Map-enhanced visual taxiway extraction for autonomous taxiing of UAVs

This item was submitted to Loughborough University's Institutional Repository by the/an author.

Citation: LU, B., LI, B. and CHEN, W.-H., 2015. Map-enhanced visual taxiway extraction for autonomous taxiing of UAVs. IFAC Papers online, 48(9), pp. 49–54.

Additional Information:

- This paper was accepted for publication in the journal IFAC Papers online and the definitive published version is available at http://dx.doi.org/10.1016/j.ifacol.2015.08.058. It was presented at the 2015 IFAC Workshop on Advanced Control and Navigation for Autonomous Aerospace Vehicles, 10th-12th June 2015, Seville, Spain.

Metadata Record: https://dspace.lboro.ac.uk/2134/18631

Version: Accepted for publication

Publisher: Elsevier / © IFAC.

Rights: This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: https://creativecommons.org/licenses/by-nc-nd/4.0/

Please cite the published version.
Map-Enhanced Visual Taxiway Extraction for Autonomous Taxiing of UAVs

Bowen Lu ∗ Baibing Li ∗∗ Wen-Hua Chen ∗

∗ Department of Aeronautical and Automotive Engineering, Loughborough University, UK (e-mail: {B.Lu, W.Chen}@lboro.ac.uk).
∗∗ School of Business and Economies, Loughborough University, UK (e-mail: B.Li2@lboro.ac.uk).

Abstract: In this paper, a map-enhanced method is proposed for vision-based taxiway centreline extraction, which is a prerequisite of autonomous visual navigation systems for unmanned aerial vehicles. Comparing with other sensors, cameras are able to provide richer information. Consequently, vision based navigations have been intensively studied in the recent two decades and computer vision techniques are shown to be capable of dealing with various problems in applications. However, there are significant drawbacks associated with these computer vision techniques that the accuracy and robustness may not meet the required standard in some application scenarios. In this paper, a taxiway map is incorporated into the analysis as prior knowledge to improve on the vehicle localisation and vision based centreline extraction. We develop a map updating algorithm so that the traditional map is able to adapt to the dynamic environment via Bayesian learning. The developed method is illustrated using a simulation study.

Keywords: knowledge-based systems, map-prior, taxiway centreline extraction.

1. INTRODUCTION

Unmanned Aerial Vehicle (UAV) was initially investigated for military applications to reduce the risk of losing precious human lives, and the cost of such a UAV was extremely high. Thanks to the rapid technological development in the recent years in a number of areas such as sensor, battery, engine, etc., the capability of a low cost UAV has been substantially increased and consequently UAVs can now be used for many civil applications. While a military UAV deserves to have a dedicated aerodrome due to its particularity, a civil UAV is more likely to share the infrastructure with existing civil aerodromes to reduce additional cost so that the technology is more affordable to public. Manned and unmanned hybrid civil aerodromes are a highly complex and dynamic environment; a UAV operating in it needs to be both efficient and robust without additional infrastructures. Autonomous taxiing in such an environment becomes a new challenging research area.

Taxiway localisation is a fundamental problem of UAV autonomous ground navigation in aerodromes. Based on Differential Global Positioning System (D-GPS), Cho et al. (2007) provide a solution for automatically taxiing, takeoff and landing. However, D-GPS is not available on every aerodrome, which limits the application of this method. In addition, it is unrealistic to assume that a UAV autonomously operates in such a dynamic environment (civil aerodrome) with a single localisation system. In order to increase the system robustness and to deal with unexpected changes in the environment, other sensing approaches should be also considered to deliver an integral system.

Various autonomous navigation solutions have been proposed with integrating different sensors. Sun et al. (2013) use an Inertial Measurement Unit (IMU) to enhance the reliability of a GPS-based navigation system. On the other hand, in a GPS denied area, a multi-sensor fusion technique is developed in Ilyas (2013) for a local navigation. In addition, a continuously rotating 2D laser scanner is used to build an allocentric 3D map of a rough terrain environment in Schadler et al. (2014). For the same environment, a 2.5D egocentric map is built with eight RGB-D sensors in Schwarz and Behnke (2014), and drivable path is assessed based on it. Although different combinations of local sensors can be used for navigation, not all of them are suitable for a flat and open environment, e.g. an aerodrome. Due to the lack of obstacle features in such an environment, range sensors (e.g. laser and RGB-D based) are not appropriate for this application.

