



MOBILIZE – Maintaining the operational safety and security of large railway systems in emergency situations

David Schreiber · Daniel Bauer · Michael Hubner · Martin Litzenberger · Andreas Opitz · Stephan Veigl · Bettina Biron

Received: 1 June 2023 / Accepted: 17 July 2023 / Published online: 1 September 2023
 © The Author(s) 2023

Abstract In this paper we describe the concept and ongoing work of the research project MOBILIZE, which addresses the operational safety and security of large railway systems to prevent sabotage and vandalism. Examples of such acts are manipulation of system components, intentional placement of objects on the tracks, theft of copper cables or damage to property such as graffiti on parked trains. The prevention of personal injuries resulting from crossing tracks or climbing on parked wagons and thereby getting too close to or even touching the overhead lines is also an important aspect. A permanent installation of video surveillance systems for the entire railway infrastructure is not feasible and, what is more, state-of-the-art video surveillance alone is currently not up to the challenges of monitoring very large areas completely. Therefore, MOBILIZE focuses on the development of a reliable portable system with multi-sensor modalities. In case of increased incidents in a specific region, the system can be deployed quickly and easily. The development of such a system raises questions that represent the main scientific challenges to be explored within MOBILIZE: which combination of sensor technologies is the most suitable to reduce false alarm rates to a minimum in practical operation, legal issues such as the changing regulations regarding

the usage of drones, usability for the operator, integration into the operational procedures of the railway operators as well as future economic exploitation of the MOBILIZE project. The current paper focuses on the work done on ground-based visual sensors as well as their fusion with other sensors employed within MOBILIZE, and an assessment of their social impact.

Keywords Operational safety · Railway systems · Provisional and relocatable surveillance system · Multi-modal sensor fusion

MOBILIZE – Aufrechterhaltung der Betriebssicherheit und Absicherung großer Eisenbahnsysteme

Zusammenfassung In diesem Artikel werden das Konzept und die laufenden Arbeiten des Forschungsprojekts MOBILIZE beschrieben, welches sich mit der Betriebssicherheit und der Absicherung großer Eisenbahnsysteme gegen Sabotage und Vandalismus befasst. Beispiele für solche Handlungen sind die Manipulation von Systemkomponenten, das absichtliche Platzieren von Gegenständen auf den Gleisen, der Diebstahl von Kupferkabeln oder Sachbeschädigungen wie Graffiti auf abgestellten Zügen. Ein weiterer wichtiger Aspekt ist die Verhinderung von Personenschaden, zum Beispiel beim unbefugten Besteigen von Gleisanlagen oder Klettern auf abgestellte Waggons wie auch beim Annähern an bzw. bei Kontakt mit den Oberleitungen. Eine dauerhafte Installation von Videoüberwachungssystemen für die gesamte Eisenbahninfrastruktur ist faktisch nicht durchführbar. Darüber hinaus sind moderne Videoüberwachungstechnologien allein derzeit nicht in der Lage, sehr große Bereiche vollständig zu überwachen. Daher konzentriert sich MOBILIZE auf die Entwicklung eines zuverlässigen, portablen Systems und der

M. Litzenberger is an OVE member.

D. Schreiber (✉) · D. Bauer · M. Hubner · M. Litzenberger · A. Opitz · S. Veigl
 Austrian Institute of Technology (AIT) GmbH,
 Giefinggasse 4, 1210 Vienna, Austria
david.schreiber@ait.ac.at

B. Biron
 University for Continuing Education Krems, Krems, Austria

Kombination unterschiedlicher Sensor-Modalitäten. Im Falle eines erhöhten Aufkommens von Vorfällen in einer bestimmten Region kann das System schnell und einfach eingesetzt werden. Die Entwicklung eines solchen Systems wirft Fragestellungen auf, welche die wichtigsten wissenschaftlichen Herausforderungen darstellen, die im Rahmen von MOBILIZE zu erforschen sind: die geeignetste Kombination von Sensortechnologien zur Reduktion der Fehlalarmrate im praktischen Betrieb, rechtliche Fragen wie die sich ändernden Vorschriften für den Einsatz von Drohnen, die Benutzerfreundlichkeit, die Integration in die Betriebsabläufe der Bahnbetreiber sowie die künftige wirtschaftliche Verwertung des MOBILIZE-Projekts. Dieser Artikel konzentriert sich auf die Arbeiten an den bodengebundenen visuellen Sensoren sowie deren Fusion mit anderen im Rahmen von MOBILIZE eingesetzten Sensoren und ihre gesellschaftlichen Auswirkungen.

Schlüsselwörter Betriebssicherheit · Eisenbahnsysteme · Provisorisches und ortsveränderliches Überwachungssystem · Multimodale Sensorfusion

1 Introduction and motivation

Maintaining the operational safety and security of large railway systems such as the infrastructure of the Österreichische Bundesbahnen (ÖBB) or Wiener Linien represents a significant challenge for the operators. Acts of sabotage, vandalism or even pranks or tests of courage by young people can significantly endanger operational safety. Such acts include the manipulation of system components, the intentional placement of objects on the tracks (such as stones, branches, shopping trolleys, or e-scooters) or the theft of copper cables. The consequences can be severe damage to vehicles, personal injury from collisions or emergency braking and the associated cessation of train services. In addition to the property damage that occurs as a result, such incidents can also put the parts of the public transport network out of operation in an uncontrolled manner, which in turn causes enormous damage to society. Moreover, personal injuries, often lethal, occur due to persons climbing on parked wagons thereby getting close to overhead power lines. Burglaries and damage to property are part of day-to-day business and thus also endanger operational safety and security.

