VISUAL SURVEILLANCE TECHNOLOGIES FOR ENHANCING ABC SECURE ZONES

Csaba Beleznai, Stephan Veigl, Michael Rauter, David Schreiber, Andreas Kriechbaum
AIT – Austrian Institute of Technology, 1220 Vienna, Austria
csaba.beleznai@ait.ac.at

Abstract: Traveler flows and crossings at the external borders of the EU are increasing and are expected to increase even more in the future; trends which encompass great challenges for travelers, border guards and the border infrastructure. In this paper we present three non-intrusive, vision-based technologies and research contributions addressing relevant security and efficiency requirements of border check procedures: (i) counting and separating humans within an eGate, (ii) robust left item detection in secure zones and (iii) estimating the queue length and the number involved persons at border crossing.
The proposed computer vision-based technologies will reduce delays and queues for travelers; improve the user experience at the border infrastructure, and at the same time support border guards in achieving a higher level of security by preventing unauthorized border crossings.

***Keywords*:** visual surveillance, ABC secure zones, pedestrian detection, left luggage detection, queue length estimation

INTRODUCTION

The increasing worldwide travel capacities at airports pose new challenges in the area of border and security control. Travelers request a reduction of delays in the immigration process and a convenient, non-intrusive, attractive border crossing, while border guards must fulfill their obligation to secure the EUs borders against illegal immigration, terrorism, crime and other threats. Infrastructure providers demand maximum border crossing throughput and minimal border crossing area. An automated border control system shall accelerate the border control process by increasing the passenger throughput while maintaining the highest level of security.

Motivated by these challenges we propose several key technology elements targeting secure zones of an ABC infrastructure and its users: travelers and operators such as border guards. In line with the proposed 'smart border package' of the European Commission issued in 2013, our goal is to speed-up, facilitate border control processes and reinforce border check procedures at the external borders of the EU.

Our paper provides a task-oriented view on three computer vision based technologies addressing relevant security and efficiency requirements of ABC secure zones such as at airports. At the same time we also provide a detailed description of the proposed methodologies by putting them into scientific context and outline their advantages over existing technologies.

The addressed key challenges and the respective proposed technology components are:

- A vision-based solution to the problem of detecting tailgating/piggybacking events (person tagging along with another person) within eGates of an Automated Border Control infrastructure. The proposed user-friendly system facilitates the work of border guards and saves time for additional measures to prevent illegal immigration.

- A reliable left item detection framework operating within the eGate and providing immediate alerts on left-behind items of various dimensions and appearances.

- A stereo vision based queue length detection framework which can successfully discriminate between waiting passengers and other slowly moving and stationary objects such as carried luggage pieces and other scene objects, also characterizing the dynamics (estimated waiting time) of the queue.

The above vision-based technologies all exploit the advantages of stereo vision based depth sensing, which allows for enhanced visual analysis in terms of more robust object detection, segmentation and tracking. Robustness in this context refers to the improved characteristics that the analysis can well cope with scene illumination variations, shadows and reflections, and occlusions between passenger-passenger and scene objects. Furthermore, the spatial sensing capability of these detection technologies allows for an easy registration into the global spatial context of the ABC infrastructure environment, thus detection results and associated alerts can be spatially referenced with respect to the infrastructure or to an existing surveillance camera network layout.

OVERVIEW

Major factors for security and mobility at airports are secure and efficient border control procedures and flexible management of traveler flows. All travelers wish to cross external borders with maximum convenience and without losing too much time at border controls. At the same time border guards must still fulfill their obligation to secure the EUs borders against illegal immigration, terrorism, crime and other threats.

Vision sensors and associated image analysis provide new means to assess relevant indicators on the presence and flow of passengers and on the specific location they are situated in. Typical observation scenarios in relevant secure zones at borders are complex: high density of passengers, many non-stationary objects (luggage, carts, and dividers) and variable illumination conditions. Traditionally, visual surveillance at borders and airports is a commonly used technology to support the task of border guards, by complementing human vigilance or providing additional information such as the estimated number of persons within an area. Fully automated surveillance, however, is nowadays still in a developing phase where it is increasingly becoming capable to meet the strict accuracy requirements imposed at border crossings.