For the navigation process on a manned aircraft, most of the local guiding information are obtained visually, including taxiway centreline, signs, pavement markings, etc. Hence a visual based information extraction turns out to be a reasonable choice. From a perspective of image processing, taxiway centreline extraction has no difference from road lane marker detection. A commonly used approach of which is to extract bright (yellow or white painted) regions from a dark background (asphalt pavement). Many research results have been reported in this area; we summarise some of the most recent/relevant...
studies below. Revilloud et al. (2013) proposes a multi-lane detection and estimation approach in which the road markers are obtained by setting a threshold on the pixel intensity. With the same intensity based extraction approach, an adaptive threshold is used in Kheyrollahi and Breckon (2010), where the extracted pixels are grouped into isolated shapes and features of road markers are further recognised. With a set of learned road marker templates, a road marker recognition approach is developed in Wu and Ranganathan (2012). As reported in Sebsadjji et al. (2010), stereo images can also be used to improve on the extraction accuracy. The most representative visual navigation for autonomous vehicle is from the VisLab, University of Parma in Italy. As reported in Broggi et al. (2012), their autonomous vehicles successfully drove across the Eurasian continent from Parma, Italy to Shanghai, China.

Although it is feasible that autonomous navigation is purely based on vision, additional information can be used to improve on its reliability. Semantic information from digital map is one of the many choices. Based on the belief that a vehicle is more likely to locate at accessible places (i.e. a road or a car park), various probability zones are assigned to map segments and used to adjust the motion model of vehicle in Oh et al. (2004). Instead of defining probability zones, Jabbour et al. (2008) use a hierarchical Bayesian framework to assess the probability of multiple localisation hypotheses. Specifically, digital map information is used as a geographical measurement (similar to other canonical sensor measurements) to improve on the GPS positioning accuracy. The Oxford mobile robotics group proposes several prior-based approaches for improving on visual localisation for road autonomous navigation. Particularly, Napier et al. (2010) use a coarse-to-fine matching scheme together with an image alignment approach to find the bias between vehicle vision and a birds-eye image. Following the work of Napier et al. (2010), Napier and Newman (2012) employ a mutual information based matching scheme in their study. Instead of using imagery digital map, an obstacle point cloud is collected as prior information and Kullback-Leibler divergence is used to compare between the prior point cloud and real-time laser scan for acquiring better localisation accuracy in Baldwin and Newman (2012).

Comparing with the previous studies mentioned above, this paper aims not only to get a better localisation by matching the visual observation with the map, but also to continuously update the map with visual observation. Specifically, GPS measurements and taxiway map are synthesised to produce a group of prior centreline distribution candidates, and an observed centreline distribution is generated from the image processing techniques. Then an error assessment is carried out with Kullback-Leibler Divergence (KLD) to correct the GPS measurement error, and improve on the centreline extraction accuracy. With the corrected locational information, this enhanced centreline extraction is then used to replace the corresponding area in the taxiway map so that the map is adapted to the environment.

The paper organises as follows. Section 2 highlights various research challenges in this area, and proposes a framework to deal with each of these research issues. Section 3 formulates and investigates the taxiway centreline extraction problem with observation and map. In Section 4 we develop a probability based map-observation matching method, and in Section 5 we discuss how to make the map adapt to a dynamic environment. Finally, this paper concludes in Section 6.

2. CHALLENGES AND FRAMEWORK

2.1 Research challenges

As stated in Section 1, the performance of image processing based taxiway centreline extraction highly depends on the quality of captured images (e.g. noise level, resolution). This section discusses various research challenges of using GPS and taxiway map to enhance the visual extraction.

The research in this paper includes two key elements, i.e. observation and taxiway map. An observation is defined to be a camera-captured image, and a taxiway map is a binary image shows taxiway centrelines. In order to exploit the map to enhance the taxiway centreline extraction, several research issues need to be addressed:

(1) Representation of map and observation: In order to pool and combine the information from the map and observation, the information need to be represented in the same format;

(2) Locational matching: Because of the noisy GPS measurements, it is not a trivial task to seek a consistent local map that matches to the observation;

(3) Updating the map: A map could be out-of-date or incomplete. Therefore, the research framework to be developed in this paper needs to be capable of updating the map on-line to adapt to the dynamic environment.