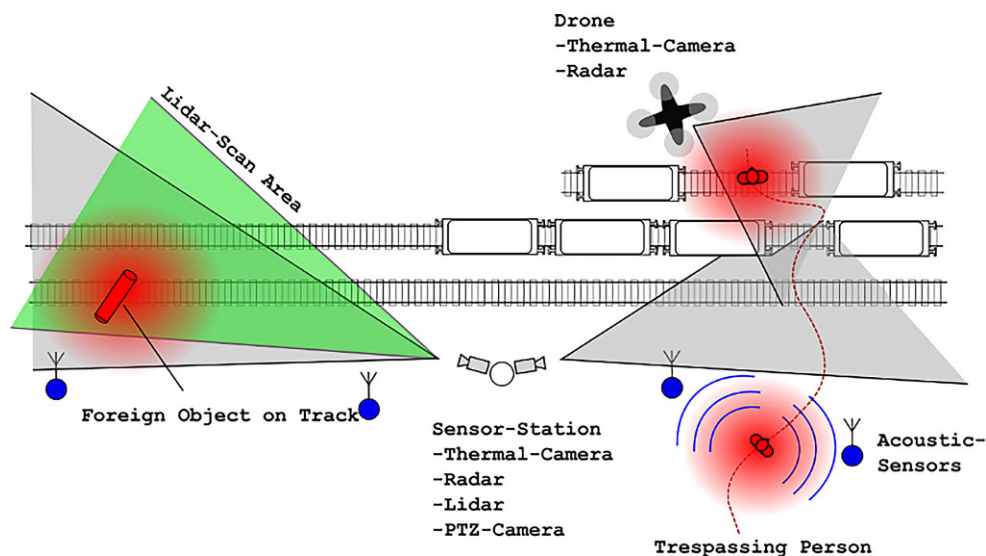
These examples demonstrate that it would be even more difficult to protect the operational security of large railway infrastructures against sabotage in times of crises. The necessary concepts for the operational safety and security of large railway infrastructures in times of crises, as well as the technical equipment for their implementation, hardly exist so far. A permanent installation of video surveillance systems for the entire railway infrastructure, including workshops,

depots, parking areas and open routes is technically hardly solvable and economically not feasible due to the geographical extent. In addition, state-of-the-art video surveillance by itself is not up to the challenges of reliably monitoring very large areas. Incidents and crises, be it the increased occurrence of vandalism on security-relevant systems or due to targeted warnings from security authorities, are presumed to increase in the future. A reliable system with advanced sensors that—in the event of an incident, such as increased observations by train drivers along the route, or information from the security authorities—can be set up quickly and easily and used for increased surveillance of the railway infrastructure, does not exist. Such a system could make a significant contribution to ensuring operational safety in the railway system.

In MOBILIZE [1], a two-year project which started in October 2021, the operational and technical concepts for a portable installation, of a technical security system for large-scale facilities are to be examined and demonstrated in operational environments. These ambitions give rise to questions that represent the main scientific challenges of the project. The usability, the integration into the operational procedures of the railway system and the data compliance framework must be examined to be able to implement such a concept in practice in the future. Technical issues need to be clarified as well, e.g., how reliably people can be detected using thermal imaging cameras over large distances. Which combination and configuration of sensor technologies is best suited to reduce false alarm rates to a minimum for practical operation? Which methods and procedures are necessary to enable easy and fast deployment at new location? In contrast to existing systems, which automatically activate cameras using simple motion sensors or camera-based motion detection, MOBILIZE will use the classification of people and events through the fusion of multiple sensors, namely, thermal cameras, radar, LiDAR scanners, acoustic ground sensors and drones, as can be seen in Fig. 1. The MOBILIZE concept is to be evaluated under a wide variety of environmental conditions over the year and during day and night. This evaluation is the decisive basis for describing the limits of the concept in a white paper. MOBILIZE will fuse the data from multiple sensors and sensor types to be able to secure a length of 200 m in each direction (effective 400 m) aiming to significantly reduce state-of-the-art false alarms rate. MOBILIZE will geolocalize the detected events within ± 5 meters on a situation map, while current systems are only capable to symbolize the position and the angle of view of the camera or sensor in situation maps.

The technical concept of MOBILIZE is based on 5 core technologies: person detection in thermal video images, LiDAR, radar and audio detection, and data fusion. In this paper, we focus on the work done within AIT concerning person detection with thermal cameras, and the concept of the multi-sensor fu-

Fig. 1 Overview of the MOBILIZE concept



sion. The rest of this paper is organized as follows. In Sect. 2, the state-of-the-art of surveillance solutions for railway systems is outlined. Section 3 presents the various technical aspects of the MOBILIZE systems, including hardware, video-based detections with thermal cameras, digital stabilization of thermal videos, sensor fusion, and, finally, data acquisition and test scenarios. Section 4 describes social science- and legal-related issues. Finally, Sect. 5 contains our conclusions.

2 Previous work

The protection of the route network and the parking spaces of public rail transport is an unsolved topic for the operators of such systems, due to an increasing risk of intentional sabotage, pranks or tests of courage by young people, or natural disasters. These acts regularly lead to life-threatening situations and interruptions in the railway system. There are already some available products on the market that enable monitoring large areas. These products range from surveillance systems with RGB or thermal video cameras (AVASYS [2] and FLIR-Monitoring System [3]) to the use of multi-sensor networks. However, most of these systems are permanently installed and therefore are easy to detect and circumvent. They are unsuitable for open routes since complete monitoring of railway routes is neither possible nor desirable. Mobile systems, such as the G4S FP360 [4], offer a solution that can be installed relatively quickly. Other manufacturers such as Securiton [5] or EXENSOR [6] offer compact, portable multi-sensor solutions. In these systems, typically the various sensor modalities operate independently from each other, and at best offer a common representation in a situation map. Consequently, detection and false alarm rates are mostly not satisfactory. In contrast, MOBILIZE will further geolocalize the detected events exactly (± 5 m) on a situa-

tion map, while current systems can only symbolize the position and the angle of view of the camera or sensor in a situation map and leave the localization of the event in a camera image to the skills of the video operator. Moreover, frequent false positives result in loss of trust among end-users, which reduces efficiency and might cause about the underlying security concept.

Current systems require high, complex mast constructions, are mounted on car trailers, and therefore cannot be used for areas that are difficult to access. The concept of the sensor station in MOBILIZE will be based on small modules that can be quickly assembled on a lightweight portable mast and put into operation within 1 h without expert knowledge. Additionally, a drone equipped with sensors is selectively activated by the other sensors. Existing ground sensors, such as those from EXENSOR, work independently of a power line for several days, but have a very limited range of around a few meters. The MOBILIZE system will be equipped with an internal battery to provide energy self-sufficiency for several hours. The technical concept of MOBILIZE provides that sensors with low energy consumption serve as a trigger for sensors with higher energy consumption for detailed analysis. This concept increases the duration of energy self-sufficiency. In addition, the energy supply is planned in such a way that commercially available photovoltaics can also be used as an energy source.

