

AVITRACK



Contract n° AST3-CT-2003-502818

D.6.3 AVITRACK Validations



Table of contents

1. INTRODUCTION	3
2. WP 1, APRON ACTIVITY MODEL	3
2.1. Geometric models	3
2.2. activities model.....	4
3. WP 2, SENSORS & VIDEO DEVICE.....	4
3.1. video sensor network	4
3.2. video sensor position	4
4. WP 5, PROTOTYPING	5
4.1. prototype architecture	5
4.2. Hardware prototype	5
4.3. HCI reporting	5
5. WP 3, SCENE TRACKING	6
6. WP 4, SCENE UNDERSTANDING.....	8
7. WP 6, EVALUATION PROCESS.....	9
8. CONCLUSION	9

1. INTRODUCTION

After intermediate evaluations concerning both tracking and understanding functions, and after the human control interface realisation, a global evaluation is processed to check the global operations AVITRACK prototype.

The final review is to analyse the achievements of the research and the prototype efficiency with regards to predicted project achievements. Coherence of the interpretation of the most complex scenarios is checked. The pertinence of the extracted information is analysed.

The capability of the prototype to produce robust and adequate reports regarding the current apron activity is to be validated.

The efficiency of the AVITRACK Prototype is verified. The improvements related to the report's note of the previous evaluation are checked.

The final evaluation checks:

- Robustness of the prototype,
- The complex operation understanding,
- Generation of automatic report activities,
- The process for all the duration of a handling operation,
- The technical evolutions,
- The operational functions,
- The users needs acceptance,
- The analysis of all activities operations treatment.

The final validation checks the fit between users-needs expression and achievement.

The following chapter presents reflection concerning each of the technical tasks, contributing to the AVITRACK prototype achievement.

2. WP 1, APRON ACTIVITY MODEL

2.1. GEOMETRIC MODELS

To design the position of the assembly of the cameras, a 3D model of the apron area has been realised to simulate the point of view of each camera. Considering that the obtained videos have been useful for all the project duration, we can consider that demarche has been successful.

The exploited software has all been used to build the 3D models of the vehicle operating on the apron area. These models are used to operate the vehicle automatic recognition involved in the events understanding.

Based on these 3D models, functional areas are defined to authorise the details events recognition. The same demarche is applied for the aircraft area definition, using a CAD model.

The exploitation of a tool creating 3D model based on several images of the objects, assure the adaptability of the process at different places and applications: the same demarche can be applied on other sites. The resolution and the precision of designed models appear suitable to realise the expected functions.

These 3D models are also used in the HCI: the 3D view contains the visualisation of the apron including 3D models of vehicles.

2.2. ACTIVITIES MODEL

Working in collaboration with the Toulouse airport actors, a handling operation model has been designed. An original approach led to build a generic representation of the handling operation based on UML diagrams. These diagrams were the foundation to structure the detailed scenarios to be recognised. The apron activities UML model reached its three objectives to become the description utility, the reference to structure the end-users needs, to guide to the operations understanding.

All along the project, scenarios and envisaged events recognition have been extracted from these documents, offering the potentiality of necessary evolution and adaptation on others apron site and handling operation control.

3. WP 2, SENSORS & VIDEO DEVICE

3.1. VIDEO SENSOR NETWORK

Starting the project a complete and detailed study reach to select a specific camera type, fix lens and to define a dedicated video network including encoder and storage units.

The quality of the video recorded during the project, the reliability of the equipment, facility of use and the accessibility offer by the airport, provide a strong added value to the project. Avitrack partners dispose of an efficiency and appropriate real test site. Closed to a real video-network for final users, including some technical particularities for research topics as the weak compression rate, high video sensor quality, the AVITRACK Toulouse-Blagnac installation offers the advantage of the industrial equipment in a real environment.

The results of the tracking software, exploiting the obtained images, for the 2 years of the project, highlight the pertinence of the initial choices and the installation efficiency.

3.2. VIDEO SENSOR POSITION

At the beginning of the project the positions of the assembly of the cameras have been fixed. 8 cameras have been installed on the Echo-40 apron on the Toulouse-Blagnac airport. Covering all the areas around the aircraft and overlapping points of view were the main criteria.

The different combinations of cameras allow:

- To observe the handling operations variety on the right side of the aircraft.
- To propose a different point of view of the aircraft front area and its specificities to observe the jet bridge stair area validate its position.
- To provide the necessary point of view for the refuelling operation control.
- To observe the objects, which they are ahead scene, under the plane or at the bottom of the scene are observed.
- The detection of the "pushback" vehicle.

From what has been observed during the project, one can think to reduce the number of cameras used to monitor the turnover. An industrial deployment could typically consist of 4 cameras per apron. Of course this depends a lot of the apron geometry and in some cases, this number will not be sufficient. It also depends greatly on the scenarios you are tracking.

4. WP 5, PROTOTYPING

4.1. PROTOTYPE ARCHITECTURE

Starting the project, a distributed architecture has been designed to allow all the development, modules integration, and evolutions setting up the prototype. The efficiency and the equation of the AVITRACK prototype architecture to the specification has been highlighted during the prototype evaluation.