We will address each of these research issues in Sections 2-5. We first of all outline an overall framework structure in Section 2.2.

2.2 Framework structure

The overall structure of the proposed framework is shown in Fig. 1. At the top of the framework, there are two inputs: measured vehicle state and camera observation. A single output is the enhanced centreline extraction.

With respect to the first research challenge, a Region of Interest (ROI) from the camera observation is projected into an orthographic view from the top, so the observation is now having a consistent view with the 2D taxiway map. Then the vehicle state is used to crop a local region from the map that matches to the ROI in camera observation. A dashed line separates the system into two levels (image level and probability level), and the enhancement of centreline extraction is undertaken at the probability level. Detailed discussion about the probability representations of observation and map, namely observation distribution and map-prior distribution, are given in Section 3.

Then a matching process is done by finding the minimised Kullback-Leibler divergence between them, so the second challenge can be solved (KLD matching). Then the optimal map-prior is used to enhance the observation
Fig. 1. Framework structure of the developed method distribution. Details of the matching process is presented in Section 4.

After a posterior distribution is produced from the optimal map-prior and observation distribution, the enhanced centreline extraction is obtained by setting a threshold to it. A dashed link updates the taxiway map with the enhanced extraction. The centreline extraction results with and without this updating mechanism are compared in Section 5, which shows the adaptivity of the proposed framework.

3. MAP-PRIOR GUIDED EXTRACTION

3.1 Representation of the observations and map

This section provides probability models of observation and map and defines the enhanced observation.

Let \( \mathbf{p} = [x, y, \theta]^T \) denote the state of aerial vehicle (position and heading) on a taxiway. Consider the following model for a sensor measurement \( \tilde{\mathbf{p}} \) on \( \mathbf{p} \):

\[
\tilde{\mathbf{p}} = f(\mathbf{p}) + \mathbf{e},
\]

where \( f(\cdot) \) is a measurement function, and \( \mathbf{e} = [\varepsilon_x, \varepsilon_y, \varepsilon_\theta]^T \sim \mathcal{N}(\mathbf{0}, \Sigma) \) is a zero mean Gaussian measurement noise of GPS. Thus, the measurement is modelled with \( \tilde{\mathbf{p}} \sim \mathcal{N}(f(\mathbf{p}), \Sigma) \).

**Observation distribution** Although the actual state \( \mathbf{p} \) is unknown in practical, the camera observation corresponding to \( \mathbf{p} \) can be obtained directly.

The first step of generating a distribution of taxiway centreline from camera observation is Inverse Perspective Mapping (IPM). This mapping is required to have a consistent perspective as the map, and can be either carried out with a perspective transform matrix, or based on vanish point detection Kheyrollah and Breckon (2010).

Then the mapped observation is further converted into a colour indicator. Specifically, by knowing the taxiway centreline painting colour (yellow), we measure the Euclidean distance (L2-norm) between this known centreline colour and the colour observed at each pixel in RGB space\(^1\) to generate an indicator at each pixel that characterises how close the colour of the observed image is to the pre-specified colour:

\[
I = \frac{1}{\sqrt{2\pi}\tau} \exp \left( -\frac{d^2}{2\tau^2} \right),
\]

where the bandwidth parameter \( \tau \) in this model controls the sensitivity of colour difference, and

\[
d = \|C_{\text{pixel}} - C_{\text{taxiway}}\|_2,
\]

where \( C_{\text{pixel}} = [R', G', B']^T \) is a colour vector of each pixel in the observed image, and \( C_{\text{taxiway}} = [R, G, B]^T \) is the given taxiway centreline colour.

We assume that the colour indicator \( I \) follows a Gaussian distribution below:

\[
q(I|m) = \mathcal{N}(I; m, \sigma^2_I),
\]

where \( \sigma^2_I \) is the variance parameter used to characterise the noise level of the image. \( m \) is the colour indicator corresponding to the ground truth of taxiway centreline in the environment.