New concepts, that enable an automatic comparison of the sensors with each other (fusion), are required. Such an approach would be able to significantly improve the quality of the alarms in the future. Almost every manufacturer of video management systems has the option of integrating maps into the system for overview purposes. The high effort for the configuration of the system (positioning/orientation of the cameras) is the biggest limitation of the currently used video and security management systems.

The use of drones is currently the most modern and effective way of monitoring large areas. The previous legal situation was replaced by a new EU drone regulation on January 1st, 2021 [7]. This new legal framework is therefore relevant for MOBILIZE. Investigating a new technology in connection with a completely new legal framework increases the scientific knowledge gain of the project, since no research results exist in this context yet.

Permanently installed video surveillance systems are currently used almost exclusively for monitoring trains, stations, and infrastructure facilities. Large and immovable towers are sometimes used for individual priority actions. By means of video cameras mounted on the tower, events can be detected during the day, but only with great restrictions at night. The video towers are connected to a control center. This type of video surveillance is not suitable for surveillance on the open route. Security personnel or, on a case-by-case basis, trail cameras are used specifically to detect stones, branches, or other objects on the track, or people trespassing on the tracks area. This requires additional resources to ensure operational security. A system that solves these challenges does not exist.

Additionally, current state-of-the-art systems on the market lack the ability to be integrated into existing railway systems and process landscape of the operational bodies; no sensors equipped with software-based detection are currently used. Upon detection (via visual contact or in camera images) of intruders, a security center is informed, which in turn notifies the police. Additionally, radio patrols from the railway company are ordered to the scene of the incident since the police are only allowed to enter the railway area after the railway company has given its approval. The current workflows and processes lead to loss of time and make timely intervention to maintain operational safety more difficult. Currently, it is virtually impossible to implement a complete and effective protection quickly and easily, which can be integrated into the railway provider's operational processes in case the security authorities issue a warning about an increased risk.

3 The MOBILIZE concept: a portable surveillance system

3.1 Hardware

The MOBILIZE system built at AIT contains two transportable sensor stations. Each sensor station comprises a sensor head mounted on a lightweight telescopic pole, which can be extended up to 5.5 meters using a manual hand crank, and a protective case containing the rest of the components (see Fig. 2). Each sensor head is connected via ethernet and power cables for power delivery. The sensor head contains a Power over Ethernet (PoE) switch as well as power distribution for all sensors and IP cameras.

It is weather resistant up to IP65 (excluding the laser scanner) and includes the following sensors:

- a differential GNSS system for positioning, providing real time IMU axes and IMU/GNSS heading measurements
- two thermal cameras having different focal lengths (wide-angle and tele) to cover both short and long distances each with a resolution of 640 × 480 pixels
- an HD RGB pan tilt zoom (PTZ) camera for the operator to visually verify the alerts from the detectors and to manually surveil the area
- a microwave radar antenna (provided by Joby Austria, <https://inras.at>) for reliably detecting moving objects
- an optional 3D lidar scanner (RIEGL VZ-400i [8]), for detection of objects on the tracks

The system is connected via a VPN over Wi-Fi or LTE 4G with all other partners' components in the project. The system can be moved fast and easily by a security team of 2–3 persons. Three audio detectors are deployed as satellite sensors along the perimeter of the depot and within the field of view of the other sensors and connected via wireless data transmission to the overall system. It is to be noted, that the overlap of the fields of view of the individual sensors is a prerequisite for the subsequent data fusion of the sensor modalities.

The other components of the sensor station are integrated in a large transportable protective case, weather resistant up to IP54. The system inside the protective case is cooled with a temperature-controlled fan and can be covered with a tarp for shielding it from direct sunlight and heat. The case contains the following components:

- two desktop i7 mini-PCs, including RTX3060 GPUs for processing the sensors data
- LTE/4G router/modem for establishing a secure VPN connection to be able to
 - notify the operator in case of an alarm
 - provide live video footage from the static thermal cameras and the PTZ camera
 - remotely control the whole system
 - provide an NTP time server for the local sensor network
- ethernet switch for connecting and powering the sensors and cameras
- 24-volt 100A hour rechargeable LiFePO4 battery
- mains AC/DC charger and 24-volt power supply
- DC/DC power converters for powering all the 48-volt PoE components in the sensor head
- a 24VDC to 230VAC converter to power the mini-PCs when on battery

3.2 Video object detection with thermal cameras

Object detection, and in particular person detection, is a fundamental task in computer vision and image and video processing. An object detection process



Fig. 2 a Illustration of a sensor station with its sensor head. b A closer look inside the protective case

takes an image as an input, localizes the objects that appear in the image (in terms of bounding boxes), assigns class labels (e.g., a person or a car) to the objects, and gives the class probability for each detected object. In the last decade object detectors based on deep neural networks (DNNs) have become the state-of-the-art in object detection. Artificial neural networks are computing systems inspired by biological neural networks. Such systems learn progressively to perform tasks by considering examples. A deep neural network is an artificial neural network with multiple layers between the input and output layers. There are different types of neural networks, but they always consist of the same components: neurons, synapses, weights, biases, and functions. One of the advantages of DNNs is their ability to extract internal representations automatically, outperforming hand-crafted features that are needed for traditional detectors. Nevertheless, object detection is a challenging task due to many factors such as clutter, imaging conditions, large number of object classes and instances, and no-

tably partial occlusion. Consequently, state-of-the-art deep learning detectors are far less robust than human beings [9].

In particular, the YOLO DNN-based detector [10] is commonly used in practical applications and systems due to its real time performance and ability to detect tens of different object classes with satisfactory precision. Since the publication of the original YOLO detector, several variants and improvements have been proposed, and one of the latest versions was chosen for MOBILIZE. However, deep learning-based detectors were developed and trained using visible light (RGB) sensors. As RGB sensors are not sufficient for outdoor surveillance under low light and during nighttime, thermal sensors were selected instead for MOBILIZE. Thermal images are more difficult to process than RGB ones, because they are of lower quality and poor contrast, and contain a smaller number of features and are noisy. Finding out how reliably can persons be detected by thermal imaging cameras over large distances is one of the scientific challenges

of MOBILIZE. Based on preliminary work and with the help of the data acquired within MOBILIZE, a person detection algorithm was developed for the MOBILIZE system. It is based on an improved implementation of the YOLO detector [10], adapted to cope with images provided by thermal cameras. To cover short and long distances with appropriate image resolution, two thermal cameras, having different focal lengths (wide-angle and tele) and overlapping field of view are employed.