This architecture provides all the expected technical functionalities, in an industrial environment. The platform is running on standard PC, which could be used for an industrial application, and can be deployed by each of the partner on its own equipment.

The AVITRACK prototype constitutes a strong base for any evolution, resulting from next research efforts, and a suitable base to envisage technical transfer for industrial solution.

4.2. HARDWARE PROTOTYPE

For the implementation of all the software set up in the Avitrack framework, a common hardware architecture has been designed. This hardware architecture is the result of a specific study, which has been led after the prototype architecture design considering the first and planned algorithms characteristics.

The number of camera, the real time processing constraints and the modularity were main criteria. The selected solution based on dual processor offers all required technical specifications. The dimension of the proposed solution has satisfied the implementation of all modules all along the project and the real time processing constraint.

Based on standard PCs configuration, the Avitrack prototype is operational, running in real time and able to be deployed in any configuration by each partner.

4.3. HCI REPORTING

A large and specific HCI has been designed for the AVITRACK prototype. The ambition of the HCI was to cover all the necessary functions imposed by the several technical modules. AVITRACK intelligent video scene understanding involves several specific modules, deployed on a distributed architecture. The challenge for the interface was to satisfy all the requirements for all the particular modules and their communication. Step by step and resulting from an efficient partners communication, the complete interface has been set up.

The interface proposes the control of the video server reading the available video channels, displays the video, displays the results of the tracking module containing all the objects description, offers the 3D view of the scene which is computed in real time and finally shows the understanding module results as scenario recognised or a specific end-users interface illustrating the schedule of the turn over and the handling operations progress.

The complete interface, displaying the results of all the modules, constitutes a useful tool setting up the prototype. The next step will be to design the end-users interface for a specific application, exploiting the prototype results.

5. WP 3, SCENE TRACKING

Motion Detection

The field of automated visual surveillance has undergone significant advancement in the last few years, with numerous methods developed by the computer vision community applied to a wide range of settings and environments.

An integral first step of visual surveillance is Motion Detection. Motion Detection algorithms attempt to locate connected regions of pixels that represent the moving objects within the scene.

Motion detection must be capable of adapting to a wide range of environmental conditions, such as lighting changes and weather conditions.

The work performed by the computer vision community on motion detection algorithms can be broadly classified into two types:

- Methods that use some form of frame to frame differencing to detect changes from one frame to the next,
- Methods that use a background model to represent the scene and detect motion by finding changes between the current frame and the background model. The latter are called background subtraction algorithms, they are the most commonly used in visual surveillance.

Frame differencing methods, while able to handle sudden light changes, suffer from problems such as not being able to fully detect objects. For this reason, frame-differencing methods tend not to be used in environments such as the apron environment of AVITRACK.

As part of the work performed on motion detection for the AVITRACK project, several of the motion detection methods were implemented in the AVITRACK project framework. The choice of methods was determined by taking into account their usefulness for the apron environment and considering the real-time constraints. The selection tried to cover all the main approaches to motion detection available currently in the visual surveillance community. The methods implemented were evaluated to determine the method that works best for the apron environment.

These are:

- W4
- Median and Morphology
- Kalman Filter
- Chromatic Background Model
- Mixture of Gaussian
- Colour Mean and Variance
- Colour and Edge Fusion
- Kernel Density Estimation
- Linear Prediction

These motion detection algorithms are described in more detail in deliverable "D 3.1 Motion Detection", and the results of the quantitative analysis performed for the AVITRACK project are given in deliverable "D 6.1a Prototype Scenes Tracking Evaluation".

From the evaluation performed so far on motion detection in the apron environment, the 5 chosen algorithms:

- Median & morphology,
- Kalman filter,
- Colour Mean & Variance,
- Gaussian Mixture Model,
- RGB mean value

were found to be robust to illumination changes, with the exception of sudden and strong changes.

The algorithms also appear to need only a few frames for the background to be adapted correctly. The main weak points are: the detection of shadow as part of objects; objects are partially detected or

fragmented when the background and the objects have low colour and are similar. Both these problems occur because the apron environment and the objects present on the apron tend to have low chromaticity.

SHADOW Detection

In real-world environments such as AVITRACK the detection of objects is downgraded by erroneous pixel labelling due to sensor noise, dynamic illumination changes etc. Near-static local illumination changes can be incorporated into the background model; dynamic illumination changes due to moving objects present a greater challenge.

The current shadow detection work in the AVITRACK project follows two distinct approaches:

- Statistical learning using *a priori* knowledge: the statistical learning method uses *a priori* knowledge about the scene objects to distinguish between the objects and the shadow; the colour probability density function of this detected shadow is modelled using a non-parametric statistical representation, which can then be used to classify the remaining shadow regions.
- Explicit shadow modelling: The explicit shadow modelling uses geometric models of the objects and the light sources to predict regions of shadow.

Object categorisation

Object categorisation can be considered as the process of assigning class ownership to moving objects. Subsequent object behaviour analysis strongly depends on reliable object classification in order to have a firm basis upon which to act.

