

Visual Surveillance of an Airport's Apron - An Overview of the AVITRACK Project ¹⁾

Philipp Blauensteiner and Martin Kampel

Vienna University of Technology,
Institute of Computer Aided Automation,
Pattern Recognition and Image Processing Group
Favoritenstr. 9, 183-2, A-1040 Vienna, Austria

{blau,kampel}@prip.tuwien.ac.at, <http://www.prip.tuwien.ac.at/>

Abstract:

This paper gives an overview on the recently granted EU project AVITRACK which stands for “Aircraft surroundings, categorized Vehicles & Individuals Tracking for apron’s Activity model interpretation & Check”, Project Number: AST-CT-2003-502818. The proposed project addresses the specific case of automatically supervising commercial aircraft servicing operations at a designated airport. The efficiency and security of the handling process are significant issues for airport authorities, aircraft operators and citizens. On the apron area, aircrafts are parked and serviced by specialized ground support vehicles (fuel, catering, baggage, small maintenance). Hardware devices and algorithms, yet to be developed, will be based on image segmentation, objects tracking and intelligent knowledge based system needs and capabilities. In order to monitor global activities on the apron modeled servicing tasks will be compared to the reality. Since the project recently started, no actual results can be presented, however, the main project goals and preliminary work on background maintenance are presented.

1 Introduction

In the last decade a rapid increase of camera-based monitoring in an increasingly diverse range of settings took place. Vehicle tracking was one of the first attempts of tracking objects in real-time [1, 8] due to the fact, that the shape of rigid objects is less variable than the shape of persons and so the features are easier to track. Current research in the area of tracking vehicles is done on real-time systems and sensors for visual traffic surveillance [2, 9]. Tracking of persons and groups of persons in real time is another objective in recent research on visual surveillance [7].

Visual surveillance of an airport’s apron, that is the area aircrafts are parked and maintained by specialized ground vehicles, includes tracking of vehicles, e.g. fueling vehicles and baggage cars, as well as tracking of individuals, e.g. workers. The importance of visual surveillance is given due to the fact, that an airport’s apron is a security relevant area. Computer aided visual surveillance

¹⁾This work was partly supported by the European Union under grant AST3-CT-2003-502818 and by the Austrian Science Foundation (FWF) under grant P-14445MAT

helps to improve transit time, i.e. the time the aircraft is parking on the apron, respecting all safety and security rules in providing automatic apron's conflict identifications. Therefore computer aided visual surveillance in this context helps to minimize costs for the company operating the airport, as personal can be deployed more efficiently, and to minimize latencies for the passengers, as the time needed for accomplishing ground services decreases.

Visual surveillance of an airport's apron has to face manifold problems: as an outdoor application a tracking system has to be robust according to weather and light changes. Ground support vehicles, as used on airports, may change their shape in an extended degree during the tracking period, e.g. a baggage car. Strictly rigid motion and object models may not work on tracking those vehicles. Another additional problem is that vehicles may build blobs, either with the aircraft or with other vehicles, for a longer period of time, e. g. a fueling vehicle during refueling process. Background maintenance and object detection algorithms have to be aware of these problems. In order to illustrate the complexity of the scene Figure 1 shows an apron with manually tracked vehicles and individuals.