In MOBILIZE, the Robot Operating System (ROS 2) [11] is employed as a middleware for combining and enabling the communication between all relevant components of the final system, as the matured and widely used ROS 2 framework provides a stable platform for the development of modular software with an increased use of sensor hardware in combination with dedicated algorithms. Accordingly, the YOLO detector was implemented as a ROS 2 node. Furthermore, MOBILIZE makes use of a GNSS receiver, which provides the exact position of the camera and its orientation. This enables an automatic mapping of the detections from the camera coordinate system to global world coordinates (WGS84) [12]. This georeferencing was implemented as separate ROS 2 node.

Figure 3 shows example detections of persons and trains on two test sequences acquired during the recording session during autumn 2022. Images captured by the wide-angle (overview) camera are in the left column, whereas images from the tele camera in the right column. In the first row, persons are walking along the track and are detected in the wide-angle image but are not (yet) visible in the field of view of

the tele camera. The second row depicts the same scene a little later when the persons are further away from the sensor station and are crossing the tracks at a distance of about 170 m. In the wide-angle image, the people are already too small (in pixel resolution) to be reliably recognized by the algorithm but are detected instead by the tele image.

3.3 Digital stabilization of thermal videos

Video stabilization is an important task of video-based outdoor surveillance, such as in MOBILIZE, where the thermal cameras are mounted on a pole and are susceptible to harsh weather conditions and high winds, resulting in camera shake which leads to unwanted motion between successive video frames. On the one hand, this unwanted motion and jitter annoys the security personnel which observes the scenes on monitors, and on the other hand, it deteriorates AI-based algorithms, such as object detection and tracking, which process these video frames. There are three basic techniques available for video stabilization: optical, electronic, and digital image stabilization. The digital stabilization method is purely software-based and independent of the camera's physical dimensions. The advantages of using this type of technique are that there are no moving and costly components (e.g., gyro sensor, motor assembly, servo electronics, gimbal), and the ability to apply different algorithms to improve the stabilization.

Two-dimensional stabilization methods are widely implemented in commercial packages due to their robustness and low cost. Two-dimensional methods are

Fig. 3 Example detections of persons and trains. **a, c** wide-angle camera images; **b, d** tele camera images. Persons walking along the track and are detected first by the wide-angle camera (**a, b**), and later by the tele camera (**c, d**)

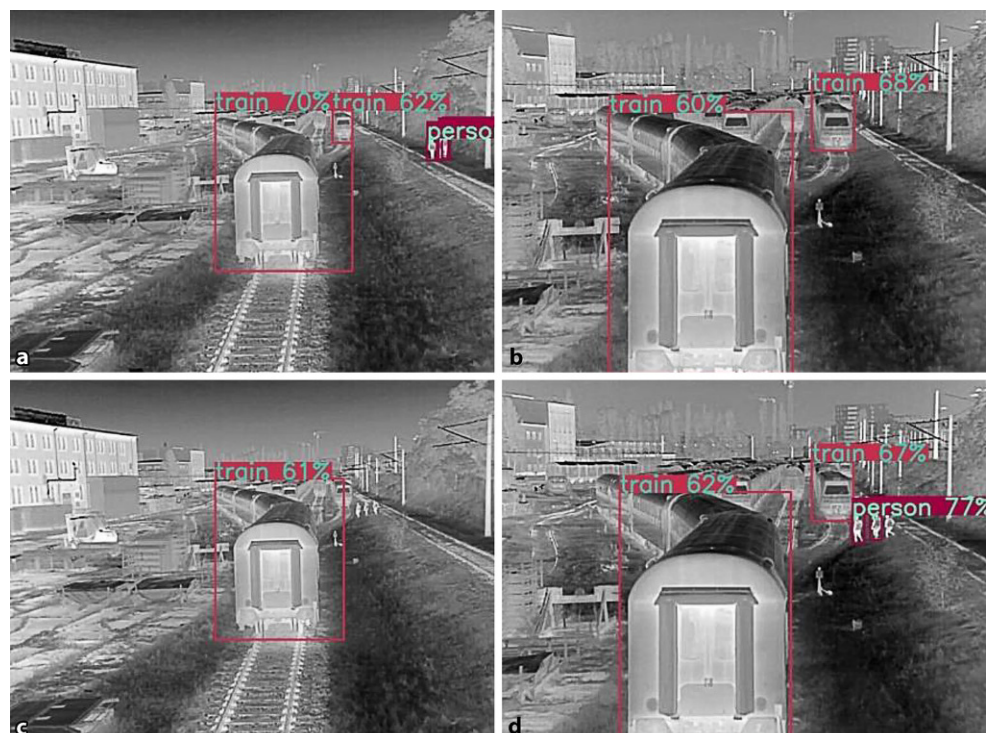
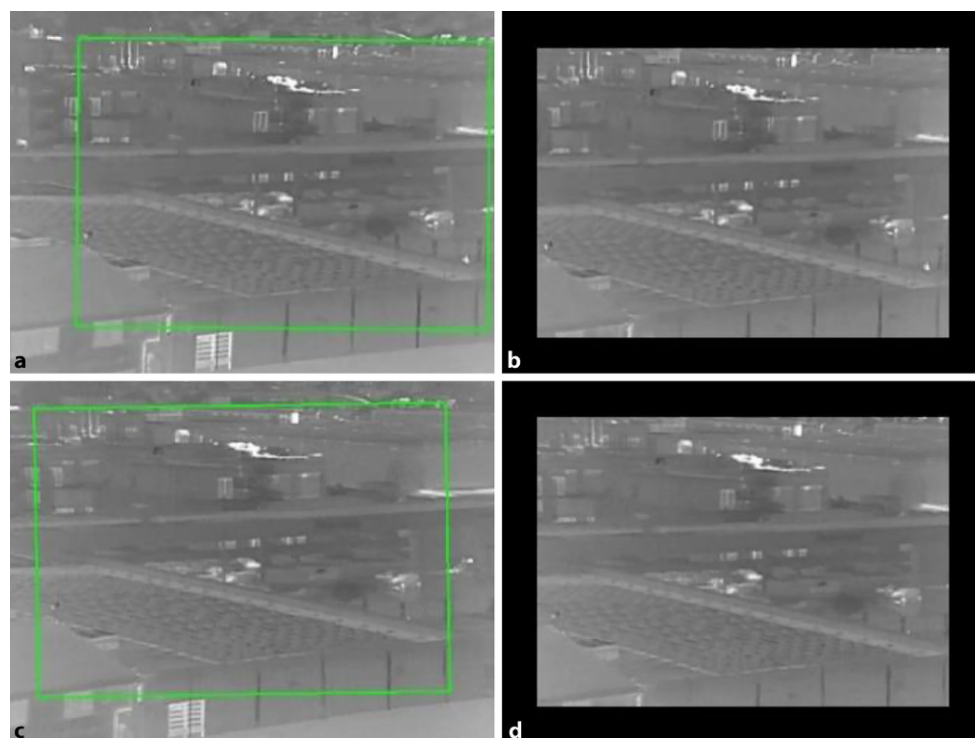