In this paper we propose the deployment of a depth-sensing stereo camera sensor, which is the result of several years of hardware and software development at the AIT [1]. A real-time stereo matching process [1] outputs depth data, which well represents the scene geometry and it remains invariant with respect to illumination variations and shadows. The combination of this depth information with color image data (originating from one camera of the stereo setup) results in a significant increase in robustness when compared to conventional vision-based solutions. In the following we present the individual depth-sensing vision-based technologies in more detail.

ABC-ENHANCING VISUAL SURVEILLANCE TECHNOLOGIES

The process of eGate operation when using the proposed vision-based technologies is depicted in Figure 1. An outward facing stereo camera setup is observing the queue in front of the eGate and a visual analysis estimates the length of the queue. The queue length estimate can be used to provide an estimated waiting time in front of the given eGate to both the border control operators and the travelers. When a traveler approaches the eGate, he places his ePassport on the passport reader, which then authenticates the ePassport, including electronic and optical security checks. If reading has succeeded, the first eGate’s door opens automatically, and the passenger goes through the eGate. During his walk, recorded live images of his face are captured, and compared against the picture stored in the chip of his ePassport. In addition, a security check is performed against the Schengen Information System (SIS). At the same time, a surveillance system (top view sensor) ensures that only a single person is present inside the eGate (person separation). Once the identity of the (single) passenger has been authenticated, the second door opens automatically, and the passenger steps out of the eGate. The opening of the second doors activates the left luggage detection module. In case that any item was left behind in the eGate, the second door opens again, enabling the passenger to return and pickup his luggage.

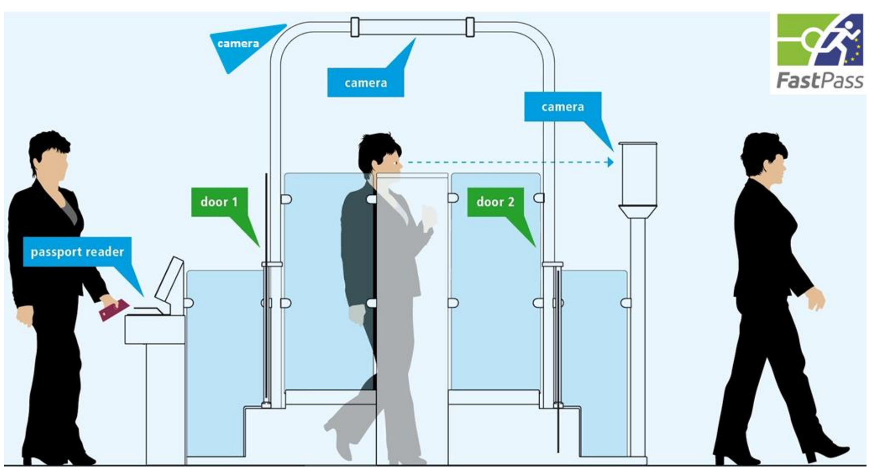

Figure 1. Illustration for an integrated two-step border control process also depicting the employed camera setups.

Person counting and separation: In order to find human candidates within the eGate’s volume, an area of interest is analyzed with respect to local maxima in the depth data which is computed from the images of the top-view stereo camera setup. The stereo camera setup is calibrated offline. The algorithm is designed to separately detect persons walking close to each other (piggy-backing) and to robustly discriminate between varying person-luggage (e.g. person carrying a backpack) and human-human configurations. Our algorithm is capable to detect and separate persons with varying body proportions, clothing and hairstyles. We have evaluated our detector on a test data set of 5184 positive (piggy-backing attempt) and 7344 negative (no abnormality) samples. The rate of successfully detecting the critical events (true positive detection rate) of 0.93 and a correct recognition rate of normality (true negative detection rate) of 0.99 were achieved on the test data set. More details on the detection system can be found in [2] and [4]. The proposed depth sensing person detection and separation framework runs at a frame rate of 15 *fps*. This observation speed implies a great number of measurements during the traveler's presence within the eGate and it enables the high accuracy of proposed critical event detection.

