OPTIMAL UAV LOCALISATION IN VISION BASED NAVIGATION SYSTEMS

Zhenlu Jin† *Xuezhi Wang*‡, *Quan Pan*† *and* *Bill Moran*‡

† School of Automation, Northwestern Polytechnical University, Xi'an, China
jinzl521@139.com *quanpan@nwpu.edu.cn*
‡ School of Computer and Electrical Engineering, RMIT University, Melbourne, Australia
xuezhi.wang@rmit.edu.au *bill.moran@rmit.edu.au*

ABSTRACT

Optimal determination of a UAV using a vision-based system to match images against a database is an important problem. It can be reformulated to the problem of using multi-region scene registration to match areas of a noisy and distorted image to a geo-referenced image. Under the assumptions that the mapping between sensed and geo-referenced images preserves gradients of straight lines cross mapping points on images and registration errors are all Gaussian distributed, we derive a two-stage weighted linear least square algorithm which localises the UAV optimally. Performance of the proposed algorithm is demonstrated via Monte Carlo multiple runs along with those available in literature.

Index Terms— UAV localisation, image registration error, multi-region scene matching, vision based navigation, weighted linear least square.

1. BACKGROUND

As an alternative to conventional GPS-based navigation systems, a vision-based system may be used to provide autonomous navigation of a UAV [1, 2]. One approach to the implementation of such a system involves the comparison of an image from an onboard ground-viewing camera against a database of images that have been geographically referenced [3, 4]. For reasons of computational complexity and to limit issues with distortion, registration of the aerial image against the geo-referenced image is often best carried out by using a multi-region/multi-block scene matching technique, where specific geo-referenced landmarks are chosen to be the reference locations [5, 6]. This approach is robust to temporal variation of the scene because of moving objects, and improves registration accuracy and efficiency [7, 8].

A deterministic multi-region image registration method is proposed in [7] without discussion on registration error. Multi-regions were extracted from an aerial image by a landmark selection method detailed in [6]. Statistically trained models for rivers, roads, sport fields and buildings are used to identify landmark areas from the aerial image. By matching the sub-regions with the reference image, the locations of a set of noisy sub-region centers on geo-referenced image are obtained and used together with prior knowledge on aerial image for UAV positioning. A pair of sub-region centers is searched over all matched sub-regions centers on the geo-referenced image which are able to infer a third sub-region center via triangulation with minimum offset to the observed one. Such a pair of sub-region centers are then used to infer the location of UAV on the geo-referenced image. This approach uses angle information as well as the distances between sub-regions in the localisation process.

In this work, we optimally find the UAV location on the geo-referenced image based on a set of noisy sub-region centers obtained from multi-region scene matching outcomes by taking account of registration error distributions. As shown in Fig.1, a set of sub-region centers on the geo-referenced image are obtained from multi-region scene matching and they preserve angles between lines across all points of the sensed image subject to a Gaussian noise. We assume that the transformation between the pair of images on the same scene preserves angles between lines across the set of sub-region centers. Furthermore, image registration error is modeled by a zero mean Gaussian distribution that is associated with each of the estimated sub-region centers on the geo-referenced image. The latter assumption can be justified from previous work in literature [9, 10, 11, 12].

In this paper, we formulate the underlying problem in a weighted linear least square framework, and eventually derive a two-stage estimator that uses information of angles between lines on the sensed image and the set of noisy location measurements to localise UAV on the geo-referenced image. In the first stage, sub-region center locations are estimated by a weighted linear least square estimation structure. These locations are then used as measurements to estimate the location of UAV in second stage. The proposed estimator is optimal in the sense that it uses all information available from the sensed

This work was supported by the US Air Force Office of Scientific Research (AFOSR) grant (FA9550-12-1-0418) via Arizona State University. It was also supported by the Major Program of National Natural Science Foundation of China (61135001), the National Natural Science Foundation of China (61473230, 61403307).