Fig. 4 Mechanical shaking of the pole: Exemplary images of un-stabilized frames (**a, c**) and stabilized frames (**b, d**)



based on the 2D apparent motion between 2 video frames (current and reference frames). In particular, global motion estimation is sufficient to compensate for camera motion in scenarios where the entire image is equally affected, i.e., when there is little depth variation or when the 3D compensation is only rotational. Most of the state-of-the-art approaches of 2D video stabilization methods were developed for RGB sensors [13, 14]. In contrast, thermal images are difficult to process as they are of lower resolution, lower quality, and contrast, contain a smaller number of features and are noisy. State-of-the-art methods developed for RGB videos are bound to perform poorly on thermal images due to these issues. In fact, only a small number of previous approaches focused on the video stabilization problem for thermal sensors. These methods were developed mainly for UAV-based applications such as monitoring of wildfires, and for border military surveillance [15].

Stabilizing thermal videos for real time applications is still an open research area. This is due to the difficulty in detecting and tracking good features in infrared videos having low resolution and poor quality. For MOBILIZE, a fast real-time online approach is needed which does not require an offline training for each individual scenario (as is the case with machine learning-based approaches), is easy to install and to set up, and which could run with modest computational resources by the end user. Consequently, a global two-dimensional motion estimation stabilization approach was employed and tested within MOBILIZE. The approach is based on *direct methods* [16, 17], which entails the estimation of mo-

tion parameters between 2 frames directly from intensities, in contrast to *feature-based methods*, which first extract a sparse set of distinct key-point features from each image separately, and then recovers and analyses their correspondences to determine the motion. The proposed algorithm is simple, yet capable of real-time speeds of about 50 FPS as well as robust, overcoming the limitations of thermal images and providing accurate frame alignment even in the case of severe camera shake, as demonstrated by experiments with a randomly added motion to the frames as well as with application of mechanical shakes to the pole. Figure 4 shows some exemplary images of the un-stabilized (left) vs. stabilized frames (right). The stabilized frames are cropped due to the existence of blank borders.

3.4 Multi-sensor fusion

In MOBILIZE, statistical fusion concepts based on Bayesian inference are being developed, analyzed, and improved with a view to reducing false alarms rate. Particular attention is paid to the combination of the individual sensors in relation to the coverage of the application scenarios. The concept of the data fusion approach is based on the aggregation of multiple density maps. A density map represents the monitored geographical area for one specific sensor modality used within the system, to allow a spatial-temporal comparison between the different sensor modalities. The approach combines probability density maps with a rule engine that links data from multiple density maps.

A combined video detection and audio classification of critical events in three-dimensional space has been described in [18]. The fusion of different wireless positioning technologies in two-dimensional space was reported in [19]. Often the temporal and spatial relationships of events in existing systems are only roughly analyzed, such as the assignment of a person detection from a camera image to acoustic cues detected in the vicinity of the same camera [20, 21]. Rule-based approaches with sensor installations to detect relevant security-related events have been described in [22]. Map-based methods can be used to determine the location of an object or security-related event. There are several approaches to achieve the necessary data association between different detections, such as feature-based and position-based maps [23]. The most common application of map-based sensor fusion today is found in the automotive sector, where a relatively small number of sensors are positioned in a relatively small spatial area around the vehicle [24]. Monitoring tasks with a very high number of sensors (e.g., up to 100) and for a very large geographical area (e.g., up to 1 km), based on evaluating the nearest neighbor relationships of elements in the map, can become problematic. When evaluating nearest neighbor relationships, the computational effort increases dramatically with the number of sensors and the generated target hypothesis. For position-based maps, where each cell of the map represents the probability of an event or object at that location, the nearest adjacent relationship can be easily evaluated by addressing the adjacent cell of the grid. This method therefore scales much better for large sensor networks. Occupancy grid maps [25] and density maps [26] are common solutions for such location-based maps.

Figure 5 illustrates the fusion concept with the example involving two sensor modalities: an acoustic-based and a video-based detectors feeding into the fusion system. The number of density maps is equal the number of the sensor modalities. Each density map is divided into equally sized grid cells. Each cell repre-

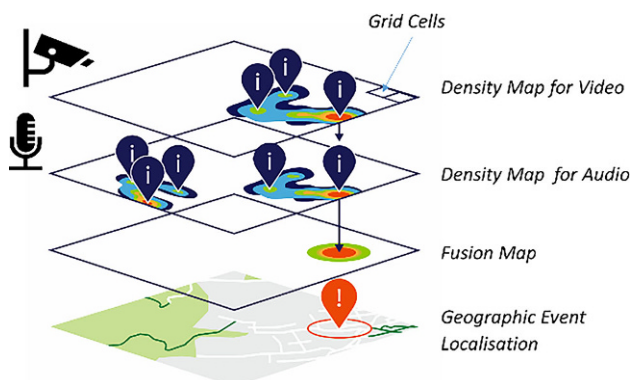


Fig. 5 Schematic of the fusion approach: an example involving a video and an audio detector

sents a coarse geographic location (of about 1×1 meter in size) in the monitored area and holds a value equal to the probability of a detected event (indicated by blue “i-pins” in Fig. 5) within this coarse location. Each detection is attributed to a geographical area by projecting the detection area provided by the detector into the corresponding cells of its density map. For the sake of the illustration, the detection regions shown in Fig. 5 are visualized as arbitrarily shaped probability distributions (the actual shape of such distributions may differ). For the cells covered by the detection area, the probability is updated using the detection's confidence value delivered by the detector.

