Abstract—In this paper a framework for an airport ground traffic surveillance system based on visual information is presented. The framework is designed to improve state-of-the-art Advanced Surface Movement Guidance and Control Systems (A-SMGCS) adding an image processing module able to act as a “gap-filler” for airport zones that traditional systems fail to reliably cover due to reflections and blind spots. Results of the approach based on a set of videos is presented demonstrating that the proposed framework is a flexible and cost-effective solution for improving existing systems’ performance.

Index Terms—air traffic control, aircraft detection, object recognition, A-SMGCS.

I. INTRODUCTION

Advanced surface movement guidance and control systems (A-SMGCS) aim at providing routing, guidance and surveillance for the control of aircraft and vehicles in order to maintain the declared surface movement rate under all weather conditions within the aerodrome visibility operational level (AVOL) while maintaining the required level of safety [7]. The necessary condition for all the algorithms (routing, guidance, and conflict avoidance) to work correctly is the provision of reliable surveillance data in all airport areas.

Usually, Surface Movement Radars (SMR) are used as the main source of information for estimating aircraft positions. However, in complex airport layouts (see Fig. 1), traditional radar-based surveillance systems are affected by limitations in their coverage due to reflections or shadows that are caused by buildings, equipment or other reflecting objects on the airport surface.

Cooperative systems for airport ground traffic surveillance, such as the Automatic Dependent Surveillance-Broadcast (ADS-B) system [12], can be involved for mitigating the problems. An ADS-B equipped aircraft determines its own position using a global navigation satellite system (GNSS) and periodically broadcasts its position as well as other relevant information to potential ground stations and to other aircrafts with ADS-B equipment.

Nevertheless, the adoption of the ADS-B system does not provide reliable data, since the above mentioned limitations can also arise in the GPS signal propagation used by the aircrafts (or vehicles) to compute their position. Furthermore, not all the aircrafts are equipped with ADS-B transponders.

When the GPS signal is not available, it is possible to adopt a Multilateration (MLAT) approach. MLAT uses a grid of ground receivers to record ADSB reports and to perform a time-difference-of-arrival analysis to estimate the aircraft position. However, MLAT can only detect cooperative targets since the aircraft transponder has to be switched on. In case of malfunction or if the transponder is switched off, the system cannot detect the aircraft [5].

For all the above discussed reasons, it is very difficult to provide a reliable coverage in all the airport areas without increasing the number of surveillance sensors (SMR radars or MLAT sensors). Moreover, electromagnetic pollution caused by radars is something that must everywhere be reduced to a minimum, therefore a “green” technology should be used to get the exact position of all parked aircrafts reliably in order to provide the airport controllers a complete view of the situation at hand.

A possible solution for increasing the overall system coverage as well as for reaching higher levels of the A-SMGCS specification [7] is to introduce a system able to act as a “gap-filler” for the airport zones that traditional sensors fail to reliably cover. Such a system should be less expensive than the electro-magnetic sensors used in the state-of-the-art systems.
Conclusions are drawn in Section V.

Both the detection and visual tracking modules are provided.

III. In Section IV a series of examples showing the output of the-of-art ground traffic surveillance system is provided in increasing the overall system coverage.

The paper is organized as follows. A classification of state-of-the-art ground traffic surveillance system is provided in Section II and the proposed framework is described in Section III. In Section IV a series of examples showing the output of both the detection and visual tracking modules are provided. Conclusions are drawn in Section V.

II. RELATED WORK

The existing systems for the monitoring of surface movements at airports can be classified taking into account whether the targets are cooperative or not [10].

Cooperative targets are equipped with a transponder that communicates its precise location to the surveillance system. There exist two main cooperative technologies:

- In ADS-B systems the aircrafts automatically broadcast their flight information to the other ADS-B capable systems. That information includes the geographic coordinates, the current speed and the flight identification (or the aircraft address). The communication of the updated location of the aircraft allows for an accurate and continuous surveillance on the target movements.

- MLAT system consists of a set of antennas installed in the operational area that broadcast query signals to the airspace. The aircrafts equipped with an adequate transponder answer with a Secondary Surveillance Radar (SSR) or a Mode-S reply signal. The round trip time between the antenna and the target is then used for estimating the position of the aircraft.

