Aerodrome situational awareness of unmanned aircraft: an integrated self-learning approach with Bayesian network semantic segmentation

Bowen Lu[1], Matthew Coombes[2], Baibing Li[3] ✉, Wen-Hua Chen[2]

[1]*Zhong-An Information and Technology Services Co., Ltd., Shenzhen, People's Republic of China*
[2]*Department of Aeronautical and Automotive Engineering, Loughborough University, Loughborough LE11 3TU, UK*
[3]*School of Business and Economics, Loughborough University, Loughborough LE11 3TU, UK*
✉ *E-mail: b.li2@lboro.ac.uk*

Abstract: It is expected that soon there will be a significant number of unmanned aerial vehicles (UAVs) operating side-by-side with manned civil aircraft in national airspace systems. To be able to integrate UAVs safely with civil traffic, a number of challenges must be overcome first. This study investigates situational awareness of UAVs' autonomous taxiing in an aerodrome environment. The research work is based on a real outdoor experimental data collected at the Walney Island Airport, the UK. It aims to further develop and test UAVs' autonomous taxiing in a challenging outdoor environment. To address various practical issues arising from the outdoor aerodrome such as camera vibration, taxiway feature extraction, and unknown obstacles, the authors develop an integrated approach that combines the Bayesian-network based semantic segmentation with a self-learning method to enhance situational awareness of UAVs. Detailed analysis of the outdoor experimental data shows that the integrated method developed in this study improves the robustness of situational awareness for autonomous taxiing.

1 Introduction

Unmanned aerial vehicles (UAVs) are increasingly used for various civil applications (e.g. monitoring gas pipelines [1] and surveillance of electrical power infrastructures [2]). It is expected that soon there will be a significant number of UAVs operating side-by-side with manned civil aircraft in civil airspace systems such as the U.S. National Airspace System of the Federal Aviation Administration [3]. The biggest challenges are the safe and effective integration of UAVs into the existing airspace [4]. This paper considers the ground surface operation safety in aerodromes and in particular, we investigate an important safety issue, UAVs' situational awareness during autonomous taxiing in aerodromes.

In the existing repository, some autonomous taxiing research has been done using the Global Hawk aircraft with DGPS and highly accurate maps to guide the aircraft in a segregated military controlled airport [5]. Nevertheless, such a system would need continuous supervision by a remote pilot as no autonomous obstacle avoidance mechanism was implemented in it and will not be able to operate without DGPS corrections or in GPS denied environments. Hence, there is a real need to develop autonomous taxiing systems for UAVs so that they can autonomously take off and land without human pilots' intervention.

Unlike military UAVs, civil UAVs are unlikely to have their own specialised aerodromes but more feasibly sharing the existing civil aerodromes with manned aircraft [6, 7]. Compared with military aerodromes, civil aerodromes can be much less tightly controlled, with greater unpredictability in ground movements of vehicles and aircraft. Since such dynamic information is not fully registered in the ground traffic control (GTC) systems, the safety of UAV taxiing cannot be guaranteed without a robust local situational awareness system. A detailed safety study on operating UAVs in civil aerodromes can be found in [3]. Using the vertical take-off and landing type of UAVs may introduce less impact on the aerodrome daily operations. However, comparing with fixed-wing UAVs, these UAVs have limited payloads and flight endurance. This paper, therefore, focuses on fixed-wing UAVs.

To be compatible with larger and more powerful sensors, the size of civil fixed-wing UAVs needs to be larger as well. Hence, in order to minimise the dedicated infrastructure for UAVs' autonomous taxiing, we assume that UAVs observe real local environment information for taxiing purposes using the onboard cameras only. On the other hand, to provide safety guarantee for UAVs' autonomous taxiing, the situational awareness systems of the UAVs are expected to have the same visual sensing capabilities as a human pilot, such as obtaining information visually from signs and taxiway markings. Vision-based approaches are required to achieve this objective [8]. Recently, Durrie *et al.* [9, 10] have used machine vision for localisation for autonomous taxiing. They use a particle filter with a known aerodrome map to localise the aircraft.

While not much dedicated research has been conducted in the area of UAV autonomous taxiing, some of the autonomous roadway traffic and lane guidance techniques could be borrowed due to the similarity of their application environments. Both are structured environments which usually have a dark surface with surface markings of bright colours (usually white or yellow). With this in mind, luminosity and colour-based approaches are the natural choices. For instance, a colour-based approach is used for tracking unmarked road lanes in [11, 12]. However, colour-only based approaches are sensitive to light conditions. By utilising the 'dark-light-dark' pattern, a lane marking extraction solution is proposed in [13]. In this solution, the camera observations are converted into light intensity images and mapped into the ground plane for matching. With a Hough transform, the lane contours are detected in [14, 15].

Following the colour-based approach in the above studies, Lu *et al.* [6] have recently investigated the extraction of the taxiway centreline in an aerodrome environment. Rather than to solely rely on visual approaches, Lu *et al.* [6] incorporate two sources of information, i.e. an aerodrome map and camera images, to improve the robustness of situational awareness. More specifically, the centreline of the aerodrome under investigation in [6] is first extracted from camera images based on the given colour, and then it is further combined with the aerodrome map to produce an enhanced centreline detection output. Although the centreline extraction is improved by utilising the aerodrome map information in [6], the real interest of such a situational awareness system is to detect obstacles. To this end, [7] presents a self-learning framework
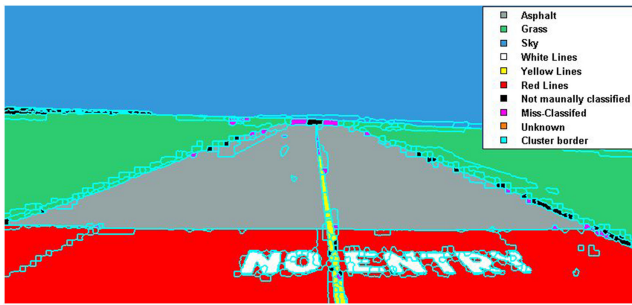
Fig. 1 *BN-semantic segmentation of an aerodrome taxiway where each cluster is classified into one of the few aerodrome classes*