The set of density maps is periodically aggregated into one fusion map, by combining weighted cell values corresponding to the same spatial position. Currently, the weights are estimated empirically, based on our experience with sensors' accuracy and reliability. It is planned, eventually, that the weights will be estimated via a machine learning approach. After aggregation, an alarm threshold is applied to all cells of the fusion map, and cells exceeding the threshold are generating an alarm (indicated by a red “!-pin”). The alarms' geographic location, that can be presented to a human operator on a map, is determined by the cells' location exceeding this threshold.

3.5 Data acquisition and test scenarios

During autumn 2022, a data recording session, together with an initial test, were performed within MOBILIZE, making use of a single sensor station. The test recordings consist of security relevant scenarios, employing 3 different sensor modalities: 2 thermal cameras, a microwave RADAR and 3 audio detectors (see Fig. 6). The recording session took place in a large open railway depot, which has frequently been subject to vandalism acts. The 11 recorded test scenarios included illegal trespassing and graffiti spraying and were performed by actors to obtain relevant sensor data for each scenario. The 3 audio detectors were deployed as satellite sensors along the perimeter of

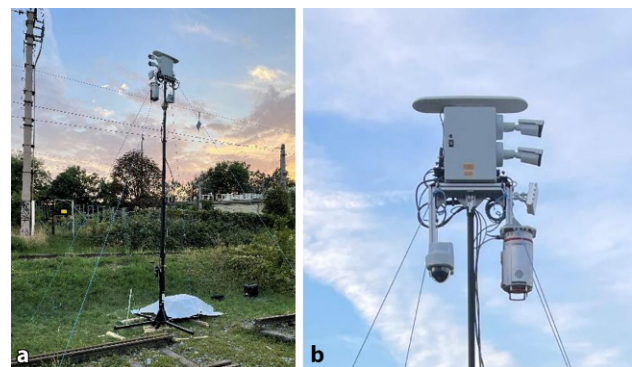


Fig. 6 The sensor station deployed on the test site during data acquisition during autumn 2022. **a** the entire sensor station; **b** the sensor head

the depot and within the field of view of the other sensors, as the overlap of the fields of view of all the individual sensors is a prerequisite for the subsequent data fusion of the sensor modalities.

In preparation for the data recordings, a ROS 2 based system for simultaneous acquisition of video frame, position and orientation data was implemented. This was necessary to generate a heterogeneous dataset from various sources that can be played back in a synchronous way in the same chronological order. The geographic locations and orientations for all included sensors were recorded with high precision via GNSS devices, to allow the full georeferencing of their detections, in the subsequent fusion step. During the acting of the different scenarios, the sensors' data streams were recorded for offline processing and classification by automatic video or audio detection and classification software. RADAR targets were recorded as is, without further classification. The output of the detectors was georeferenced using the known sensors' geolocations and orientations.

The upcoming work in the MOBILIZE is to fuse the sensors' detections, and to evaluate the improvement in the detection and false alarm rates, induced by the fusion, compared to the performance of the individual sensors. Next, the fusion software will be implemented and tested in a final test operation planned for summer 2023, where various scenarios (persons crossing tracks, objects on tracks, vandalism, etc.) will be acted by volunteers. This testing phase will last for several days and nights and will make use of both sensor stations built by AIT. Based on these test recordings, the performance of the individual sensors, as well as of the whole system, will be evaluated.

4 Social science and legal approach

While MOBILIZE technology partners are concerned with the feasibility and implementation of the technical solutions, the social science or GSK (Geistes-, Sozial- und Kulturwissenschaften) partners are concerned about the human factor, the implementation of a technical system in everyday work, the risks, opportunities, and the effects of these technological developments on users in the Business to Business (B2B) sector and society as a whole. Thus, it is important to weigh the advantages and disadvantages of technological changes in a theoretical manner from a sociological, legal, and ethical perspective. Thus, the main goal of the legal and sociological analysis in MOBILIZE is to explore the implications of unintended side effects when deploying sensor fusion and large-scale surveillance technologies. Considerations of risk governance and mitigation, as well as ensuring privacy compliance, also play an important role [27–29].

From a legal perspective, particular attention has been paid to the data protection impact assessment under Art 35 GDPR (General Data Protection Regulation of the European Union). In particular, the rel-

evant aviation law provisions for the planned use of drones also represent a legal limit to what is technically feasible. As the planned system is portable, general recommendations are given for the deployable sensor station, as it is in the nature of things not to be able to conclusively examine from a legal point of view every conceivable location. Since the video data processed in this project is not intended to reveal sensitive information about individuals, the images to be recorded are classified as personal data, but not as sensitive data per se. If face or voice are identifiable, it is legally considered as information related to identifiable persons. The results of the legal analysis, which is beyond the scope of this paper, are included in the social science approach [30].

4.1 Technological impact assessment

On this basis, the various approaches to technology assessment are concerned with how to increase the chances of technologies addressing societal issues without causing unintended risks and side effects. In practice, however, from an application-oriented perspective, the primary concern is to find out the attitudes of those people who are to interact directly with the new technology. After all, they are the ones who will later decide whether and how to use it and integrate it into their existing activities. It is important that the actors on the ground are considered as experts in their field of activity. It is them who, based on their experience, know best where support from technical systems would be helpful and what it might look like—or where it might not. It is the task of social science to identify these attitudes and experiences and to use this additional knowledge to positively influence technology development.

