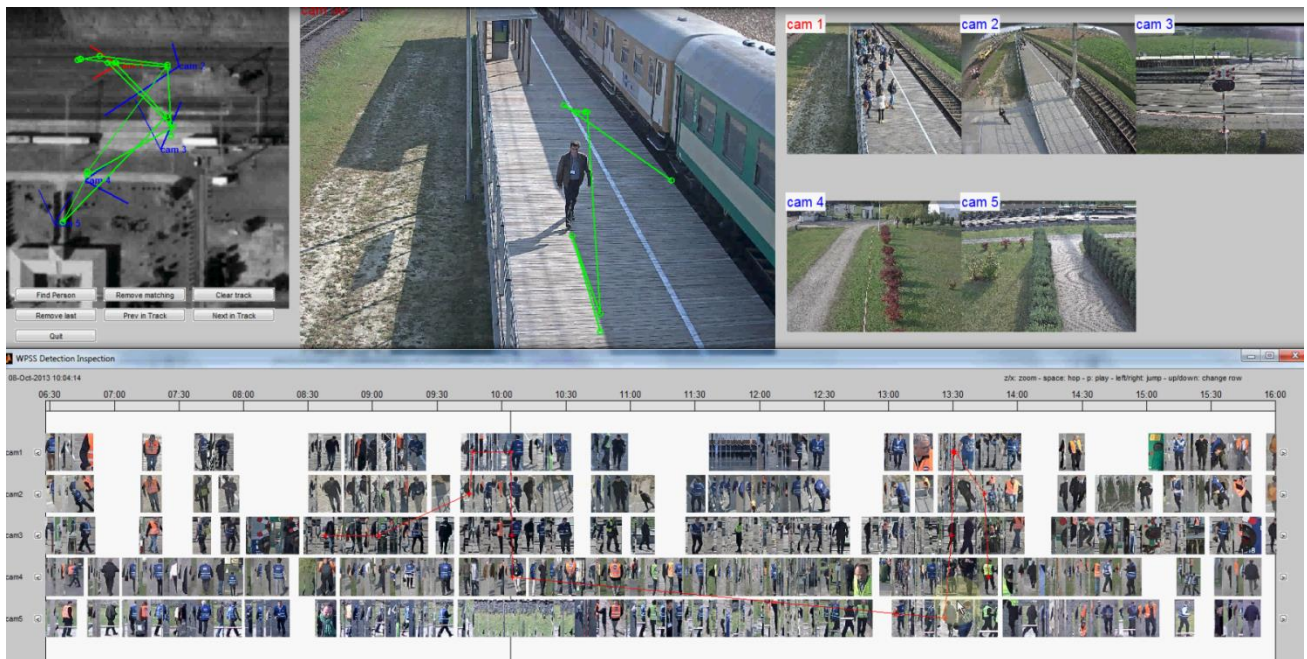


Figure 8: When all candidates are shown, it is much harder to find the person of interest.

4. CONCLUSIONS

We presented a system for real-time tracking and fast interactive retrieval of persons in video streams from multiple static surveillance cameras. The system performs real-time track generation in multiple cameras and live extension of candidates as was live demonstrated in Poland. The system provides clear benefits for retrieving a person's past or current location, by providing A) a semi-automatic search (by showing only candidates that are similar to the suspect) and B) an overview of where and when pedestrians are. An operator can find the origin or destination of a person more efficiently with fewer misses due to e.g. toilet/shop visits, especially over large time and space intervals.

ACKNOWLEDGEMENT

The work for this paper was supported by the EU FP7 project PROTECTRAIL and the project WPSS.

REFERENCES

- [1] An, L., Kafai, M., Yang, S., Bhanu, B., "Reference-based person re-identification," IEEE AVSS, (2013).
- [2] Bialkowski, A., Denman, S., Sridharan, S., Fookes, C., Lucey, P., "A database for person re-identification in multi-camera surveillance networks," IEEE DICTA, (2012).
- [3] Botts, M., Percivall, G., Reed, C., Davidson, J., "OGC sensor web enablement: Overview and high level architecture," GeoSensor networks LNCS 4540, 175-190 (2008).
- [4] Bouma, H., Baan, J., Borsboom, S., Zon, K., Luo, X., Loke, B., Stoeller, B., Kuilenburg, H., Dijk, J., "WPSS: Watching people security services," Proc. SPIE 8901, (2013).
- [5] Bouma, H., Baan, J., Landsmeer, S., Kruszynski, C., Antwerpen, G., Dijk, J., "Real-time tracking and fast retrieval of persons in multiple surveillance cameras of a shopping mall," Proc. SPIE 8756, (2013).
- [6] Bouma, H., Borsboom, S., Hollander, R., Landsmeer, S., Worring, M., "Re-identification of persons in multi-camera surveillance under varying viewpoints and illumination," Proc. SPIE 8359, (2012).
- [7] Bouma, H., Vogels, J., Aarts, A., e.a., "Behavioral profiling in CCTV cameras by combining multiple subtle suspicious observations of different surveillance operators," Proc. SPIE 8745, (2013).