a   *b*

Fig. 2 *Experiment platform*
(a) Jetstream, *(b)* Fire truck

that extends [6] in several aspects. First, Lu *et al.* [7] considered both the static taxiway features (e.g. the centreline of aerodrome taxiways) and moving obstacles. In addition, instead of using a colour-based approach, it adopts the frequency tuned saliency detection method (see, e.g. [16]) to improve the quality of image processing. More importantly, a recursive learning mechanism is introduced to process the information extracted from the camera images, where the detection of the centreline, obstacles and so on is based on not only the aerodrome map and current camera observation, but also on the image observations obtained in the previous time periods. This makes the detection results less sensitive to noise in the segmentation of individual images. The performance was validated in an indoor experiment in [7].

This paper further tests the self-learning framework proposed in [7] in a real outdoor aerodrome environment in the Walney Island Airport. Compared with the indoor experiment in [7], several new challenges arise when processing the real-world data, including camera vibration, extraction of the aerodrome's taxiway features and unknown moving obstacles, as detailed in Section 2.

To address various practical issues arising from outdoor aerodrome environments, we develop a new approach in this paper by integrating the Bayesian network (BN) semantic segmentation with the self-learning method in [7]. The former is originally proposed in [17] and it is used in this research to extract real-world features of the aerodrome (e.g. taxiways, grass, and the centreline) more reliably than that used in [7]. This is because in general certain types of objects are easier and can be more accurately detected due to the prominence of the features: the colour-based approaches can detect the centreline more accurately than the texture-based ones, whereas the texture based techniques are good at differentiating the surfaces types. By combining multiple techniques, the performance can be improved. We incorporate the method developed in [17] to fuse various (including colour, texture, luminance and spatial relationship based) features together with a Bayesian network; see an example shown in Fig. 1. With the BN-semantic segmentation technique, we also identify the reference horizon, upon which we can subsequently calibrate the obtained images and hence address the issue of camera vibration.

In practice, it is unrealistic to assume the availability of the visual attributes of moving obstacles beforehand; their appearances usually differ from one to another. We extend the work in [7] by the integration of the BN-semantic segmentation in [17] and the self-learning process in [7] to make use of knowledge acquired in the previous time periods through a recursive Bayesian learning process to improve the robustness of obstacle detection.

This paper is structured as follows. In the following section, we outline various research challenges for autonomous taxiing in real outdoor aerodrome environments. The integration of the BN-semantic segmentation and the self-learning process is investigated in Section 3. A detailed analysis of the outdoor data is undertaken in Section 4. Finally, we conclude this paper in Section 5.

2 Research challenges for autonomous taxiing in real outdoor aerodrome environments

Real-world outdoor aerodrome data was captured at Walney Island Airport by BAE Systems. A fire truck (Fig. 2*b*), with a monocular camera (GoPro), mounted to the dash alongside a commercial GPS/IMU module, was used to drive around the aerodrome taxiways to simulate UAV taxiing. This platform configuration was chosen according to the specification of the targeted UAV platform where a Jetstream aircraft (Fig. 2*a*) is used as a surrogate UAV for the development and test purposes.

Comparing with indoor environments, several new challenges need to be addressed to make the self-learning framework proposed in [7] work in an outdoor aerodrome environment:

- *Camera vibration:* The typical taxiing speed of aircraft is between 30 and 60 km/h, and the aircraft may need to accelerate and decelerate during the taxiing, follow signage, GTC commands or aerodrome traffic and so on. Therefore, the pitch angle of the camera is frequently changing, and the roll angle also changes when the vehicle turns. These angles will dramatically affect the accuracy of the inverse perspective mapped camera observations. This problem has a much smaller impact on indoor experiments due to slower speeds and smaller camera fields of view for indoor experiments.
- *Taxiway features in a real complex aerodrome environment:* In contrast to an indoor environment, more taxiway features can be detected and more interference factors exist in an outdoor environment. Since simple colour or saliency-based detection approaches are not able to detect all kinds of taxiway features robustly in such a complex environment, a better image process approach is required.
- *Unknown obstacle detection:* Taxiway features are static and follow a common set of standards, whereas moving obstacles could be anything. Consequently, the detection accuracy for static features can be improved with a supervised learning approach, whereas for the obstacle detection it is more reasonable to use approaches without such supervision.

3 Integration of the BN-semantic segmentation with the self-learning framework

Fig. 3 shows the overall structure of the system with the proposed integration, where the self-learning framework proposed in [7] is used as the foundation, and the BN-semantic segmentation is integrated to enhance situational awareness.

As it can be seen from Fig. 3, there are three major inputs for the system, i.e. taxiway map, GPS and camera observations. We outline the inputs/outputs and the major functions of each component shown in Fig. 3.

First, the taxiway map is used as prior knowledge in the system. A taxiway feature map and an obstacle map will be generated from it. The taxiway feature map will be used as an anchor point to find the spatial relationship between each of the camera observations during the taxiing. The obstacle map is normally initialised as an empty map unless some obstacle information is already known, and it will be used to keep the history of observed obstacles.