Non-cooperative targets do not communicate any information to the surveillance system. Their positions are calculated without the need of installing any electronic device within the aircraft. There exist three main non-cooperative technologies:

- SMR is based on the transmission of radio frequency signals that are echoed by moving objects. Typically, the transmitter is a rotating antenna that broadcasts electromagnetic waves that are rejected by any object present on their path. Then, the rejected signal is received by a sensor that measures the location, shape and direction of the moving object. Depending on the size of the monitored area, one or more antennas should be used in order to cover the entire surface. The cost of an SMR system is quite high.

- Magnetic Sensing is a technology that allows detecting the passage of ferromagnetic objects (e.g., vehicle motors, aircraft components) on the basis of their interaction with the earth’s magnetic field [5]. The sensors are usually installed on the ground and are able to detect the local change of the earth’s magnetic field produced by ferromagnetic objects. Moreover, analyzing the magnetic signature it is possible to differentiate among different classes of targets (the signature of a vehicle is usually stronger than an aircraft due to its closer distance to the ground [10]). An European project, ISMAEL [4], used magnetic sensing technology to improve the capability of larger airports to detect aircrafts and to provide a stand-alone and inexpensive solution to small and medium airports.

- Image-based tracking system are designed to use cameras in order to detect aircrafts. The main advantages in using video surveillance are: 1) the chance of using very low cost equipment 2) the possibility to make use of existing equipment already installed in the airport premises. There exist various video surveillance system for airport monitoring. The INTERVUSE project [11] proposes to use a detection scheme derived from a successful road traffic monitoring system that receives images from a network of cameras. The TRAVIS [9] system consists of a scalable network of autonomous tracking units that use cameras to capture images, detect moving objects and provide results to a central sensor data fusion server (SDF). The SDF server is responsible for tracking and visualizing moving objects in the scene as well as collecting statistics and providing alerts for dangerous situations. The TRAVIS prototype system was installed at Thessaloniki (Greece) airport for monitoring the traffic at aircraft parking areas. In [3], a system for detecting and tracking the aircraft and land vehicles is presented. The identification system recognizes the characters that form the aircraft identification number, in order to assign a correct label to each track. The main drawbacks of the above mentioned systems are the use of a detection scheme based on the creation of a background model that limits the detection to moving objects only and the need of calibrated cameras.

In this paper an image-based tracking system will be presented. The system aims at using visual information for improving the accuracy of the aircraft detection and tracking in parking areas. In particular, differently from other systems, we propose to use a classifier to detect the aircraft, in order to be able to detect not moving objects also. Moreover, our approach does not need calibrated cameras.

III. COVERING BLIND SPOTS

Parking zones in airports (Fig. 1) represent a challenging scenario for automatic video surveillance due to [3]:

- quite fast changes of illumination (e.g., due to cloud movement);

- reduced visibility caused by adverse meteorological conditions (e.g., rain, fog);

- the presence of not moving objects of interest. Indeed, the relevant targets should be detected both while moving and while they are steady (e.g., parked aircrafts).
Moreover, in order to limit the number of sensors, this work is based on the use of PTZ cameras. Since PTZ cameras can be moved (automatically or manually), the conventional foreground/background modeling approach is ineffective being based on the assumptions that all the targets of interest are moving and that the camera is static.

To overcome the above mentioned limitations, the proposed system architecture makes use of an aircraft detection module based on a classifier (see Fig. 2).

The Video Processing Unit (VPU) is designed to detect and track aircrafts. It takes as input a single image coming from the video stream (current frame). The Data Fusion (DF) module aims at associating the tracks coming from the VPU (visual tracks) and the tracks elaborated by the Multi Sensor Data Fusion (MSDF) unit (system tracks). Indeed, the MSDF module collects and computes all the data coming from the existing electro-magnetic sensors (ADS-B, SMR, MLAT). In the reminder of this Section, each module of the framework is described in details.

A. Aircraft Detection

The detection module aims at finding the aircrafts in the current frame obtained from the PTZ camera. Since the camera is controlled by a human operator, a foreground/background modeling approach is ineffective. Thus, we decided to adopt a classifier based detection. In order to obtain a fast (quasi-realtime) detection, a Haar-like feature-based approach [13] has been chosen.