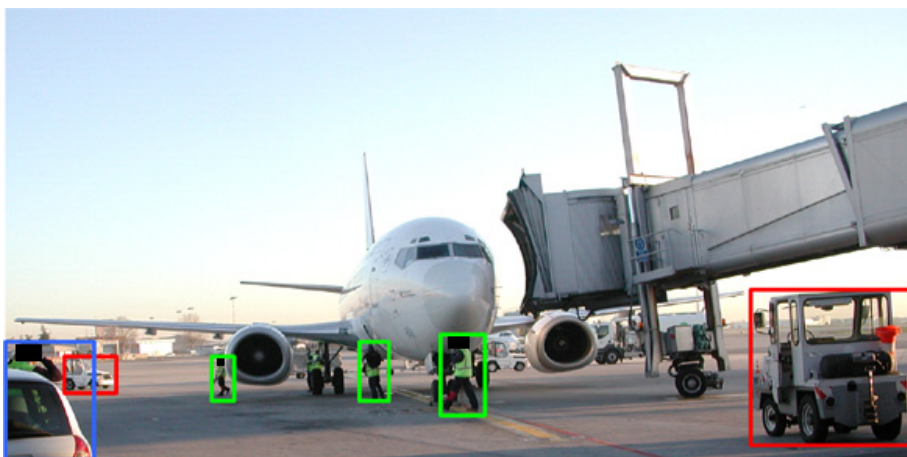


Figure 1: Airports apron: tracking of vehicles and individuals

The AVITRACK project addresses the specific case of automatically supervising commercial aircraft servicing operations at a designated airport. It is intended to add an additional technological layer to a specific airport in order to improve efficiency and security of the handling services. This paper is organized as follows: Section 2 describes the partners of the project and their main contributions. Section 3 addresses the specific case of visual tracking, especially background maintenance. At the end of the paper conclusions are drawn.

2 Partners and Main Objectives

To give a short overview on the project we introduce the partners in more detail and describe the main activities and their goals.

- **SILOGIC:** This company is an independent IT Services Consultancy at a very high technological level. It coordinates the project and will be the main partner in designing the activity model, in prototyping and in project management.

- **University of Reading:** The Computer Vision Group of the University of Reading has more than 20 years experience in the field of model-based vision and has produced one of the world's leading model based tracker. They will be the main partner in scene tracking.
- **Institut National de Recherche en Informatique & Automatique:** The National Institute is a French public-sector scientific and technological institute. They will be the main partner in scene-understanding.
- **Toulouse Airport:** This is the airport on which the AVITRACK system is installed. It has an essential role in the evaluation of the results.
- **FEDESPACE:** This NPO/NGO is coordinating major developments for the airport of the future, airplane and traffic control automation. It is involved as research co-ordination in Aeronautical and Space field. Fedespace will participate in the apron activity model and will co-ordinate the end users group composed of airports, airlines and airport handlers.
- **EURO-INTER:** The company gives advice in the European Union approach when dealing with research and technical development. It will design the dissemination strategy and organize the dissemination process of the achieved results.
- **IKT:** IKT System Partner AS sets it's activity in scientific information technology (like modelization, simulation and optimization) using radio frequency instrumentation. It will be involved in the development of the project's sensor devices.
- **TEKEVER:** This company is a European information technology SME, based in Portugal. The main activity is software development for embedded systems, Internet and wireless networks (GSM, GPRS, UMTS). It will participate in the development of the project's sensors and video devices, in prototyping and in validation.
- **ARC Seibersdorf GmbH:** ARC is Austria's largest Applied Research & Development Center with a broad spectrum of competencies. It will be the main partner of the project according to the development of sensors and video devices. It will further participate in prototyping and validation.
- **Vienna University of Technology, PRIP:** Scene tracking and scene understanding are the major concerns of PRIP in this project. It will further be involved in prototyping and validation.

The AVITRACK project is subdivided into eight clearly defined activities organized into eight work-packages (WP), namely: within **WP1 Apron Activity Model** the end-users needs are carefully designed and all apron activities are modeled. The users needs contribute to the specifications and are part of the final validation references. The model is the basis for the numerical simulation of all apron ground operations.

WP 2 Sensor & Video Devices represents all research and developments necessary for installing the equipment for the apron surveillance. The apron is observed by 8 sensors in order to see a maximum of operations around the aircraft. Figure 2 schematically shows an apron with the sensor positions.