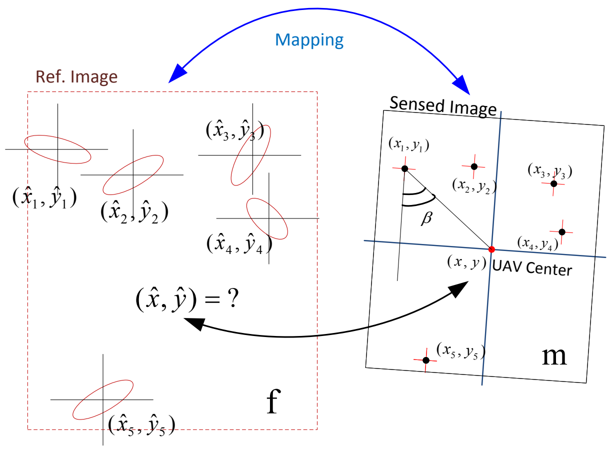

Fig. 1. Transformation between sensed and geo-referenced images that preserves angles between lines across sub-region centers.

image evidenced by the fact that the error covariance is identical to the inverse of Fisher information matrix. In addition, we also obtain the estimates of the set of sub-region centers with reduced uncertainties.

2. PROBLEM FORMULATION

In this work, the transformation between pair of images on the same scene is assumed to preserve angles between lines. This means, as shown in Fig. 1, the slope of the line between (x_1, y_1) and (x, y), $\tan(\beta) = \frac{x_1 - x}{y_1 - y}$, is a constant during the mapping while translations, scalings and rotations are allowed.

Statistical approach for the vision based UAV localisation requires the characterization of image registration errors. The performance of image registration can be affected by inconsistent color, object details and also the registration technique used, etc. As aforementioned, many researchers have studied the registration error analysis in various contexts and statistical models were proposed in [9, 10, 11, 12]. The basic results from the existing work are as follows.

1. Registration error is proportional to the averaged intensity differences between the pair of images.
2. Registration error is also inversely proportional to the signal to noise ratio.
3. Under mild assumptions, the displacement error can be modeled by a zero-mean Gaussian distribution.

In this paper, we assume that all sub-region registration processes in multi-region scene matching are independent with each other and displacement error of a sub-region is of zero-mean Gaussian distribution.

Let $\boldsymbol{m}$ and $\boldsymbol{f}$ represent the sensed and geo-referenced images, respectively. The UAV localisation problem can be described as follows.

What we know on the sensed image $\boldsymbol{m}$ are:

1. $\boldsymbol{U}_m = \{(x_1^m, y_1^m), \cdots, (x_r^m, y_r^m)\}$ – a set of r noisy free locations of sub-region centers;
2. and $\boldsymbol{X}_m = (x^m, y^m)$ – the center of UAV.

What we "observed" on the geo-referenced image $\boldsymbol{f}$ is a noisy version of the set of r sub-region centers

$$\boldsymbol{U}_f = \{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_r\} = \{(x_1^f, y_1^f), \cdots, (x_r^f, y_r^f)\} \quad (1)$$

which correspond to those sub-region centers in $\boldsymbol{m}$. They corrupt with zero-mean, independent and additive Gaussian noisy with variances, say,

$$\sigma_i^2 = \begin{bmatrix} \sigma_{i,x}^2, & 0 \\ 0 & \sigma_{i,y}^2 \end{bmatrix}, \qquad i = 1, \cdots, r \quad (2)$$

For the transformation with invariant angles between lines,

$$\frac{x_i^m - x^m}{y_i^m - y^m} = \tan\beta_i^m \equiv \tan\beta_i^f = \frac{x_i^f - x^f}{y_i^f - y^f} \quad (3)$$

and this applies to all lines across the sub-region centers.

The objective of UAV localisation is to find the UAV coordinate (x^f, y^f) on the geo-referenced image $\boldsymbol{f}$.

3. UAV LOCALISATION

3.1. Direct approach

This UAV localisation problem may be solved directly using least square method as follows. Denoted by sp_i, $i = 1, \cdots, r$ the slope of the line between (x_i^m, y_i^m) and (x^m, y^m). We can directly compute the slopes based on knowledge of the sub-region locations on $\boldsymbol{m}$:

$$sp_i = \frac{x_i^m - x^m}{y_i^m - y^m}, \quad i = 1, \cdots, r \quad (4)$$

and these slopes are invariant after image transformation from $\boldsymbol{m}$ to $\boldsymbol{f}$.