However, the Haar-like classifier has been originally designed for face recognition, thus we had to verify the applicability of the method for aircraft detection creating the classifier.

To this aim, we used the OpenCV HaarTraining functions. A set of 1000 images not containing aircrafts and a set of 550 images including different types of aircrafts (with different views) taken from the Internet have been used as input for the off-line training stage obtaining a 15 levels classifier (the training was stopped when the false alarm rate reached $5 \times 10^{-6}$). We experimentally found that good detection results can be obtained using a search window of 90×30 pixels. It is worth noting that since the aircrafts can be viewed by different angles, the dimension of the search window is a crucial parameter for creating an accurate classifier.

The output of the aircraft detection module is a list of observations. Each observation is a bounding box representing a detected aircraft. In Fig. 3 some examples of detection of different types of aircrafts are shown.

B. Visual Tracking

The visual tracking module takes in input the image flow and the output of the detection module and returns a set of visual tracks, i.e., bounding boxes with an identification number. Its main role is the temporal filtering of the false positives. In order to deal with the movements of the camera, we decided to use a particle filter (PF) based approach [2].

PF [8] is a three step online procedure for Multi-Target Tracking (MTT): 1) In the sampling step, several events (particles) describing the state of the system, i.e., the displacement of the targets by sampling a candidate probability distribution (pdf) over a state space, are hypothesized; 2) In the update step, dynamic is applied to the particles; 3) In the observational step each hypothesis is evaluated given the observations of the system. The best observations of the system and the best fitting ones are selected, thus avoiding brute-force search in the prohibitively large state space of the possible events. In this way, the candidate pdf is refined for the next filtering step.

PFs offer a probabilistic framework for recursive dynamic state estimation that fits MTT [1]. The goal is to determine the posterior distribution $p(x_{t}|z_{1:t})$, where $x_t$ is the current state, $z_t$ is the current measurement, and $x_{1:t}$ and $z_{1:t}$ are the states and the measurements up to time $t$, respectively. We denote as $x_t$ the state of a single object, and $x_t = \{x_1^t, x_2^t, \ldots, x_K^t\}$ the joint state (for all objects).

The Bayesian formulation of $p(x_{t}|z_{1:t})$ and the Chapman-Kolmogorov equation enable us to find a sequential formulation of the problem:

$$p(x_t|z_{1:t}) \propto p(z_t|x_t) \int_{x_{t-1}} p(x_t|x_{t-1})p(x_{t-1}|z_{1:t-1})dx_{t-1}$$  (1)

The PF is fully specified by an initial distribution $p(x_0)$, the dynamical model $p(x_t|x_{t-1})$, and the observation model $p(z_t|x_t)$. The posterior distribution at previous time $p(x_{t-1}|z_{1:t-1})$ is approximated by a set of $N$ weighted particles, i.e., $\{(x_{t-1}^{(n)}, w_{t-1}^{(n)})\}_{n=1}^{N}$, because the integral in Eq. (1) is often analytically intractable. Equation (1) can be rewritten using the Monte Carlo approximation:
The update of the weights is computed according the following relation:

\[
p(x_t|z_{1:t}) \approx \sum_{n=1}^{N} w_{t-1}^{(n)} \delta(x_t - x_t^{(n)}). \tag{2}
\]

where \( q \) is called proposal distribution. The design of an optimal proposal distribution is a critical task. A common choice is \( q(x_t^{(n)}|x_{t-1}^{(n)}, z_t) = p(x_t^{(n)}|x_{t-1}^{(n)}) \) because simplifies equation (3) in \( w_t^{(n)} \propto w_{t-1}^{(n)} p(z_t|x_t^{(n)}) \). Thus, the weight at the current time is updated using the weight at the previous time and evaluating the likelihood of the observation with respect to the hypothesis \( x_t^{(n)} \).

In our implementation, the prior distribution \( p(x_0) \) is given by the output of the detector, a set of 100 windows (particles) around the current state are considered to compute the update step (see Fig. 4), and an HSV histogram-based model is used in the observational step [6].