In a first step, expert interviews were conducted with those responsible for the topic of safety at ÖBB and Wiener Linien. The aim was to obtain information from the employees themselves about their needs and any obstacles for the implementation. Above all, it is essential to consider the impact of new technologies on existing processes and work routines: “Decisions about the introduction of such technologies are often made with limited prior knowledge about the extent to which such systems will change existing work practices and influence work performance (negatively as well as positively)” [31]. It is crucial, they argue, to understand the processes and activities of potential users to involve them in technology development. This is the only way to ensure that the functions required by the users can be covered by the new technologies. [32] speak of a technology being used when it is accepted. This in turn is related to trust in the technology, whether it can reliably fulfill the desired functions and means more efficient work.

Based on this, a questionnaire will clarify how security experts estimate the risk potential for employees and customers, as well as the opportunities. Top-

ics include attitudes toward surveillance drones and cameras in the context of privacy issues. The topics include a question, among other things, regarding a possible fundamental skepticism based on empirical data which indicates that new technologies do not necessarily mean a simplification of everyday work, but in some cases can even lead to more work without any discernible improvement in quality. However, it should be noted at this point that an analysis revealed that the target groups are by and large technology-savvy and technology-friendly groups of people. Initial results show that the employees of the transport companies are very positively inclined to new developments in monitoring, also when it is related to their working reality.

5 Conclusions

This paper outlines the concept and ongoing work of the project MOBILIZE, which addresses the operational safety and security of large railway systems against acts of sabotage, vandalism, or pranks, including manipulation of system components, intentional placement of objects on the tracks or the theft of copper cables, as well as preventing personal injuries resulting from climbing on parked wagons and getting too close to or even touching overhead lines, and damage to property such as graffiti on parked trains. Permanent installations of video surveillance systems for the entire railway infrastructure are technically and economically hardly feasible. Moreover, state-of-the-art video surveillance alone is currently not up to the challenges of completely monitoring very large areas. MOBILIZE focuses on the development of a reliable system with advanced sensors that—in the event of increased observations by train drivers along the route or security authorities—can be set up quickly and easily and used for improved surveillance of the railway infrastructure.

The technical concept of MOBILIZE is based on 5 core technologies: person detection in thermal video footage, LiDAR, radar, and audio detection, as well as data fusion. The technical concepts and application-relevant questions, as well as the integration into the operational procedures of the railway operators, are examined and demonstrated in a real operational environment. In this paper, we focused on the work done within AIT concerning person detection with thermal cameras, and the concept of the multi-sensor fusion. The development of such a system gives rise to many questions which represent the main scientific challenges and the innovations of MOBILIZE, to be explored within MOBILIZE. This includes:

Which legal framework must be considered for putting a temporary system for the protection of large railway systems into operation? In the project, the basic requirements regarding data protection are to be researched to be able to put the entire system into operation.

How can alarm messages and sensors such as cameras providing live information best be integrated into existing railway processes to improve safety- and security-relevant incidents? Additionally, the project should evaluate the effects of new systems on operational management concepts.

How can the sensors be calibrated under the constraint that the MOBILIZE system should be deployed and ready for use within one hour by a team of non-experts?

Which methods and algorithms for real-time person detection in thermal images are suitable for detection of persons over large distances, and how can video stabilization methods contribute by enhancing the video quality?

Which fusion concepts are suitable for creating a spatiotemporal coincidence of detections from multi-sensor modalities to localize events, and to minimize the false alarm rate? Particular attention needs be paid to the combination of the individual sensors in relation to the coverage of the application scenarios.

Funding The project MOBILIZE is funded by the Austrian “Mobility of the Future” program—an initiative of the Federal Ministry for Climate Protection, Environment, Energy, Mobility, Innovation and Technology (BMK)



Funding Open access funding provided by AIT Austrian Institute of Technology GmbH

Conflict of interest D. Schreiber, D. Bauer, M. Hubner, M. Litzberger, A. Opitz, S. Veigl and B. Biron declare that they have no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. MOBILIZE. <https://projekte.ffg.at/projekt/4105746>
2. AVASYS. <http://www.avasys.at/>. Accessed 10.05.2023
3. FLIR-Monitoring System. <https://www.flir.com/browse/industrial/monitoring-systems/>. Accessed 10.05.2023
4. G4S FP360. <https://www.g4s.com/en-cz/nase-sluzby/electronic-monitoring>. Accessed 10.05.2023