**Map-prior distribution** By knowing the aerial vehicle state \( \mathbf{p} \) and camera intrinsic and extrinsic parameters, one can find a corresponding area on taxiway map that matches with the observation distribution derived in previous section. Due to the noises in GPS measurements, the noise model in Eq. (1) needs to be considered. Therefore, instead of using actual state \( \mathbf{p} \), the measured state \( \tilde{\mathbf{p}} \) should be used to generate the map-prior distribution.

To be clear, the map in this paper is a binary image that contains white labelled taxiway centrelines with black labelled background, and a *map-prior* is a distribution generated from the map with the following two steps. The first step is to crop a local map from the global map \( \mathcal{M} \), followed by a convolution of the local map and a Gaussian kernel (termed Gaussian blur in the literature). This generates the information on the local map \( \tilde{\mathcal{M}} \). The prior distribution of the ground-truth map is specified as follows:

\[
q(m) = \mathcal{N}(m; \bar{m}, \sigma^2_m),
\]

where \( \sigma^2_m \) is the variance parameter used to characterise the noise level of the map prior.

**Enhanced distribution** Given the prior distribution (5) and the observation distribution (4) at each pixel, the enhanced point-wise belief can be obtained by Bayesian learning via the following Bayes’ rule, as given in the following posterior distribution:

\[
q(m|I) \propto q(I|m)q(m) = \mathcal{N}(I; m, \sigma^2_I)\mathcal{N}(m; \bar{m}, \sigma^2_m).
\]

The ratio between \( \sigma^2_I \) and \( \sigma^2_m \) controls the adaptivity of the map. An image output \( \tilde{\mathcal{M}} \) (extracted centreline) from the enhanced distribution can be obtained by setting a threshold to the above posterior distribution.

\(^1\) Other colour spaces could also be used, such as HSL in Sotelo et al. (2004), CIELab in Ma and Xie (2010), etc. Since this research is not focusing on sophisticated image processing techniques, here we use RGB space for simplicity.
Fig. 2. A taxiway of Stansted Airport, UK

Fig. 3. Colour indicator of observation

Fig. 4. Taxiway map

3.2 An example of map-prior enhanced extraction

A photograph taken at Stansted Airport, UK, is used as the camera observation in this example. Fig. 2a shows the original image, and a trapezoid in which marks the ROI. The actual size of ROI (Fig. 2b) is measured from satellite map as 17.5m (width) x 26m (depth).

By setting $C_{\text{taxiway}} = [171, 147, 75]^T$ and applying Eqs. (3) and (2) to the data displayed in Fig. 2b, the colour indicator of observation is obtained, as shown in Fig. 3.

To work out a map-prior, a map of the taxiway is used as the global map, as presented in Fig. 4a. With given measurements of position and heading, a cropped local view is obtained, as displayed in Fig. 4b. Then a Gaussian blur is applied to incorporate noise information into the local map. Fig. 5 shows the generated colour indicator of map.

Following Eq. (6), the enhanced distribution (posterior) is a product of observation distribution and map-prior distribution. Fig. 6a shows the result obtained from the enhanced distribution where and the taxiway can be extracted conveniently by setting an appropriate threshold. Fig. 6b gives the extracted centreline from Fig. 6a. For comparison, Fig. 6c shows the extraction result from the original observation distribution in Fig. 3. Different choices for the threshold were explored in the extractions for the best performance. It can be observed that the enhanced extraction shows a clearer centreline with less noise.

4. KULLBACK-LEIBLER DIVERGENCE BASED OBSERVATION-PRIOR MATCHING

Section 3 discussed an approach to combining the observation and map to produce a better extraction result, and an instance was given to demonstrate its performance. Next, we turn to consider the second challenge in Section 2.1, as the enhancement performance also depends on whether the observation and map-prior are locationally matched.