In the absence of image registration error, the UAV location (x^f, y^f) on the geo-referenced image $\boldsymbol{f}$ can be written as a function of each registered sub-region centers from (3) and (4),

$$x_i^f - sp_i y_i^f = x^f - sp_i y^f, \quad i = 1, \cdots, r. \quad (5)$$

Now we write (5) in the following vector form:

$$\boldsymbol{Z} = \begin{bmatrix} z_1 \\ \vdots \\ z_r \end{bmatrix} = \begin{bmatrix} 1 & -sp_1 \\ \vdots & \vdots \\ 1 & -sp_r \end{bmatrix} \begin{bmatrix} x^f \\ y^f \end{bmatrix} \quad (6)$$

where

$$z_i = x_i^f - sp_i y_i^f, \quad i = 1, \cdots, r.$$

We call the vector $\boldsymbol{Z}$ as the measurement about the unknown UAV location on the geo-referenced image $\boldsymbol{f}$. In the presence of image registration error, the measurement equation (6) becomes

$$\boldsymbol{Z} = \begin{bmatrix} z_1 \\ \vdots \\ z_r \end{bmatrix} = \begin{bmatrix} 1 & -sp_1 \\ \vdots & \vdots \\ 1 & -sp_r \end{bmatrix} \begin{bmatrix} x^f \\ y^f \end{bmatrix} + \begin{bmatrix} v_1 \\ \vdots, \\ v_r \end{bmatrix} \tag{7}$$

where

$$v_i = \Delta x_i^f - sp_i \Delta y_i^f.$$

Since both Δx_i^f and Δy_i^f are zero-mean Gaussian distributed with variances $\sigma_{i,x}^2$ and $\sigma_{i,y}^2$, we may write

$$v_i \sim \mathcal{N}(0, \sigma_{i,f}^2) \tag{8}$$

From definition, the variance $\sigma_{i,f}^2$ is derived as below.

$$\sigma_{i,f}^2 \triangleq \mathbf{E}[(\Delta x_i^f - sp_i \Delta y_i^f)^2] = \sigma_{i,x}^2 + sp_i^2 \sigma_{i,y}^2. \tag{9}$$

Now, let $\boldsymbol{X}_f = [x^f, y^f]^T$ and we write (7) in compact form as

$$\boldsymbol{Z} = \boldsymbol{H}\boldsymbol{X}_f + \boldsymbol{V} \tag{10}$$

where

$$\boldsymbol{H} = \begin{bmatrix} 1 & -sp_1 \\ \vdots & \vdots \\ 1 & -sp_r \end{bmatrix}, \quad \boldsymbol{V} = \begin{bmatrix} v_1 \\ \vdots, \\ v_r \end{bmatrix}.$$

Eq. (10) demonstrates that the underlying UAV localisation problem is formulated as a linear parameter estimation problem. It can be optimally solved by the linear least square method, or equivalently, by the maximum likelihood estimation method.

The estimation of $\boldsymbol{X}_f$ under LS criterion is to minimize the quadratic error

$$\boldsymbol{J} = [\boldsymbol{Z} - \boldsymbol{H}\boldsymbol{X}_f]^T \boldsymbol{R}^{-1} [\boldsymbol{Z} - \boldsymbol{H}\boldsymbol{X}_f] \tag{11}$$

where $\boldsymbol{R}$ is the stacked measurement errors of the form

$$\boldsymbol{R} = \text{diag}\left[\sigma_{1,f}^2, \cdots, \sigma_{r,f}^2\right].$$

By setting the gradient of (11) with respect to $\boldsymbol{X}_f$ to zero, we obtain the LS solution for the UAV location on the geo-referenced image as

$$\hat{\boldsymbol{X}}_f = [\boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{H}]^{-1} \boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{Z} \tag{12}$$

The associated covariance matrix $\boldsymbol{\Sigma}$ is given by

$$\boldsymbol{\Sigma} = [\boldsymbol{H}^T \boldsymbol{R}^{-1} \boldsymbol{H}]^{-1} \tag{13}$$

Remarks:

- The estimator implemented via (12) and (13) is optimal in the sense that it achieves Cramer Rao Lower Bound, which can be shown identical to (13).
- It can be easily shown that the solution under maximum likelihood criterion is identical to the LS solution as in (12) and (13).