C. MSDF System

The Multi Sensor Data Fusion (MSDF) unit is the part of the system in charge of fusing data coming from all the radar devices (in the case of A-SMGCS systems, radar sources are MLAT, ADS-B, and SMR). MSDF algorithms try to fuse the contributions of the different sensors in a single system track (see Fig. 2).

Fusion algorithms could be based on single track kinematic, but on the knowledge of the environment also. For example, one might adapt the tracking model to the different airport zones or to assign different associations weights to the sensors based on the airport zones. A possible strategy consists in assigning a reduced weight to the MLAT position in the airport zones, such as the parking bay, where the MLAT is known not to work properly (the system could be also adaptive).

D. Data Fusion

Even with complex fusion schemes, problems may arise from the fact that, in some areas, the radar sensors providing target position are limited (e.g., for not ADS-B aircrafts in the parking zones not covered by an SMR). In those cases, the use of an EO gap fillers could be very useful.

The module responsible to add the EO contribution to the system tracks is the Data Fusion module. In this case, the fusion to be performed is not the same one produced by the MSDF system (based on kinematical parameters), in the sense that EO systems usually provide other features (e.g., state of motion, color, type).

IV. EXAMPLES

In order to test the proposed image-based detection approach, we used 12 videos downloaded from the Internet. The classifier is able to detect targets of different sizes and in presence of partial occlusions (see Fig. 5).

We used a feature extraction function to improve the detection quality. Indeed, a Haar-like feature-based approach typically suffers of an high false alarm rate, that is a measure of the number of false detection made by the classifier. The idea is to extract key points (e.g., SIFT or SURF) to validate an observation. When the classifier finds a possible target, the bounding box is analysed in order to extract the key points. If the number of key points present in the bounding box is below a threshold, then the observation is rejected. In Fig. 5 the above presented algorithm can be used to filter out a false positive detection (black bounding box).

To initialize the PF, we automatically selected strong observations as follows. Given a frame at time \( t_0 \), we put all the observations in a set named \( obs \). At time \( t_1 \), we try to associate the new observations with the previously registered ones in \( obs \), i.e., the old observations. The association between old observations and new observations is made on the basis of a nearest-neighbour policy with the Bhattacharyya distance between the HSV value histograms of the observation belonging to \( obs \) and the current new observation as measure. If a new observation is associated to an old one, then the observation is maintained in \( obs \) and a counter is updated. A
strong observation is an observation with a counter that has become greater than a predetermined threshold (e.g., 10).

In Fig. 6 three frames taken from one of the considered video sequences are shown. In the first row the detection results are reported, while the visual tracking results obtained using a PF-based approach are shown in the second row.

After the initialization, the particle filter is able to track the portion of the image defined by the observation’s bounding box. We use an HSV histogram-based model as observational model. Using a PF-based approach makes the tracker robust with respect to missing detections, since the PF continues to evolve even in absence of observations. As an example, the third frame in the first row of Fig. 6 presents a missing detection: in this case, the particle filter continues to correctly track the aircraft. It is worth noting that when a new observation is available, the PF tracker is reinitialized.

The performance of the algorithm in terms of frame per seconds (fps), using an Intel Core 2 Duo SU7300 CPU 4 GB RAM, is shown in Table I.

<table>
<thead>
<tr>
<th>Frame Dim.</th>
<th>FPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>320 × 180</td>
<td>16.26</td>
</tr>
<tr>
<td>320 × 240</td>
<td>14.54</td>
</tr>
<tr>
<td>640 × 360</td>
<td>5.23</td>
</tr>
</tbody>
</table>

V. CONCLUSIONS

In this paper a framework for an airport ground traffic surveillance system based on visual information has been presented. The framework is designed to improve state-of-the-art Advanced Surface Movement Guidance and Control Systems (A-SMGCS) adding an image processing module able to act as a “gap-filler” for airport zones that traditional systems fail to reliably cover due to reflections and blind spots.

Results of the approach based on a set of videos is presented demonstrating that the proposed framework is a flexible and cost-effective solution for improving existing systems performance.

As future work we intend to quantitatively evaluate the detection and tracking modules using publicly available dataset. In order to deal with the weather issues, such as heavy rain or snow, and the issue of night-time we will add infrared cameras to the system.

REFERENCES