Although there are various methods that can be used for matching images in image processing field, many of them have a high computational cost. Therefore, instead of employing an additional image processing based algorithm, we investigate the feasibility of directly matching the observation and map-prior distributions in this section. A commonly used approach for measuring the similarity of two distributions is that of Kullback-Leibler Divergence (KLD). More specifically, let $q(m; \hat{p}_{\text{test}}, X)$ be the map-prior distribution at each pixel $X$ with testing state $\hat{p}_{\text{test}} = \hat{p} + [dx, dy, d\theta]^T$, and $q(l|m; p, X)$ be the observation distribution at each pixel $X$ with state $p$. We solve:

$$\hat{p}_{\text{opt}} = \arg\min_{\hat{p}_{\text{test}}} \sum_{X \in I} D_{\text{KL}}(q(m; \hat{p}_{\text{test}}, X)||q(l|m; p, X)), \quad (7)$$

where $I$ denotes the observed area and $\hat{p}_{\text{opt}}$ is the best matched state.

In order to demonstrate the changing nature of KLD measurements with respect to space, a mesh plot of KLD is given in Fig. 7, where $dz$ and $dy$ (in metres) show the testing area around the measured state $\hat{p}$ ($d\theta$ is not showing in this figure). Practically, the gradient method can be used to find $\hat{p}_{\text{opt}}$. However, the local optimal issue needs to be carefully handled. With a known GPS accuracy specification, a grid based KLD check within the GPS error bounded area (e.g. Fig. 7) can be used to provide a good initial point to avoid local minima in the gradient descending algorithms.

\[^{2}\text{http://www.geograph.org.uk/photo/2481393}\]

\[^{3}\text{The heading (bearing) follows the measurement method in land navigation: using north as 0° reference, and increasing to 359.9° in clockwise direction.}\]
Given an observation indicator in Fig. 8a and a map indicator from a noisy position measurement in Fig. 8b, the calibrated vehicle state $\hat{p}_{opt}$ is obtained via Eq. (7) and a corrected map indicator is shown in Fig. 8c.

To demonstrate the matching performance in a moving scenario, a frame sequence is created from Fig. 2b, and GPS measurements are assigned to these frames accordingly. The size of each observation is now reduced to 17.5m (width) × 16m (depth). By setting that the vehicle moves straightforwardly with a constant speed of 10m/s, the frame sequence is generated with a 0.1s interval. Fig. 9 shows half of the sequence with an interval of 0.2s.

With the proposed map-observation matching, a centreline extraction result of the frame sequence is shown in Fig. 10. In this simulation study, the extracted centreline at each time step only relies on current observation and map-prior. It can be seen that the extracted centreline is clear and consistent through the whole frame sequence.

To visually assess the performance of the proposed KLD matching in this moving scenario, a trajectory comparison is given in Fig. 11. The figure uses Universal Transverse Mercator (UTM) coordinate system which maps longitude and latitude into a two dimensional Cartesian coordinate system and uses meter as its unit. In addition, an offset vector ($x$: 311100 meters, $y$: 5752700 meters) is subtracted from the UTM coordinates so that the performance result is easier to assess from the figure. An asterisk labelled line is a reference trajectory that used to generate the frame sequence in Fig. 9, and a circle on it marks the initial point (time = 0s). By adding a zero mean Gaussian noise with covariance $\Sigma = \text{diag}(0.5, 0.5, 0)$ onto the reference positions, the simulated GPS measurements are obtained in triangles. Squares are corrected states after the KLD matching. It can be clearly seen that corrected states are much closer to the reference trajectory (ground truth).

5. ADAPTIVE MAP

Considering that a map could be out-of-date and dynamic obstacles (e.g. logistic vehicles or pedestrians) cannot possibly be included in the map in advance, a real-time updating strategy is worth investigating.

The issues of map-observation matching and enhancing framework were investigated in Section 3 and 4. Given a camera observation and the global map at time $t$, the best matched local map is found by the KLD matching, and an enhanced extraction is obtained by combining the observation and the best matched local map with Bayesian learning. In this paper, the global map at time $t + 1$ is generated by substituting the best matched local map with the enhanced extraction. The same frame sequence in Fig. 9 is used for testing the performance. Fig. 12 shows the...
enhanced extraction at different time steps. Comparing with the results with static map in Fig. 10, the pavement marking on centreline is enhanced and gradually become solid. Figs. 13a and 13b give the static and adaptive global map at time = 1s, and the rectangles show the observation-covered area in the one second. It can be seen from the results that more detailed features are extracted and updated into the global map, and hence the KLD matching and extraction enhancing performance in next step are improved.