5. SECURITON. <https://www.securiton.com/>. Accessed 10.05.2023
6. EXENSOR. <https://www.exensor.com/>. Accessed 10.05.2023
7. STRUCINSPECT. <https://strucinspect.com/knowledge/blog/drone-regulations-2021>. Accessed 10.05.2023
8. RIEGL VZ-400i Data Sheet. <https://www.riegl.com>. Accessed 16.06.2023
9. Kaur J, Singh W (2022) Tools, techniques, datasets and application areas for object detection in an image: a review. *Multimed Tools Appl* 81:38297–38351
10. Redmon J, Divvala J, Girshick R, Farhadi A (2016) You only look once: unified, real-time object detection. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas
11. ROS2. <https://docs.ros.org/en/foxy/index.html>. Accessed 10.05.2023
12. WGS84. https://de.wikipedia.org/wiki/World_Geodetic_System_1984. Accessed 10.05.2023
13. e Souza MR, de Almeida Maia H, Pedrini H (2023) Survey on digital video stabilization: concepts, methods, and challenges. *ACM Comput Surv* 55:3
14. Wang Y, Huang Q, Jiang C, Liu J, Shang M, Miao Z (2023) Video stabilization: a comprehensive survey. *Neurocomputing* 516:205–230
15. Yousaf A, Khurshid K, Khan MJ, Hanif MS (2017) Real time video stabilization methods in IR domain for UAVs—a review. *Fifth International Conference on Aerospace Science & Engineering (ICASE)*
16. Irani M, Anandan P (2000) About direct methods. *Vision Algorithms: Theory and Practice*. IWVA 1999. *Lecture Notes in Computer Science*, vol 1883. Springer, Berlin, Heidelberg
17. Baker S, Gross G, Matthews I, Ishikawa T (2003) *Lucas-Kanad 20 Years On: A Unifying Framework: Part 2*, Tech. Report, CMU-RI-TR-03-01
18. Kooij J, Liem MKJ, Andringa T, Gavrilă D (2016) Multi-modal human aggression detection. *Comput Vis Image Underst* 144:106–120
19. Bohn J, Vogt H (2003) Robust probabilistic positioning based on high-level sensor-fusion and map knowledge. *ETH, Zurich*
20. Andersson M, Ntalampiras S, Ganchev T, Rydell J, Ahlberg J, Fakotakis N (2010) Fusion of acoustic and optical sensor data for automatic fight detection in urban environments. *13th International Conference on Information Fusion*
21. Lefter I, Burghouts GJ, Rothkrantz LJM (2012) Automatic audio-visual fusion for aggression detection using meta-information. *IEEE Ninth International Conference on Advanced Video and Signal-Based Surveillance*
22. Marsala C, Museux N, Coffi J-R (2012) Adaptive complex event processing for harmful situation detection. *Evol Syst* 3:167–177
23. Lundquist C, Hammarstrand L, Gustafsson F (2011) Road intensity based mapping using radar measurements with a probability hypothesis density filter. *Ieee Trans Signal Process* 59(4):1397–1408
24. Bouzouraa ME, Hofmann U (2010) Fusion of occupancy grid mapping and model based object tracking for driver assistance systems using laser and radar sensors. *Intell Veh Symp IEEE* 294–300
25. Elfes A (1989) Using occupancy grids for mobile robot perception and navigation. *Computer* 22:46–57
26. Erdinc O, Willett P, Bar-Shalom Y (2009) The BinOccupancy Filter and Its Connection to the PHD Filters. *IEEE Transactions on Signal Processing*, pp 4232–4246
27. Joly PB, Matt M (2022) Towards a new generation of research impact assessment approaches. *J Technol Transf* 47(3):621–631
28. Van Den Ende J, Mulder K, Knot M, Moors E, Vergragt P (1998) Traditional and modern technology assessment: toward a toolkit. *Technol Forecast Soc Change* 58(1–2):5–21
29. Dusseldorp M (2013) *Technikfolgenabschätzung. Handbuch Technikethik*, pp 394–399
30. Simonis G (2013) *Konzepte und Verfahren der Technikfolgenabschätzung*, 1st edn.
31. Pennathur P, Bisantz A, Fairbanks R, Perry S, Zwemer F, Wears R (2007) Assessing the Impact of Computerization on Work Practice: Information Technology in Emergency Departments. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*
32. Grabowski M, Rowen A, Rancy J (2018) Evaluation of wearable immersive augmented reality technology in safety-critical systems. *Saf Sci* 103:23–32

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

David Schreiber, is a scientist working at the Austrian Institute of Technology since 2001, in the field of computer vision. He has received B.Sc. in Physics, and an M.Sc. and a Ph.D. in theoretical High-Energy Physics. Before joining AIT, he worked within several industrial companies, in the areas of computational geometry, image processing and computer vision.



Daniel Bauer, works as a senior technician at the AIT—Austrian Institute of Technology. In 1996 he graduated from the technical school for electronics, worked in industry and aviation and has been working in electronics development and mechatronics in the field of biometrics, optronic systems for automated quality assurance tasks in industry and surveillance since 2000.



Michael Hubner, is employed as a Research Engineer at the AIT—Austrian Institute of Technology. He graduated with focus on numeric of differential equations in 2015 in Technical Mathematics and received his MSc from TU Wien. In 2016 he joined the New Sensor Technology Team at AIT, where he focuses his research on development in the field of multimodal sensor data fusion methodologies of sensor data. Mr. Hubner is also a qualified software architect (ISAQB),

which helps him in planning and developing large scale sensor systems.



Martin Litzenberger, received a diploma in physics and a Ph.D. degree in electrical engineering from Vienna University of Technology in 1999 and 2003, respectively. His diploma and Ph.D. thesis were on the design and optical characterisation of CMOS ESD-protection devices. After finishing his Ph.D. he joined the newly founded business unit New Sensor Technologies at the AIT Austrian Institute of Technology in Vienna, Austria where he has been developing novel embedded systems for industrial applications combining biologically-inspired analogue VLSI sensors and digital electronics. His current research is focused on algorithms for sensor data analysis and data fusion in safety and security applications and fiber optic distributed acoustic sensing. Martin Litzenberger is a member of IEEE and ÖVE. He is author and co-author of more than 50 scientific publications and holds 4 patents. He has received the 2010 ISCAS Best Live Demo Award and the 2011 Outstanding Reviewer Award of AIP—Journal Review of Scientific Instruments. He was involved in the EU projects ADOSE, eMorph, FOSIBLE, SEAMOVES, MANTIS, DayGuide and national projects ASaP, InnoSense, smartcountplus, Eventdetector, ARIS. He has been a technical coordinator in the EU AAL-JP project CARE and project coordinator of the FP7 project CARBOTRAF and AAL-JP project “FreeWalker”

Andreas Opitz, is a Research Engineer at AIT Austrian Institute of Technology GmbH. He received his MSc in Computer Science at TU Wien, Vienna, Austria in 2009. He is working at AIT Austrian Institute of Technology GmbH since 2010 on developing applications in the context of security, process- and data management in combination with various computer-vision-based technologies.



Stephan Veigl, has an MSc in Computer Science from Vienna University of Technology. He is a project manager and research engineer at AIT focusing on modern concepts and architectures for video and image processing. Mr. Veigl has researched and implemented computer vision algorithms and has developed the software design for a modular, service based video analytics system capable of processing both, offline and live data. His projects expertise covers border control & border surveillance projects, as well as projects in the area of fight against terrorism and protection of critical infrastructure.

Bettina Biron, is a researcher at the Danube University Krems—Center for Infrastructural Security. She previously worked as researcher at the Institute for Communication Science at the University of Vienna. She was also a lecturer at the Faculty of Computer Science at the University of Vienna for communication and media science. Her research focuses on infrastructural safety, crisis and war communication, radicalization, social media and gender studies.



Bettina Biron, is a researcher at the Danube University Krems—Center for Infrastructural Security. She previously worked as researcher at the Institute for Communication Science at the University of Vienna. She was also a lecturer at the Faculty of Computer Science at the University of Vienna for communication and media science. Her research focuses on infrastructural safety, crisis and war communication, radicalization, social media and gender studies.