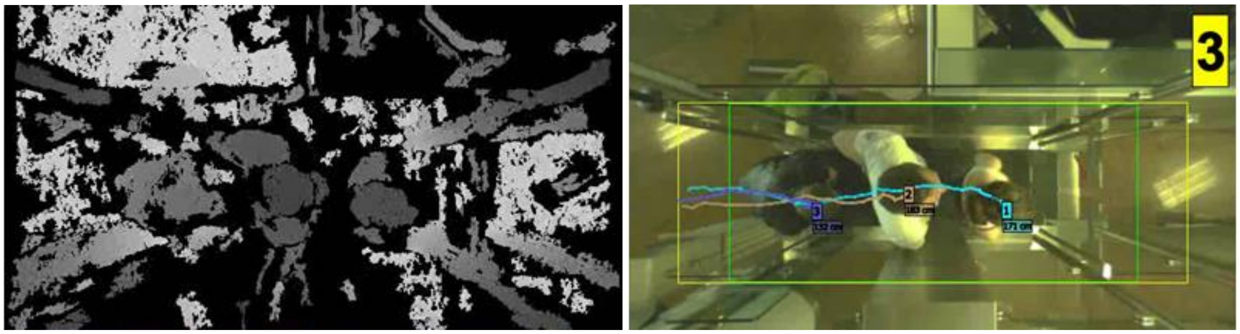

**Figure 2. Left: An example depth image depicting three persons. Bright areas are far from the camera (ground floor), dark areas are close. Entirely black regions do not contain any depth information.
Right: Corresponding detection, tracking and counting results and the estimated number of persons within the eGate.**

Left luggage detection: Our left luggage detection module is based on the fusion of color- and depth-based change detection and static object delineation, both employing a reliable and efficiently computable background subtraction method [3], [4]. An image region is considered as static if it persistently reappears as a foreground region (deviation from a learned background) over a longer period of time. Each (depth and color) static object detection procedure produces an object segmentation estimate, independently from each other. Due to this independence the delectability in complex situations is enhanced: flat objects (such as a dropped passport) generate a change in the color image, but not in the depth image; while poorly contrasted objects (e.g. a trolley with the same color as the eGate floor) are difficult to detect in the color image, but they produce a marked depth deviation. The individual static object estimates (segmented regions) obtained for color and depth cues are fused via a union operation resulting in a reliable detection performance for various types of left objects (see Figure 3). The left luggage detection module is synchronized with the eGate's door signals. Thus, the analysis is stopped during the period where the traveler is inside the eGate, and detection of left items is activated for a short period of time after the person has left the eGate.

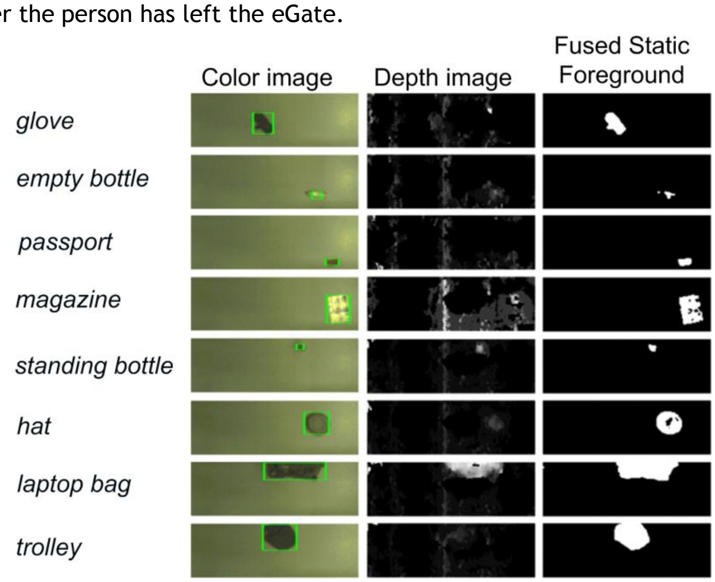

Figure 3. Examples for left luggage detection. Color image with detected bounding box (left), depth image (middle) and resulting static object detection results after fusing color and depth information (right).

