Activity and Threat Recognition for Commercial Transport

Ronnie Johansson and Maria Andersson
Swedish defence research agency (FOI)

***Abstract*—We are investigating how to build a surveillance system for commercial transport using relatively inexpensive equipment and information processing from individual sensors to recognized threats. In this paper, we report on some results and ongoing work on activity and threat recognition.**

I. Introduction And Background

Commercial transport with trucks is continually exposed to criminal activities to a large extent over whole Europe. Today there is no monitoring system available that can continuously protect the truck and driver from such threats.

Additionally, in recent years, there have been a number of incidents where terror organizations have caused disruption to mass transportation networks and other areas of critical infrastructure. A very real threat is that the same (or other) terror organizations will seek to disrupt the transit of (or to destroy or capture vehicles containing) hazardous or dangerous materials (e.g., chemical liquids or gas), including radioactive (nuclear) material, or simply vehicles of huge economic value.

In an ongoing EU project (EU FP7 ARENA), we are investigating the design of a flexible surveillance system for the deployment on mobile assets, such as on-road commercial transport based on inexpensive optical sensors. We are not aware of any similar approaches in the literature where the full chain from individual images to event detection and threat recognition is investigated. For experimentation and testing, a scenario including a parked truck (the mobile asset) in a parking lot is studied. Based on image data, objects (e.g., people) are extracted, labeled and tracked. Many simple object activities may occur in the parking lot, e.g., people moving about between different locations (such as vehicles and buildings), standing waiting, etc. However, some activities may be indicative of a threat to the truck, such as physical assault to the truck or its driver.

In a previous paper [1], we report on the recognition of various simple activities in video data: *context-specific* activities based on soft-computing methodologies (e.g., person moving towards the asset), *individual action-specific* activities based on Bag-of-Word models (e.g., walking, running, loitering, and entering or exiting a vehicle) and *group-specific* activities based on *K*-means clustering (e.g. merging, splitting and fighting).

Using recognized activities, threats can possibly be recognized too. We identify three approaches to detecting threats: *modus operandi patterns* (known threats formalized as sensor recognizable activities by domain experts), *recall* (recognition of patterns of activities that in historical data have co-occurred with confirmed threats), *and anomalies* (detection of a significant activity deviation from what is expected based on historical data). The two latter approaches rely on available real-world historical data, which is not easily accessible to the project so it is primarily the modus operandi approach that has been pursued in the project.

II. Results

The algorithms for simple activity recognition have been evaluated with real tracking data [1]. Visual sensor data, from one overview camera, have been recorded for 23 realistic scenarios on a truck parking lot. The evaluation shows promising results despite the complicated environment that causes every now and then false detections. For instance, for context-specific activities, the recognition accuracy is 90%, for action-specific activities 71%, and for group-specific activity 66%.

For threat recognition, one of the project partners, TNO, is pursuing a stochastic parser approach [2], where recognized activities act as words in a flow of text generated by the employed sensors and activity recognition. At FOI, we are pursuing the following two directions: Bayesian networks (BN) and Markov logic networks (MLN). Our BN approach uses simple activities such as loitering to estimate the "suspiciousness" of a person. Unlike the stochastic parser, the BN approach cannot discern a particular order of activities, only the co-occurrence of activities which makes it false-positive prone. On the other hand, the time complexity of inference is relatively low and the problem modeling well known. MLNs have the great advantage that the structure of the problem can be expressed explicitly in first-order logic, but its computational requirement is demanding. Results of the threat recognition approaches have yet to be evaluated. Note that it may not necessarily be that a single approach is chosen for the resulting system; the results of all approaches can be displayed in parallel or even fused.

References

[1] Andersson, M., Patino, L., Burghouts, G., Flizikowski, A., Evans, M., Gustafsson, D., Petersson, H., Schutte, K., Ferryman, J., "Activity recognition and localization on a truck parking lot", *The 10th IEEE International Conference on Advanced Video and Signal-based Surveillance* (AVSS 2013), Krakow, 27-30 August, pp. 263-269, 2013.

[2] Sanromà, G., Burghouts, G., Schutte, K., "Recognition of Long-Term Behaviors by Parsing Sequences of Short-Term Actions with a Stochastic Regular Grammar", *Structural, Syntactic, and Statistical Pattern Recognition*, Lecture Notes in Computer Science Volume 7626, pp. 225-233, 2012.