Second, a forward-facing camera is assumed to be the only observation source of this system. Unlike the pixel-based saliency method in [7], in this research, each original camera image is first passed into the BN-semantic segmentation module, from which two important outputs can be obtained: the horizon and semantic segmentation information. The BN-semantic segmentation method takes the HSV (Hue, Saturation, Value) of the input image, segments it, and classifies the segments based on common airfield objects. So instead of 'important pixels' used in [7], we have a pre-
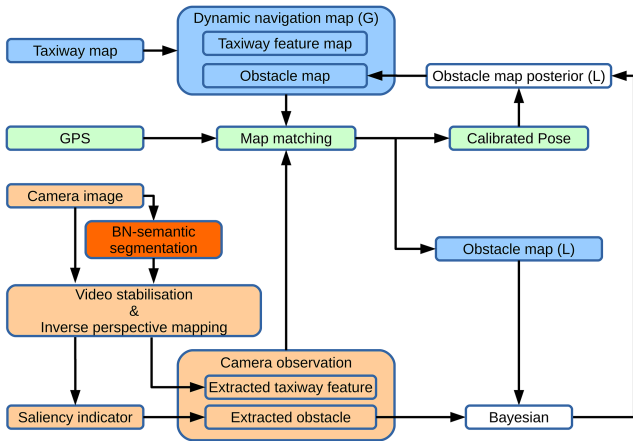
Fig. 3 *Self-learning framework integrated with the BN-semantic segmentation module*
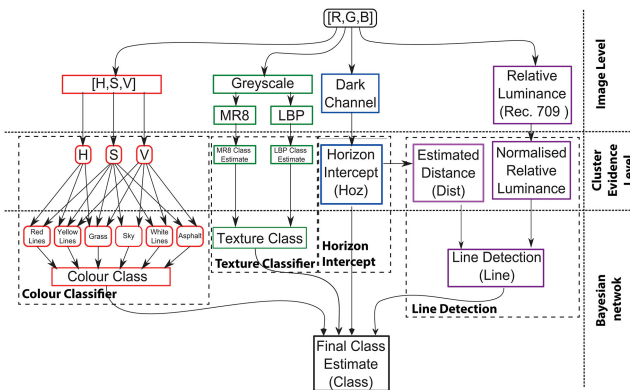

Fig. 4 *BN structure*

segmented image where each of those segments has been classified as airport objects. This has two big advantages: (i) the boundaries around objects will be much clearer with less noise, which in turn helps with the map matching phase; and (ii) the map will have known classes (e.g. building or taxiway) so image data can be associated with the map for the map matching phase with much greater ease.

The horizon information obtained from the BN-semantic segmentation is used to stabilise both the original and the segmented images, and then the inverse perspective mapping (IPM) is employed to transform them into the top down view; this ensures the transformed images are consistent with the taxiway map. Finally, the IPM output of the segmented images becomes the extracted taxiway features, while the IPM output of each original image will be further processed into the saliency indicator for obstacle detection.

Next, the GPS measurement provides an initial point for matching the extracted taxiway features to the prior distribution of the feature map. Since it is common to have multiple taxiways with very similar sizes and shapes (e.g. identical junction layouts) in an aerodrome, the initial GPS measurement is usually required to avoid the locational ambiguity. As the camera observation at each time step is already transformed into the top down view, the matching process can be conveniently done with the rigid point set registration. The successfully matching gives a calibrated pose of the vehicle and a local obstacle map as the counterpart of the obstacle observation.

Finally, in the Bayesian self-learning phase, the obstacle map (as a prior) and the extracted obstacles from the current image observation will be pooled together and processed, resulting in a posterior obstacle map. This posterior map at the current time point is regarded as a prior obstacle map for the Bayesian self-learning at the next time point.

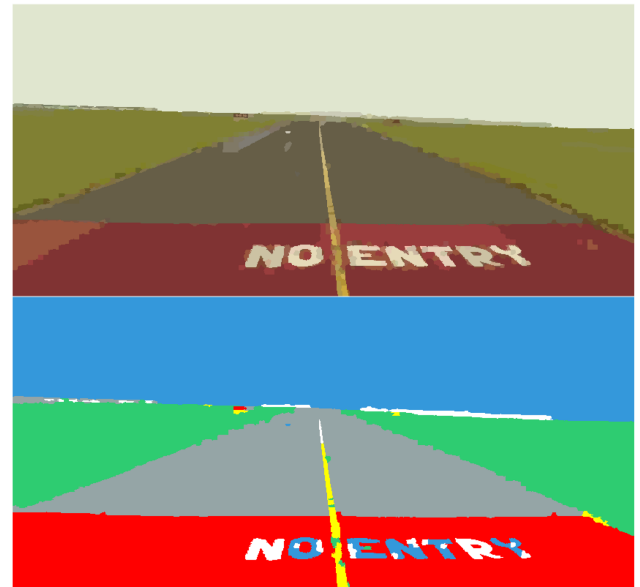This process is detailed below.

Fig. 5 *Discretised HSV colour image and subsequent BN colour classifier output*

3.1 Taxiway feature extraction

To perform UAV navigation, features need to be extracted from the images so that they can be matched to an aerodrome map. The most distinctive and robust features are the taxiway centreline, and the edge of the taxiway, de-marked by the transition between asphalt and grass. The method used to extract the centreline in [6] was a simple colour-based algorithm, and in [7] was the frequency tuned saliency indicator. In the indoor environment, these single-feature based methods worked very well, but at a real aerodrome, they would not be reliable enough, due to light and weather effects and worn or faded markings. Hence, a more robust feature extraction method is required. In this paper, we use BN-semantic segmentation to aid feature extraction. In the process of BN-semantic segmentation, each segmented cluster in an image is classified into a small number of classes, e.g. grass, asphalt, yellow centre line, and white lines.