3.2. Complete approach

While the above approach provides optimal solution for $p(\boldsymbol{X}_f|\boldsymbol{Z})$, we note that available information about $\boldsymbol{X}_f$ is only partially explored in this problem. In fact, we may derive a two-stage least square algorithm to improve localisation performance. The idea and algorithm implementation are the following:

1. We first estimate the set of sub-region centers on the geo-referenced image $\boldsymbol{f}$ using all slopes of available lines on sensed image $\boldsymbol{m}$, i.e., we find $p(\boldsymbol{u}_i|\boldsymbol{Z}_{u_i})$ based on the slopes $sp_j^{\boldsymbol{u}_i}$ using least square method, where $sp_j^{\boldsymbol{u}_i}$ are defined as
$$sp_j^{\boldsymbol{u}_i} \triangleq \frac{x_j^m - x_i^m}{y_j^m - y_i^m}, \quad i \neq j, \ i, j = 1, \cdots, r. \tag{14}$$
Similar to (10), $\boldsymbol{Z}_{u_i}$ can be formulated via (14). Thus, $p(\boldsymbol{u}_i|\boldsymbol{Z}_{u_i})$, distributions of the set of r sub-region centers are able to be estimated under least square criterion through (12) and (13).
2. Each of the set of estimated sub-region centers $\hat{\boldsymbol{u}}_i$ is then combined with the original observation $(\hat{x}_i, \hat{y}_i)$ to obtain a new set of estimated sub-region centers $\hat{\boldsymbol{U}}_c$. Assume two independent measurements $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$ on a unknown constant parameter with covariances P_1 and P_2 respectively. The fusion rule for combing the two is
$$P = \left(P_1^{-1} + P_2^{-1}\right)^{-1} \tag{15}$$
$$\boldsymbol{y} = P\left(P_1^{-1}\boldsymbol{y}_1 + P_2^{-1}\boldsymbol{y}_2\right) \tag{16}$$
3. We then apply the direct approach described in 3.1 to find $p(\boldsymbol{X}_f|\boldsymbol{Z}_c)$, where $\boldsymbol{Z}_c$ is formulated and computed using the newly estimated set of estimated sub-region centers $\hat{\boldsymbol{U}}_c$.

We call this two-stage least square estimator as a complete approach because the measurement set $\{\boldsymbol{Z}, \boldsymbol{Z}_c\}$ take all slopes of lines across sub-regions centers on sensed image $\boldsymbol{m}$ into account in calculating $\boldsymbol{X}_f$.

Note that the noise term v_i in (7) is derived under the assumption that the x and y components of registration error of a sub-region are independent as demonstrated by (2). However, once the locations of the set of registered sub-region on $\boldsymbol{f}$ are updated in the first stage LS, the two error components are correlated. Now, we derive a method to incorporate cross terms of registration error covariance into the measurement variance similar to (9) in the second stage LS for finding $\boldsymbol{X}_f$.

In the following derivation, we use the observation on the ith sub-region center location (x_i^f, y_i^f) as an example. Assume $\hat{x}_i^f$ and $\hat{y}_i^f$ are jointly Gaussian distributed random variables with covariance matrix $\boldsymbol{Q}_i$. The variance of weighted sum of two correlated normal random variables can be calculate as

$$\text{Var}(p\hat{x}_i^f + q\hat{y}_i^f) = p^2\sigma_{i,x}^2 + q^2\sigma_{i,y}^2 + 2pq\sigma_{i,xy}^2 \tag{17}$$

where p and q are the weights of two correlated location errors. The covariance matrix of the noise is

$$\boldsymbol{Q_i} = \begin{bmatrix} \sigma_{i,x}^2 & \sigma_{i,xy}^2 \\ \sigma_{i,yx}^2 & \sigma_{i,y}^2 \end{bmatrix} \quad (18)$$

Denoted by Γ, the weight matrix is given by

$$\Gamma = \begin{bmatrix} p^2 & pq \\ qp & q^2 \end{bmatrix}. \quad (19)$$

(17) can then be written as

$$\text{Var}(p\hat{x}^f + q\hat{y}^f) = \Gamma \circ \boldsymbol{Q}_i = \text{tr}(\Gamma \boldsymbol{Q}_i^T) \quad (20)$$

where $(A \circ B)_{ij} \triangleq A_{ij} \cdot B_{ij}, \ \ A, B \in \mathbb{R}^{m \times n}$ denotes the Hadamard product of matrices A and B [13]. The sum of all elements in the Hadamard product is equal to the trace of AB^T.