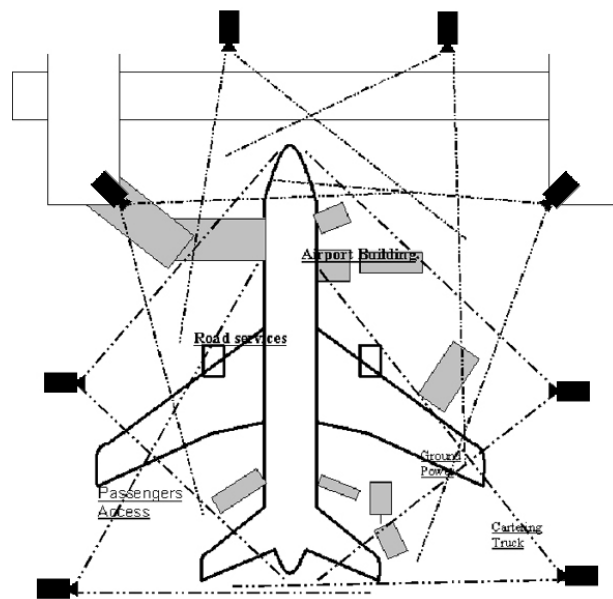


Figure 2: Airport apron: data acquisition

In **WP3 Scene Tracking** algorithms for people and vehicles tracking and recognition are developed. The algorithms are tested on recorded videos. The research concerns the efficiency of tracking several vehicles, people and/or aircrafts in a complex and outside scene. WP3 has to assure accurate results for tracking on a long period to allow robust interpretation of the scene recorded.

Within **WP4 Scene Understanding** algorithms to perform a high level interpretation of the results from WP3 are developed. The approach uses cognitive vision techniques based on spatio-temporal reasoning and use a priori knowledge of the observed environment (3D models of the empty scene) and predefined activity models (representative scenarios).

WP5 Prototyping represents the central point for implementing all produced software and algorithms. WP5 generates the necessary tools for video recording and the numerical model construction.

WP6 Validations & Perspectives verifies the achievements of the prototype. Each integration module precedes an evaluation task. These evaluations are independent of the developer teams.

WP7 Dissemination policy and exploitation plan is in charge of the dissemination of the project's results. A users group is formed. This group is consulted for professional user's needs specifications and to inform airports operators external to the project for awareness.

WP8 Project Management assures a level of international project management suitable to the European Commission and to the project partnership.

3 Visual tracking

Visual Tracking as our main contribution to the project is to analyze in real time the recorded videos in order to allow robust interpretation of the scene. The tracking tasks are organized as follows:

- *Motion detection*: The core of a surveillance system is the capability to perform reliable detection of targets in the image. In order to provide operations, the system must be capable of adapting to a wide range of environmental conditions (changing illumination and weather). In an outdoor setting, the effect of rain and the resulting reflections from the ground, snow and fog have to be considered. So one aim of the project is to analyze and apply robust, statistical techniques for motion detection which should be able to adapt to short, medium and long-term temporal and spatial changes in the scene.
- *3D scene reconstruction*: A 3D reconstruction of the apron area will be constructed, based upon known metric measurements in the world. A 3D-calibration tool will be designed for simultaneous calibration of multiple cameras. Statistical learning techniques will be used to learn calibration information based on monitoring of the scene over an extended period of time.
- *Object Categorization*: Identifying classes of object is an aim of the AVITRACK-project. For example, to differentiate between single individuals, groups of individuals and ground support vehicles.
- *Tracking of vehicles*: The aim is to robustly track vehicles (i.e. rigid motions) in the scene. Besides motion information a library of statistical 3D deformable models will be used to obtain more accurate ground-plane position estimated within a calibrated 3D world.
- *Tracking of individuals*: Robustly tracking of individuals (i.e. non-rigid motions) will be provided by an 2D statistical active shape model approach. Existing algorithms (e.g. [7]) will be extended. The position of people will be mapped into 3D in order to perform localization with respect to vehicles.
- *Complex scenes tracking*: Tracking of multiple objects (i.e. individuals and vehicles) simultaneously across multiple cameras is the concluding aim of the project. Effective motion detection, tracking of basic elements, 3D localization and elements recognition and categorization is necessary to achieve the project's aim.

Tracking of both, individuals and vehicles relies on an efficient motion detection algorithm. Background maintenance is a basic principle of most motion detection algorithms [10].