The BN-semantic segmentation method is originally developed in [17] to combine colour, texture, luminance and contextual information probabilistically to improve classification performance, as shown in Fig. 4. Here we integrate this method with the self-learning framework for an aerodrome environment. During the autonomous taxiing, the captured image at each time step is first segmented into clusters; this is achieved using SLIC superpixels to get an initial fine segmentation, and then the density-based spatial clustering of applications with noise algorithm combines similarly coloured adjacent superpixels into clusters of pixels of the same class. The colour and texture are extracted and discretised from each cluster and used to give an initial estimate of the class of each cluster. An example is shown in Fig. 5, where first the HSV colour of each cluster is discritsed, and then the trained colour classifier subsection of the BN estimates the class of each cluster. Luminance is used to find the high reflectance surface markings. If the position of the cluster is known relative to the horizon, logic can be applied to better find close obstacles that appear on both sides of the horizon and differentiate between a ground object and sky classes. The advantage of pre-segmenting the image is that the boundary between classes will be more accurate and much better defined than it would be if a per-pixel classification was used. These smoother and more coherent boundaries will make map matching much easier.

At this stage, the horizon is also extracted as a reference to address the camera vibration issue, as detailed in the following section.

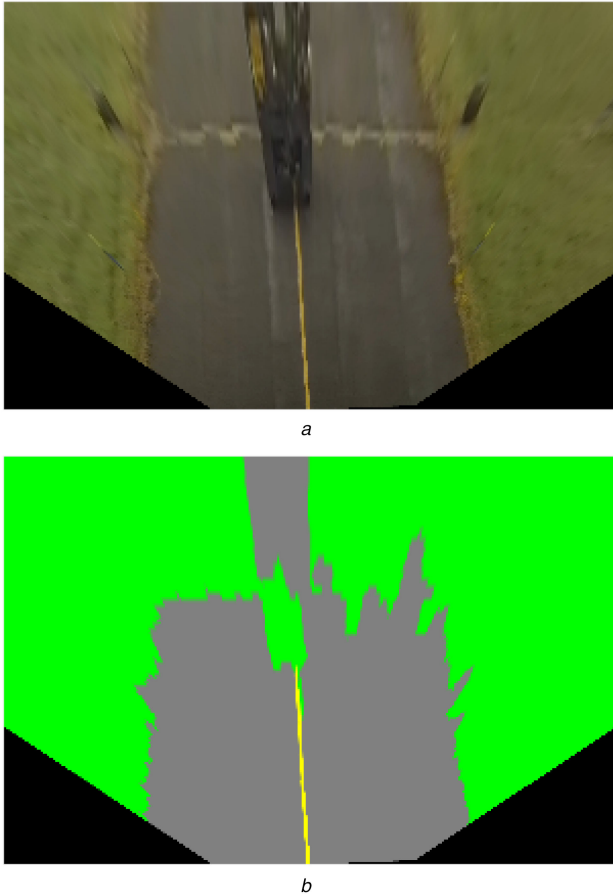**Fig. 6** *Stabilised video frame with the reference horizon*

a

b

Fig. 7 *Inverse perspective mapped camera view*
(a) Original, *(b)* Segmented

3.2 Horizon-based video stabilisation

As stated in the first challenge, camera vibration is not an issue in a highly controlled indoor experiment, where the vehicle moved in a relatively slow, constant speed on a smooth indoor floor surface. However, in the outdoor test, the vehicle's moving speed is much higher with frequent acceleration and deceleration on a relatively rough surface. In addition, the pitch (while accelerating/decelerating) and roll (while turning) angles will affect the accuracy of the IPM drastically, especially for distant objects.

Many video stabilisation approaches in the literature are based on local feature points (corners, edges etc.); they may be not suitable for large open environments. However, we also note that, as aerodromes are in large open-air locations, the horizon is almost always a feature that can be relied upon to give a pitch and roll reference. Fig. 6 shows an example of using horizon for video stabilisation: here the red dashed line is the reference horizon; this is achieved using dark channel detection, further detailed in [17].

From the detected horizon, the registration can be achieved with a rotation and a translation. In this example, the image is rotated counter-clockwise and translated downward so the horizon of the image matches to the reference horizon.

Fig. 7 gives a pair of typical IPM outputs from the original and segmented camera observations, in which green, grey, and yellow colours indicate the detected grass surface, asphalt surface and centreline, respectively. It can be observed that the centreline is the most robust feature, while the detected grass surface contains some false positives (FP) and asphalt surface has some false negatives (FN). The map matching in this paper relies on the extracted centreline, but the interface between the grass and asphalt could be matched with the taxiway map.

In summary, the BN-semantic segmentation and the image stabilisation are inter-connected to each other; they work jointly for the improvement of performance. During the stage of the BN-semantic segmentation, various features are extracted, including the taxiway centreline, taxiway boundaries, and horizon. The horizon is then used to address the image stabilisation issue. In addition, as the boundaries between objects will be clearer due to the BN-semantic segmentation, the map matching phase will be able to match the image to the map even though the camera will not be perfectly stable.

3.3 Obstacle extraction with Bayesian learning

The frequency tuned saliency detection method in [18] can be implemented with (1) and is applied to the original camera observation, as displayed in Fig. 7a:

$$\mathcal{S}(X) = \| \boldsymbol{I}_\mu - \boldsymbol{I}_G(X) \|, \tag{1}$$

where X denotes all the pixel locations in the image, $\boldsymbol{I}_\mu = [L_\mu, a_\mu, b_\mu]$ is the average colour vector of the observation image in the $L*a*b*$ colour space, and $\boldsymbol{I}_G = [L_G, a_G, b_G]$ is the blurred observation image with a Gaussian filter. The blurry process removes fine texture details and high spatial-frequency noise (see e.g. [7]). $\| \cdot \|$ is the L_2 norm (Euclidean distance).

For the autonomous taxiing system developed in this paper, the aerodrome layout is assumed to be known and a UAV only requires obstacle detection within the asphalt taxiway area. Therefore, the average colour vector $\boldsymbol{I}_\mu$ can be computed within the asphalt area. The advantage of this is that the contrast between asphalt and the obstacle will be more significant. We then apply the saliency detection with the two average colour vector definitions to the image in Fig. 7a. A comparison is given in Fig. 8 which shows that the saliency indicator of the asphalt area is much lower when $\boldsymbol{I}_\mu$ is defined with the asphalt mask, and the saliency indicators outside the asphalt area are ignored.