Scene Tracking

Automated object tracking is a widely studied problem and several methods and algorithms are available in the visual surveillance community.

Generally, tracking methods can be classified into two strategies:

- A top-down or model-based tracking: use of knowledge of what objects will appear in the scene, and given a 3D model of the object, hypothesis are generated and tested against the image for the presence of that object. One such system is the work done for the ESPRIT VIEWS project.
- A bottom-up or low-level tracking: the image is segmented into a number of objects and uses only low-level information available in the image to identify and track objects, such as the work of Collins et al.

The initial approach selected for AVITRACK was to use a bottom-up approach for tracking, as algorithms of this type are normally more generic in nature and require little contextual information about the scene. Model-based methods are generally slower and were not considered to satisfy the real-time processing demands of AVITRACK. It is envisioned that future work will address the issue of combining both approaches together, so that the advantages of one approach will complement the other. Four different bottom-up algorithms were implemented:

- **Region Based Tracker**

The method identifies and extracts regions of interest in an image, “blobs”, and tracks them over time. A connected components algorithm is used to connect mobile pixels identified during the motion detection phase into respective blobs. Significant and persistent object features model the blobs, in order to handle occlusions.

The following features have been calculated for each blob:

- Position: centroid coordinates
- Shape: the best-fit ellipse is calculated.
- Dominant colour in RGB space

The method is able to handle partial occlusions and temporal lost objects.

- **Local Feature based Tracker**

This method is based on the Kanade-Lucas-Tomasi (KLT) feature tracker, and it tracks an object by selecting a set of sparse local features for that object and tracks the features from one frame to the next.

- *Colour based Tracker*

This method uses an object's histogram as the global colour model for tracking.

- *Cluster Tracker*

This method is based on the work of Pece, which is based on grouping clusters of moving points.

These algorithms are described in more detail in deliverable "D 3.4a Scene Tracking". Initial evaluation of these methods give good results for most of the trackers. Some problems include loss of identity during occlusion, and tracking errors due to low colour information.

Data Fusion

In the AVITRACK project the Nearest Neighbour method has been applied and analysed, a fundamental problem in this work is not the actual fusion mechanism but the object localisation and association difficulty due to interactions between moving and static objects. The data fusion has been achieved using a formal statistical approach to combine the measurement uncertainties into a fused estimate for each target. Current work is focussed on the JPDAF filter which gives a probabilistic estimate of the data association. This is expected to resolve some of the difficulties experienced due to the sensitivity to the order of processing in the Nearest Neighbour filter. Object localisation is related to the problem of categorisation in that the object type and pose is required before a meaningful location estimate can be recovered, precise localisation can be achieved using the fused result from the data fusion.

Ghost detection

A *ghost* is defined as a set of connected pixels, detected as a mobile object but not corresponding to any real moving object.

An object is integrated into the background when becomes stationary. In these cases, ghosts are created when stationary objects start to move again. Furthermore, ghosts are produced when parts of the background start moving.

A movement density measure introduced by Ruiz-del-Solar *et al* is adopted to detect ghosts in the scene.

Movement Density Module

The movement density module receives the detections from the motion detection algorithm. Movement pixels are identified and connected by means of 8-connectivity into blobs. For each blob is defined a movement density value. The movement density measures for the blob the average change in the last frame. Ghosts should have a low movement density, while the moving objects should have a larger movement density.

6. WP 4, SCENE UNDERSTANDING

The aim of Scene Understanding is to provide a high level interpretation of the tracked mobile objects trajectories in term of human behaviours, vehicle activities, or their interactions. This process consists in detecting video events, which have been learned through examples or predefined by application experts.

Two main categories of approaches are used to recognise video events based on:

- A probabilistic/neural network
- A symbolic network corresponding to the events to be recognised.

Vu et al have combined the previous approaches to optimise the temporal constraint resolution by ordering in time the components of the video events to be recognised.

We have extended this last method to address Complex Activity recognition involving several physical objects of different types over a large space observed by a camera network and over an extended period of time.

The results obtained in the AVITRACK project show that new complex activity recognition can be handled and challenged with success. The several publications written during the project will contribute to the state of the art of next video understanding systems, which will probably deal with more and more complex scenes above what was done in AVITRACK.

7. WP 6, EVALUATION PROCESS

Specific evaluation tasks were planned during the AVITRACK project, regularly checking the achievements and the milestones. This process has been led efficiently by the concerned technical partners. At each step, jointed to the technical meetings, evaluations highlighted reached achievements, necessary improvements and fixed the next prior steps for research. The final achievements have taken a great advantage of the demarche realised in with transparency and positive way.

8. CONCLUSION

Globally, each technical objectives of the Avitrack project has been reached: Operational prototype, software, test site, video network.

Constituting an innovative approach for the handling operation supervision, the interest of all airport actors is evident. The robustness and the design of specific operations to be check should be the next steps, exploiting the AVITRACK results.

The way of exploitation, research and dissemination are exposed in the “integration report” (D6.4 ref DL-AVI-1-075).