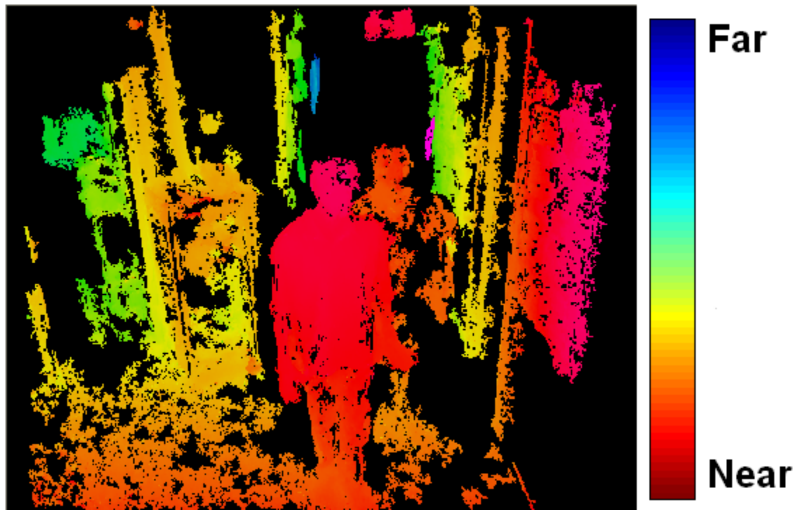

Figure 4. A sample depth image showing two persons waiting in front of an eGate. Color coding represents the distance from the camera, as indicated by the color scale (right).

Queue length estimation: The depth sensing capability of the outward facing stereo camera setup (see in Figure 1) can be well applied to estimate and monitor the number and dynamics of travelers waiting in front of an eGate up to a distance of approximately 12 meters. Depth images provide valuable visual hints where individual travelers are located (see Figure 4), since the measured depth ordering reveals how the individual, partially occluding persons are spatially arranged within the queue. Combining depth information with color image data furthermore enables a reliable long-term tracking of various parts of the queue, thus characterizing queue dynamics and providing information with respect to an estimated waiting time. Our queue length estimation framework is currently in development; nevertheless, initial results show that difficult indoor and outdoor situations can be successfully analyzed; scenarios where conventional single-camera surveillance solutions typically fail.

CONCLUSIONS

In this paper we presented a set of relevant vision-based technologies for supporting border guards in achieving a higher level of security by preventing unauthorized border crossings, and at the same time improving user experience at the border infrastructure. The presented concepts incorporate reliable hardware and algorithmic components in form of depth sensing vision sensors and illumination-invariant representations, which at the same time encode the appearance of humans and their environment in a highly specific manner.

Based on our extensive current testing results in applied settings and future plans for trials, jointly performed with end users and infrastructure providers we are confident that the proposed technology elements will achieve a significant impact on rendering border control procedures safer and an individual’s travel experience more enjoyable.

ACKNOWLEDGEMENT

The work has been supported by the FastPass project. The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 312583.

REFERENCES

1. M. Humenberger, C. Zinner, M. Weber, W. Kubinger and M. Vincze, “A fast stereo matching algorithm suitable for embedded real-time systems”, CVIU, Vol. 114, Issue 11, November 2010, 1180-1202.
2. M. Rauter, “Reliable Human Detection and Tracking in Top-View Depth Images”, In 3rd International Workshop on Human Activity Understanding from 3D Data, CVPR, 2013.
3. D. Schreiber and M. Rauter, “GPU-based non-parametric background subtraction for a practical surveillance system”. In ECV Workshop, CVPR, pages 870–877, 2009.
4. D. Schreiber, A. Kriechbaum and M. Rauter, “A Multisensor Surveillance System for Automated Border Control (eGate)”, 1st Workshop on Activity monitoring by multiple distributed sensing (AMMDS) at the 10th IEEE Int. Conf. on Advanced Video and Signal-based Surveillance (AVSS 2013)