To make the self-learning process possible for moving obstacle detection purposes, map matching based on the BN-semantic segmentation plays a key role: the BN-semantic segmentation significantly enhances the accuracy of map matching. A pose $\boldsymbol{p}_{\text{opt}}$ can be obtained from the map matching via careful calibration.

On the basis of the map matching, the global obstacle map $q(\boldsymbol{M})$ is cut off to obtain a local map, denoted as $q(\boldsymbol{M}; \boldsymbol{p}_{\text{opt}})$, upon which the self-learning-based obstacle detection is undertaken. This is detailed below.

Specifically, each pixel in the obstacle map and obstacle observation is assumed to follow a Gaussian distribution, respectively,

$$\text{Obstacle map:} \quad q(\boldsymbol{M}; \boldsymbol{p}_{\text{opt}}) = \mathcal{N}(\boldsymbol{M}; \mathcal{M}, \sigma^2_{\text{map}}), \tag{2}$$

$$\text{Obstacle observation:} \quad q(\mathcal{S}|\boldsymbol{M}) = \mathcal{N}(\mathcal{S}; \boldsymbol{M}, \sigma^2_{\text{obs}}), \tag{3}$$

where $\boldsymbol{M}$ denotes the parameter matrix corresponding to the ground truth of the obstacle layout, $\mathcal{M}$ is the obstacle map, and $\mathcal{S}$ is the saliency indicator. σ^2_{map} and σ^2_{obs} are the variances specifying the noise levels of the map and observation.

Based on the pose $\boldsymbol{p}_{\text{opt}}$, the observed obstacle is now ready to be updated into the local obstacle map. Applying Bayes rule yields the following posterior distribution of the obstacle map $q(\boldsymbol{M}|\mathcal{S}; \boldsymbol{p}_{\text{opt}})$:

$$q(\boldsymbol{M}|\mathcal{S}; \boldsymbol{p}_{\text{opt}}) \propto q(\mathcal{S}|\boldsymbol{M})q(\boldsymbol{M}; \boldsymbol{p}_{\text{opt}}). \tag{4}$$
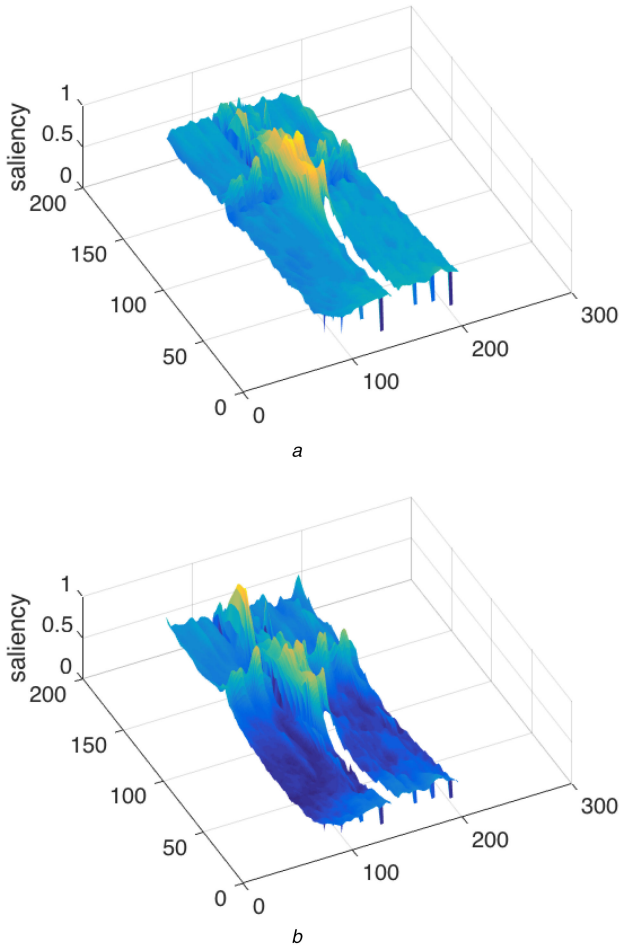
Fig. 8 *Saliency indicator computed from Fig. 7a*
(a) Original saliency, *(b)* With map information
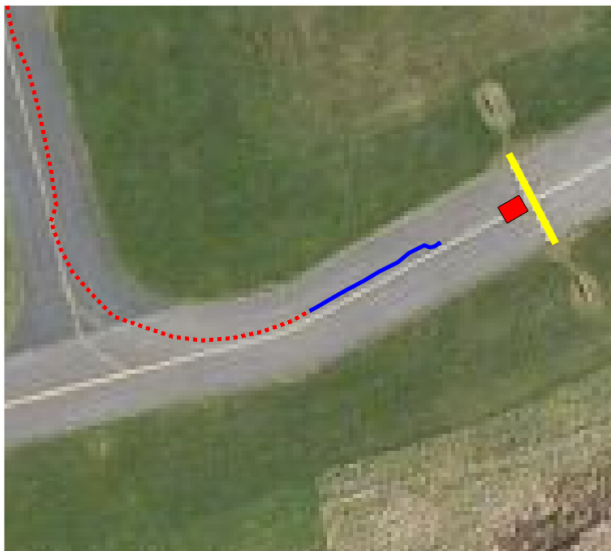
Fig. 9 *Aerodrome map*