To start the tracking task in the AVITRACK project the following background maintenance algorithms were evaluated. The algorithms were implemented in C with OpenCV libraries and tested on a 2266 MHz / 512MB RAM PC running redhat Linux 9.0. The algorithms were tested on two datasets. One dataset consists of a 3000 frame sequence of extensively noised gray scale images with a resolution of 768x576 pixels showing a parking lot scenery (PETS2001 - DATASET 1, University of Reading). 10 reference frames were chosen from this dataset and foreground objects were

detected manually in those frames. The other dataset consists of a 100 frame sequence of gray scale images with a resolution of 348x288 pixels without initialization period showing a motor highway. 3 frames were chosen as reference frames and foreground objects were detected manually. These objects are compared with the foreground objects detected by the algorithms and false positive and false negative object pixels are counted and summed up over the chosen frames and compared with each other.

Median and Morphology: A reference image is computed as background model by taking the median of 21 consecutive images. Thus, moved objects can be better eliminated from the reference image. The reference image is updated according to a learning rate set by the user and background subtraction is performed. The difference image is improved by morphological filtering; morphological opening with a 3x3 structure element to eliminate isolated pixels is performed as well as morphological closing with a 5x5 structure element to fill the holes of moved objects. Our experiments have shown that this algorithm yields best results in terms of minimum total errors using our datasets.

Adjacent frame differencing: Each frame is subtracted from the previous one. To reduce noise the image is convolved with a 5x5-Gaussian kernel. A threshold is set manually for the difference image. Frame differencing yields to 51319 total errors in PETS2001 dataset and 15791 total errors in highway dataset. Nevertheless it faces the problem that only movement is detected, and as soon as an object does not move on it is recognized as background.

Adjacent frame differencing and morphology: The difference image is computed in the same way as in Frame Differencing. To achieve a better result the difference image is subject to morphological opening with a 3x3 structure element and morphological closing with a 5x5 structure element. Morphological operations improves the result (975 total errors less than frame differencing without morphological operations in PETS2001 dataset, 653 total errors less in highway dataset). However it faces the same problem concerning objects becoming motionless as frame differencing.

Mean and Threshold: A weighted sum of input images is used to calculate the background model. Background subtraction is performed and the threshold is set manually. The resulting difference image is subject to morphological operations as described in Frame Differencing and Morphology. In comparison with frame differencing segmentation quality decreases (2239 total errors more in PETS2001, 3663 total errors more in highway dataset). However, the problem of foreground objects becoming motionless is solved using this algorithm, due to the fact that the background model is calculated as a mean of more frames (depending on the learning rate of the weighted sum).

Mean and Covariance (Single Gaussian): The background model is established in the same way as Mean and Threshold. Additionally to the mean the covariance is calculated. The threshold is determined automatically by the standard deviation of each pixel. A pixel is considered foreground if it is out of the range of 2 standard deviations. Quality of background/foreground-segmentation

Algorithm	PETS2001 dataset				highway dataset			
	false pos.	false neg.	total	fps	false pos.	false neg.	total	fps
Median & Morphology	7456	41497	48953	21	5700	8959	14659	> 24
Frame Differencing	3600	47719	51319	> 24	6362	9429	15791	> 24
Frame Diff. & Morph.	2417	47927	50344	> 24	6062	9121	15138	> 24
Mean & Threshold	15820	36763	52583	24	10909	7892	18801	> 24
Mean & Covariance	29395	46520	75915	22	15818	8672	24490	> 24
W^4	36844	29086	65930	20	14863	6891	21754	> 24
Frame Diff. & Gaussian	116586	24944	141530	19	14507	8127	22634	> 24

Table 1: Performance results of the algorithms

decreases in comparison with a manually set threshold (23332 total errors more in PETS2001, 5689 total errors more in highway dataset), however there is no need to set a threshold manually.

W^4 : W^4 was first presented by Toyama et al. [6] in 1998, an improved algorithm was presented in 2000 [7]. In the experiments we implemented the background maintenance task of W^4 . For each pixel a minimum and maximum intensity value is calculated as well as the maximum inter-frame intensity change. If a pixel is outside of these parameters it is recognized as foreground. Morphological operations are performed to improve the result. Performance of the background maintenance part of W^4 is better than Mean & Covariance (9985 total errors less in PETS2001 dataset, 1856 total errors less in highway dataset) and has the same advantage according to thresholds.