- [8] Bouma, H., Burghouts, G., Penning, L., et al., "Recognition and localization of relevant human behavior in videos," *Proc. SPIE* 8711, (2013).
- [9] Burghouts, G.J., Schutte, K., Bouma, H., Hollander, R., "Selection of negative samples and two-stage combination of multiple features for action detection in thousands of videos," *Mach. Vision and Appl.*, (2013).
- [10] Burghouts, G., Eendebak, P., Bouma, H., Hove, J.M., "Improved action recognition by combining multiple 2D views in the bag-of-words model," *IEEE AVSS*, 250-255 (2013).
- [11] Christensen, E., Curbera, F., Meredith, G., Weerawarana, S., "Web services description language (WSDL) 1.1," W3 Consortium, (2001).
- [12] Dijk, J., Rieter-Barrell, Y., Rest, J. van, Bouma, H., "Intelligent sensor networks for surveillance," *Journal of Police Studies: Technology-Led Policing* 3(20), 109-125 (2011).
- [13] Fagette, A., Courty, N., Racoceanu, D., Dufour, J.Y., "Unsupervised dense crowd detection by multiscale texture analysis," *Pattern Recognition Letters*, (2013).
- [14] Farenzena, M., Bazzani, L., Perina, A., Murino, V., Cristani, M., "Person re-identification by symmetry-driven accumulation of local features," *IEEE CVPR*, 2360-2367 (2010).
- [15] Ferryman, J., Hogg, D., Sochman, e.a., "Robust abandoned object detection integrating wide area visual surveillance and social context," *Pattern Recognition Letters* 34(7), 789-798 (2013).
- [16] Gray, D., Brennan, S., Tao, H., "Evaluating appearance models for recognition, reacquisition, and tracking," *IEEE Int. Workshop Performance Evaluation of Tracking and Surveillance PETS*, (2007).
- [17] Gray, D., Tao, H., "Viewpoint Invariant Pedestrian Recognition with an Ensemble of Localized Features," *Proc. European Conference on Computer Vision ECCV*, (2008).
- [18] Hamdoun, O., Moutarde, F., Stanculescu, B., e.a., "Person re-identification in multi-camera system by signature based on interest point descriptors collected on short video sequences," *Distr. Smart Cameras*, (2008).
- [19] Hu, N., Bouma, H., Worring, M., "Tracking individuals in surveillance video of a high-density crowd," *Proc. SPIE* 8399, (2012).
- [20] Martinez-del-Rincon, J., Herrero-Jaraba, J. E., Gomez, J. R., Orrite-Urunuela, C., "Automatic left luggage detection and tracking using multi-camera UKF," *IEEE Workshop PETS*, 59-66 (2006).
- [21] ONVIF, "ONVIF Core Specification," version 1.01, (2009).
- [22] Parameswaran, V., Chellappa, R., "View invariance for human action recognition," *IJCV* 66(1), 83-101 (2006).
- [23] Rest, J. van, Grootjen, F., Grootjen, M., et al., "Requirements for multimedia metadata schemes in surveillance applications for security," *Multimedia Tools and Applications MTAP*, (2013).
- [24] Satta, R., "Dissimilarity-based people re-identification and search for intelligent video surveillance," PhD thesis Univ. Cagliari Italy, (2013).
- [25] Schulzrinne, H., Columbia, U., et al., "Real time streaming protocol RTSP," *IETF RfC* 2326, (1998).
- [26] Soori, U., Yuen, P., Han, e.a., "Target recognitions in multiple-camera closed-circuit television using color constancy," *Optical Engineering* 52(4), (2013).
- [27] Thida, M., Eng, H., Monekosso, D., Remagnino, P., "Contextual analysis of videos," *Synthesis Lectures on Image Video and Multimedia Processing* 6(2), (2013).
- [28] Tian, Y., Feris, R. S., Liu, H., Hampapur, A., & Sun, M. T., "Robust detection of abandoned and removed objects in complex surveillance videos," *IEEE Trans. Systems Man and Cybernetics* 41(5), 565-576 (2011).
- [29] Bak, S., Zaidenberg, S., Boulay, B., Bremond, F., "Improving person re-identification by viewpoint cues," *AVSS*, (2014).
- [30] Bedagkar-Gala, A., Shah, S., "A survey of approaches and trends in person re-identification," *Image and Vision Computing* 32(4), 270-286 (2014).
- [31] Bouma, H., Baan, J., Burghouts, G., et al., "Automatic detection of suspicious behavior of pickpockets with track-based features in a shopping mall," *Proc. SPIE* 9253, (2014).
- [32] Burghouts G., Schutte, S., Hove, R., et al., "Instantaneous threat detection based on a semantic representation of activities, zones and trajectories," *Signal Image and Video Processing SIVP*, (2014).
- [33] Cancela, B., Hospedales, T. M., & Gong, S., "Open-world person re-identification by multi-label assignment Inference," *BMVC*, (2014).
- [34] Gong, S., Cristani, M., Loy, C., Hospedales, T., "The re-identification challenge," *Person Re-Identification*, 1-20 (2014).
- [35] Liu, C., Gong, S., Loy, C., "On-the-fly feature importance mining for person re-identification," *Pattern Recognition* 47(4), 1602-1615 (2014).