It is straightforward to show that the posterior distribution $q(\boldsymbol{M} \,|\, \mathcal{S} ; \boldsymbol{p}_{\text{opt}})$ is a Gaussian distribution, $\mathcal{N}\left(\boldsymbol{M} ; \mathcal{M}_{\text{post}}, \sigma_{\text{post}}^2\right)$, with

$$\mathcal{M}_{\text{post}} = \frac{\sigma_{\text{obs}}^2}{\sigma_{\text{map}}^2 + \sigma_{\text{obs}}^2}\mathcal{M} + \frac{\sigma_{\text{map}}^2}{\sigma_{\text{map}}^2 + \sigma_{\text{obs}}^2}\mathcal{S}, \qquad (5)$$

$$\sigma_{\text{post}}^2 = \frac{\sigma_{\text{map}}^2 \sigma_{\text{obs}}^2}{\sigma_{\text{map}}^2 + \sigma_{\text{obs}}^2}. \qquad (6)$$

Then the local obstacle posterior $q(\boldsymbol{M} \,|\, \mathcal{S} ; \boldsymbol{p}_{\text{opt}})$ is updated back into the global obstacle map $q(\boldsymbol{M} \,|\, \mathcal{S})$. This updating process is recursively undertaken whenever a new obstacle observation is obtained. In addition, a forgetting factor $0 < \lambda < 1$ is introduced to inflate the variance of the obstacle map at each time step, so the new obstacle observation is not overwhelmed by the previous observations; see [7] for details.

4 Outdoor experiment

In this section, we provide a detailed analysis of the outdoor experiment.

The outdoor experiment consisted of two phases. Phase-1 was the data collection phase. In this phase, a satellite map of the Walney Island airport was obtained from Microsoft Bing Maps and was further processed into the taxiway feature map (Fig. 3). This feature map contained the taxiway centrelines, taxiway boundaries and stop signs. In order to simulate a taxiing UAV, we used a fire truck as a surrogate. The autonomous flight capable BAE Jetstream was used as the example aircraft in this experiment. The camera used was mounted on the fire truck at the height of the Jetstream's cockpit. The fire truck taxied around the aerodrome adhering at all times to standard taxiing procedures and rules, moving at a speed within the range of common aircraft taxiing speed. Three types of data were collected: video (from camera), attitude (from IMU) and positioning (from GPS) information.

Phase-2 was the analysis phase based on the collected data. By using the positioning and attitude information as an initial position, a gradient-based search was then applied to match the actual view (video) from the camera with the map features. By doing so, the vision and the taxiway map were aligned. With this alignment as a prerequisite, the self-learning process was then carried out and the dynamic navigation map was updated continuously based on the posterior distribution of the obstacle map.

During the outdoor experiment of testing the real-world taxiing process, the fire truck was driven along an aerodrome taxiway, as shown in Fig. 9. We focus on one particular scenario, where a yellow line marks the taxiway stop sign on the map, and an obstacle vehicle (marked with a red rectangle) was stopped in front of it. A curve in the figure shows the trajectory of the fire truck, where the solid blue line indicates when the obstacle vehicle appeared in the camera's view.

The performance of the self-learning improved saliency detection for various types of obstacles (large/small) in different light conditions (bright/dark) were compared against the original saliency detection in [7], which shows that the self-learning improved detection gives a robust and consistent result in the indoor environment. By integrating this self-learning framework with the BN-semantic segmentation and applying it to the outdoor aerodrome environment, our experiment shows that the system remains a performance as well as in the indoor environment.

To demonstrate the performance in the outdoor environment, we follow a similar routine as in the indoor environment by first giving an intuitive detection result in Fig. 10. Both of the results, obtained by using the original saliency and the self-learning improved saliency methods, respectively, are illustrated with the 'false colour' fused images, where the green colour indicates the camera observation and magenta marks the detected obstacles. We can see from the figures that the detection result with self-learning covers the true obstacle roughly the same as the original saliency detection, but with far fewer FP detections.

Next, we conduct a quantitative assessment for the obstacle detection analysis. We note that the obstacle vehicle appeared in the camera observation only when the fire truck was approaching the obstacle (as indicated by the solid blue section of the curve in Fig. 9). Hence, the following analysis focuses on this section of taxiing only.

In the literature, there are several commonly used measures for the assessment of pattern recognition results, including recall and false discovery rate (FDR).

Recall (also termed sensitivity) measures the proportion of positives that are correctly identified as such. FDR, on the other
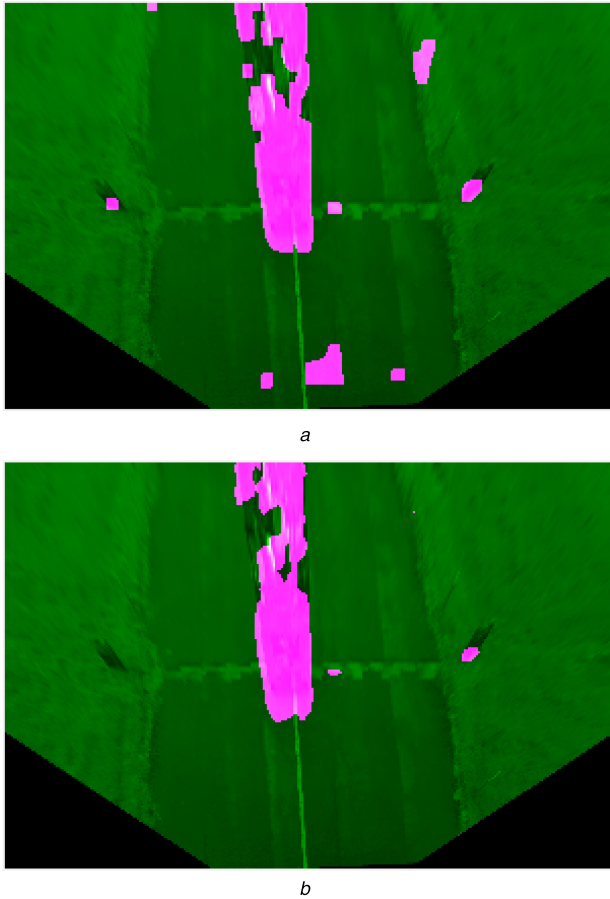
872

IET Intell. Transp. Syst., 2018, Vol. 12 Iss. 8, pp. 868-874

a

b

Fig. 10 *Image frame of the obstacle detection*
(a) Original saliency, *(b)* Self-learning improved saliency

hand, is the proportion of discoveries that are false among all. More precisely, we have

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}},$$

$$\text{FDR} = \frac{\text{FP}}{\text{FP} + \text{TP}},$$

where TP and FP are defined to be true positives and false positives, respectively, and FN is defined to be false negatives.