Frame Differencing and Gaussian: Bramberger et al. [2] proposed a Gaussian background model in connection with frame differencing to detect long-term background change. The background distribution is adopted with regard to the observed distribution. The experiments have shown that this algorithm is not robust against noise (141530 total errors in PETS2001 dataset; 65615 more total errors than the Mean & Covariance and 92577 more total errors than Median & Morphology).

The results of the experiments are listed in table 1. Beside the tracking quality the average frame rate of the algorithms was subject to our research, the results are listed in tables 1, too. Initialization periods of the algorithms were not considered at calculation of the frame rate.

The results of the experiments, as they can be seen in the table 1, have shown that the high-level methods (W^4 and Frame Differencing & Gaussian) are not as robust against noise as simplistic methods (frame differencing, Median & Morphology, Mean & Threshold, Mean & Covariance). On images with less noise (highway dataset) the higher-level background maintenance systems perform better as it can be seen in table 1. On smaller images (384x288 pixels in our experiments) frame rates are higher than 24 frames per second for each of the implemented algorithms, on larger images (768x576 pixels in our experiments), frame rate of the frame differencing algorithms is higher than 24 fps, all implemented algorithms yields better than 19 fps. The Median & Morphology algorithm

presented in this paper yields tolerable performance and may be used as a starting point in the described project.

4 Conclusion

In this paper an overview of the AVITRACK-project was given. It was shown that the described project intends to enhance safety and security on the airport and to help to reduce the duration that a commercial aircraft is spending on an airport's apron. We presented an overview of the work-packages. As preliminary work a comparison of background maintenance algorithms was given. It is intended to make use of the results of AVITRACK-project on other airports to improve overall safety and security at the activity on an airports apron.

5 Acknowledgment

The authors want to thank the whole AVITRACK project proposal team, since an overview on such a project can only be given if techniques and ideas of many researchers involved are used. Especially we want to thank David Cher, Luc Barthelemy, James Ferryman, Monique Thonnat, Florent Fussier, Francois Bremond, Nuno Cunha, Jean-Louis Bonafe, Joao Portela, Norbert Brändle and Helmut Schwabach for their contribution.

References

- [1] D. Beymer, P. McLauchlan, B. Coifman, and J. Malik. A Real-Time Computer Vision System for Measuring Traffic Parameters. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1997.
- [2] M. Bramberger, R. P. Pflugfelder, B. Rinner, H. Schwabach, and B. Strobl. Intelligent Traffic Video Sensor: Architecture and Applications. In *Telecommunications and Mobile Computing Workshop on Wearable Computing*, Graz, Austria, March 2003.
- [3] M. Bramberger, R. P. Pflugfelder, B. Rinner, B. Strobl, and H. Schwabach. A Smart Camera for Traffic Surveillance. In *Telecommunication and Mobile Computing Workshop*, Graz, Austria, March 2003.
- [4] J. Ferryman, A. Worrall, G. Sullivan, and K. Baker. Visual Surveillance Using Deformable Models of Vehicles. *Robotics and Autonomous Systems Journal*, 19:315 – 335, 1997.
- [5] W. Grimson. Using Adaptive Tracking to Classify and Monitor Activities in a Site. In *IEEE Conference on Computer Vision and Pattern Recognition*, Santa Barbara, CA, 1998.
- [6] I. Haritaoglu, D. Harwood, and L. S. Davis. w^4 : Who? When? Where? What? a real time system for detecting and tracking people. In *3. International Conference on Face and Gesture Recognition, Nara, Japan*, Apr 1998.
- [7] I. Haritaoglu, D. Harwood, and L. S. Davis. W^4 : Real Time Surveillance of People and their Activities. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 22(8), pages 809–830, 2000.
- [8] D. Koller, J. Weber, and J. Malik. Towards Realtime Visual Based Tracking in Cluttered Traffic Scenes. In *IEEE Intelligent Vehicles Symposium*, pages 201–206, 1994.
- [9] R. P. Pflugfelder. Visual Traffic Surveillance Using Real-time Tracking, Diploma Thesis. Technical Report PRIP-TR-071, PRIP, TU Wien, 2002.
- [10] K. Toyama, J. Krumm, B. Brumitt, and B. Meyers. Wallflower: Principles and Practice of Background Maintenance. In *ICCV (1)*, pages 255–261, 1999.