Fig. 12. Adaptive-map-guided centreline extraction

Fig. 13. Static (a) and adaptive (b) maps

6. CONCLUSIONS

By representing camera observation and taxiway map with probability distributions, this paper proposes a framework for enhancing the performance of vision based taxiway centreline extraction. Without involving additional image feature extraction techniques, an image matching process is performed with Kullback-Leibler divergence, and the localisation difference from the matching is used as a reference for correcting the error of GPS measurement. In addition, an updating algorithm is developed to ensure the map up-to-date. The future work of this research will be testing the algorithm on a practical platform for map updating and obstacle detection.

REFERENCES

Baldwin, I. and Newman, P. (2012). Road vehicle localization with 2D push-broom LIDAR and 3D priors. 2012 IEEE International Conference on Robotics and Automation, 2611–2617.

Broggi, A., Cerri, P., Felisa, M., Laghi, M.C., Mazzei, L., and Porta, P.P. (2012). The VisLab Intercontinental Autonomous Challenge: an extensive test for a platoon of intelligent vehicles. International Journal of Vehicle Autonomous Systems, 10(3), 147–164.

Cho, A., Kim, J., Lee, S., Choi, S., Lee, B., Kim, B., Park, N., Kim, D., and Kee, C. (2007). Fully automatic taxing, takeoff and landing of a UAV using a single-antenna GPS receiver only. In 2007 International Conference on Control, Automation and Systems, 821–825.

Ilyas, M. (2013). Dependable navigation in GPS denied environment: A multi-sensor fusion technique. In IEEE ISR 2013, 1–6.

Jabbour, M., Bonnifait, P., and Cherfaoui, V. (2008). Map-Matching Integrity Using Multihypothesis Road-Tracking. Journal of Intelligent Transportation Systems, 12(4), 189–201.

Kheyrollahi, A. and Breckon, T.P. (2010). Automatic real-time road marking recognition using a feature driven approach. Machine Vision and Applications, 23(1), 123–133.

Ma, C. and Xie, M. (2010). A Method for Lane Detection Based on Color Clustering. 2010 Third International Conference on Knowledge Discovery and Data Mining, 8(1), 200–203.

Napier, A. and Newman, P. (2012). Generation and exploitation of local orthographic imagery for road vehicle localisation. 2012 IEEE Intelligent Vehicles Symposium, 590–596.

Napier, A., Sibley, G., and Newman, P. (2010). Real-time bounded-error pose estimation for road vehicles using vision. 13th International IEEE Conference on Intelligent Transportation Systems, 1141–1146.

Oh, S.M., Tariq, S., Walker, B.N., and Dellaort, F. (2004). Map-based priors for localization. In 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566), volume 3, 2179–2184.

Revilloud, M., Gruyer, D., and Pollard, E. (2013). An improved approach for robust road marking detection and tracking applied to multi-lane estimation. In 2013 IEEE Intelligent Vehicles Symposium (IV), IV, 783–790.

Schadler, M., Stäckler, J., and Behnke, S. (2014). Rough Terrain 3D Mapping and Navigation Using a Continuously Rotating 2D Laser Scanner. KI - Künstliche Intelligenz, 28(2), 93–99.

Schwarz, M. and Behnke, S. (2014). Local navigation in rough terrain using omnidirectional height. Proceedings of the German conference on robotics (ROBOTIK), (June).

Sebsadji, Y., Tarel, J.P., Foucheur, P., and Charbonnier, P. (2010). Robust road marking extraction in urban environments using stereo images. 2010 IEEE Intelligent Vehicles Symposium, 394–400.

Sotelo, M.A., Rodriguez, F.J., Magdalena, L., Bergasa, L.M., and Boquete, L. (2004). A Color Vision-Based Lane Tracking System for Autonomous Driving on Unmarked Roads. Autonomous Robots, 16(1), 95–116.

Sun, D., Petovello, M.G., and Cannon, M.E. (2013). Ultratight GPS/Reduced-IMU Integration for Land Vehicle Navigation. IEEE Transactions on Aerospace and Electronic Systems, 49(3), 1781–1791.

Wu, T. and Ranganathan, A. (2012). A practical system for road marking detection and recognition. 2012 IEEE Intelligent Vehicles Symposium, 25–30.