FDR is related to precision: precision $= 1 - $ FDR. Recall and FDR (and hence precision) measures are sometimes used together in the F_1 score to provide a single measurement

$$F_1 = \frac{2\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}.$$

To test the self-learning ability of the system in the outdoor environment, we compare the two methods, i.e. the original saliency method and self-learning improved saliency method, in terms of recall, FDR, and F_1 score. The obtained recall, FDR, and F_1 score in the outdoor experiment are displayed in Figs. 11–13.

The video data was captured at 30 frames per second. To calculate the assessment measures (recall, FDR, and F_1 score), we evenly chose 3 frames per second to manually mark the ground truth, as displayed in Figs. 11–13, where the tick unit of the *x*-axis was 1/3 of a second. This covers a detection period of about 15 s and corresponds to the taxiing trajectory section indicated the solid blue curve in Fig. 9.

From Fig. 11, it can be seen that the two methods maintain a similar level of recall rates. This observation also applies to the F_1 scores displayed in Fig. 12. Comparing Fig. 10*a* with Fig. 10*b*, this is not surprising: a majority of the TP were captured by both methods.
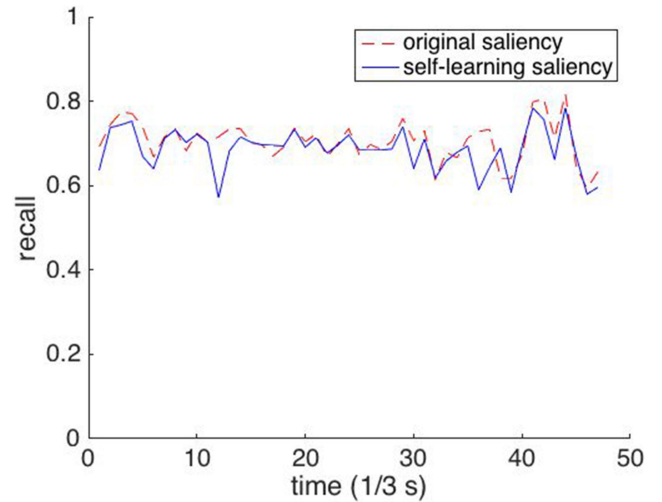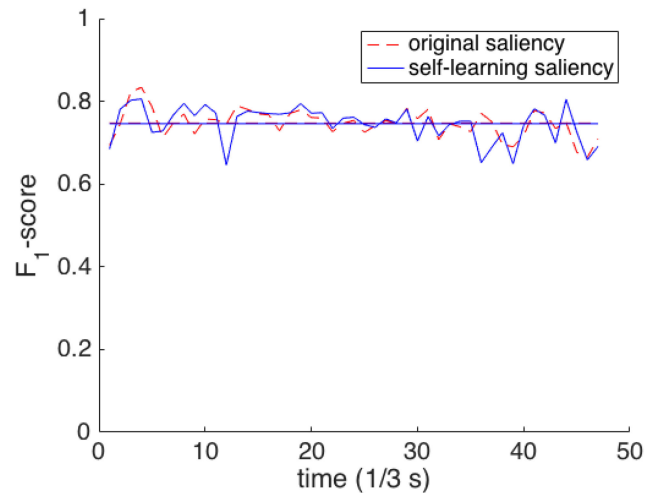

Fig. 11 *Recall*

Fig. 12 *F_1 score*

From Fig. 13, on the other hand, we can see a very large performance difference. To highlight this difference, we mark the mean value of the FDR for the original saliency detection (8.35%) and the mean value of the FDR for the self-learning saliency method (3.07%) with two horizontal lines in Fig. 13, respectively. It can be seen from Fig. 13 that the FDR for the 44th data point is particularly high for the original saliency method. From the corresponding image frame for the obstacle detection displayed in Fig. 10, we can see clearly that the original saliency detection method had a much higher FP detection level, resulting in a higher FDR. Overall, with the self-learning method, the FDR was reduced by 63.23% in this detection period.

The above analysis shows that, by controlling for the recall rate level, the self-learning method can substantially reduce FDR.

In summary, with the integrated method that combines the self-learning method with the BN-semantic segmentation, the outdoor experiment has demonstrated its improved robustness: the camera vibration problem in the outdoor aerodrome is overcome and aerodrome features can be extracted more accurately. Furthermore, based on the more accurately extracted features, images can be precisely matched with the maps. Finally, knowledge about obstacles can be self-learned and obstacles can be detected in a robust manner.

5 Conclusions

This paper investigates autonomous taxiing of UAVs in a real outdoor aerodrome environment. The self-learning framework developed in [7] was tested only in an indoor laboratory. When applying [7] in an outdoor aerodrome environment, various research challenges arose. To address these practical issues, we
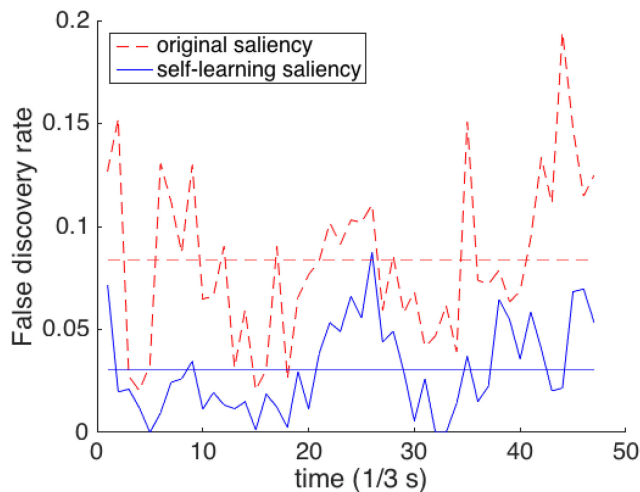
Fig. 13 *False discovery rate*

have integrated a BN-semantic segmentation image processing technique with the self-learning framework in [7] to enhance situational awareness in autonomous taxiing. Through testing against the real aerodrome environment, we have demonstrated that the integrated approach in this paper can overcome the camera vibration problem, and better extract taxiway features of an aerodrome. The enhanced self-learning framework also improves the robustness of the obstacle detection by taking into account the obstacle observations acquired in the previous time periods.

In the current integration structure, the self-learning framework takes the output of the BN-semantic segmentation module, but the enhanced result from the self-learning framework is not fed back to the BN-semantic segmentation module. Adding an interaction element between them may lead to a more robust result; this is a potential future work to be explored in our future research.

6 Acknowledgment

7 References

[1] Hausamann, D., Zirnig, W., Schreier, G., *et al.*: 'Monitoring of gas pipelines – a civil UAV application', *Aircr. Eng. Aerosp. Technol.*, 2005, **77**, (5), pp. 352–360

[2] Li, Z., Liu, Y., Hayward, R., *et al.*: 'Knowledge-based power line detection for uav surveillance and inspection systems'. 2008 23rd Int. Conf. Image and Vision Computing, Christchurch, New Zealand, 2008, pp. 1–6

[3] Loh, B.Y.R., Roe, T.: 'UAVs in civil airspace: safety requirements', *IEEE Aerosp. Electron. Syst. Mag.*, 2009, **24**, (1), pp. 5–17

[4] Dalamagkidis, V.K.P. K., Piegl, L.A.: 'Current status and future perspectives for unmanned aircraft system operations in the us', *J. Intell. Robotic Syst.*, 2008, **52**, (2), pp. 313–329

[5] Loegering, G., Harris, S.: 'Landing dispersion results - global hawk auto-land system'. AIAA's 1st Technical Conf. and Workshop on Unmanned Aerial Vehicles, Portsmouth, VA, USA, 2002

[6] Lu, B., Li, B., Chen, W.H.: 'Map-enhanced visual taxiway extraction for autonomous taxiing of uavs', *IFAC-PapersOnLine*, 2015, **48**, (9), pp. 49–54

[7] Lu, B., Coombes, M., Li, B., *et al.*: 'Improved situation awareness for autonomous taxiing through selflearning', *IEEE Trans. Intell. Transp. Syst.*, 2016, **17**, pp. 3553–3564

[8] Eaton, W., Chen, W.H.: 'Image segmentation for automated taxiing of unmanned aircraft'. 2015 Int. Conf. on Unmanned Aircraft Systems (ICUAS), Denver, CO, USA, 2015, pp. 1–8

[9] Durrie, T.G.E.W.F.J., Pledgie, S.: 'Vision-aided inertial navigation on an uncertain map using a particle filter'. Proc. of the IEEE Int. Conf. on Robotics and Automation, 2009, pp. 4189–4194

[10] Durrie, J., Gerritsen, T., Frew, E.W., *et al.*: 'Vision aided inertial navigation on an uncertain map using a particle filter'. IEEE Int. Conf. on Robotics and Automation, 2009, ICRA 09, Kobe, Japan, 2009, pp. 4189–4194

[11] Sotelo, M.A., Rodriguez, F.J., Magdalena, L., *et al.*: 'A color vision-based lane tracking system for autonomous driving on unmarked roads', *Auton. Robots*, 2004, **16**, (1), pp. 95–116

[12] Cheng, H.Y., Jeng, B.S., Tseng, P.T., *et al.*: 'Lane detection with moving vehicles in the traffic scenes', *IEEE Trans. Intell. Transp. Syst.*, 2006, **7**, (4), pp. 571–582

[13] Satzoda, R.K., Trivedi, M.M.: 'Drive analysis using vehicle dynamics and vision-based lane semantics', *IEEE Trans. Intell. Transp. Syst.*, 2015, **16**, (1), pp. 9–18

[14] Gaikwad, V., Lokhande, S.: 'Lane departure identification for advanced driver assistance', *IEEE Trans. Intell. Transp. Syst.*, 2015, **16**, (2), pp. 910–918

[15] Zhang, Y., Gong, J., Tian, J.: 'Robust lane detection and tracking using improved Hough transform and Gaussian mixture model', in: Zhang, T., Sang, N., (Eds): '*Seventh international symposium on multispectral image processing and pattern recognition (MIPPR)*', vol. 8003 (International Society for Optics and Photonics, Guilin, China, 2011), pp. 1–8

[16] Achanta, R., Hemami, S., Estrada, F., *et al.*: 'Frequency-tuned salient region detection'. IEEE Conf. on Computer Vision and Pattern Recognition, 2009, CVPR 2009, Miami, FL, USA, 2009, pp. 1597–1604

[17] Coombes, M., Eaton, W., Chen, W.H.: 'Unmanned ground operations using semantic image segmentation through a Bayesian network'. Int. Conf. on Unmanned Aircraft Systems (ICUAS), 2016, pp. 868–877

[18] Wang, H., Wei, Z., Wang, S., *et al.*: 'A vision-based obstacle detection system for unmanned surface vehicle'. IEEE 5th Int. Conf. on Robotics, Automation and Mechatronics (RAM), Qingdao, China, 2011, pp. 364–369

874

IET Intell. Transp. Syst., 2018, Vol. 12 Iss. 8, pp. 868-874