(20) indicates that as long as the weights p and q are known, we can calculate the variance of (9) for correlated registration error of jointly Gaussian distribution. For the underlying problem, the weights p and q are calculated as follows.

$$p = \frac{x_i^m - x^m}{\sqrt{(x_i^m - x^m)^2 + (y_i^m - y^m)^2}} \quad (21)$$

$$q = \frac{sp_i^2(y_i^m - y^m)}{\sqrt{(x_i^m - x^m)^2 + (y_i^m - y^m)^2}} \quad (22)$$

4. SIMULATION RESULT AND DISCUSSIONS

The proposed algorithms were implemented in Matlab under a localisation scenario shown in Fig. 2, which involves 5 sub-region centers (numbered as 1,2,3,4,5) over a scene of 500×500 pixels. The sub-region centers are measured by a multi-region scene matching algorithm subject to a zero mean Gaussian distributed (registration) error. The standard deviations of registration errors are given in Table 1.

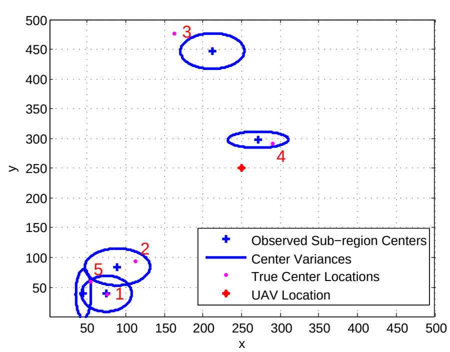

Fig. 2. Simulation Scenario

1000 Monte Carlo runs were carried out for either the Direct LS and complete LS algorithms and Root Mean Squared Errors (RMSE) for the UAV localisation are compared in Table 2 with those achieved by the deterministic algorithm [6].

Region Index (i)	1	2	3	4	5
$\sigma_i(x)$ (pixels)	32.50	42.36	41.39	39.34	9.98
$\sigma_i(y)$ (pixels)	30.09	30.97	29.93	13.75	41.08

Table 1. Standard deviation of registration errors for sub-region centers.

Algorithm	Direct LS	Complete LS	Deterministic
$RMSE$ (pixels)	15.46	9.38	19.94

Table 2. RMSE comparison over 1000 runs.

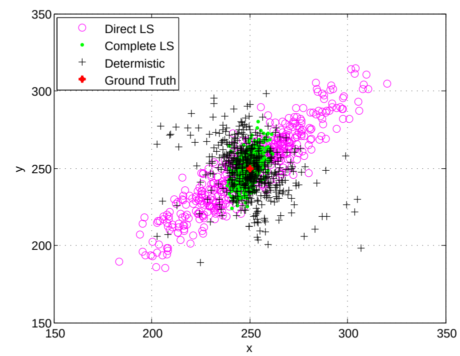

Fig. 3. Comparison of estimated UAV location spread over 500 runs.

To get a better idea of the error performance, we compared the estimated UAV location over 500 runs for all algorithms under consideration in Fig. 3.

According to the above simulations, in addition to the best RMSE performance on UAV location estimation, the Complete LS algorithm also provide sub-region center estimates with reduced covariances.

5. CONCLUSIONS

In this paper, UAV localisation using measurements from a vision based navigation system is considered. The measurements can be thought as the outcomes of a linear mapping of a set of points on an aerial image onto a geo-referenced image with additive Gaussian noise. A two-stage weighted linear least square estimator is derived which uses the measurements as well as knowledge of the aerial image to optimally localise the UAV. The performance of the proposed algorithm is illustrated in simulation. Finally, while the proposed algorithm is efficiently formulated and behavior as expected, it does not properly handle the rotational error in image registration. Approach which takes full consideration of the conformal mapping between the pair of images involve nonlinear formulation will be addressed in a separate paper elsewhere.

6. REFERENCES

[1] Matthes Rieke, Theodor Foerster, Jakob Geipel, and Torsten Prinz, "High-precision positioning and real-time data processing of uav-systems," *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 38, pp. 1–C22, 2011.

[2] Gianpaolo Conte and Patrick Doherty, "A visual navigation system for uas based on geo-referenced imagery," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 38, pp. 1, 2011.

[3] Aurelien Yol, Bertrand Delabarre, Amaury Dame, Jean-Emile Dartois, and Eric Marchand, "Vision-based absolute localization for unmanned aerial vehicles," in *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*. IEEE, 2014, pp. 3429–3434.

[4] Gianpaolo Conte and Patrick Doherty, "Vision-based unmanned aerial vehicle navigation using geo-referenced information," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, pp. 10, 2009.

[5] Lin Song, Yong-mei Cheng, Lu Yu, and Liang Yu, "A method of waypoint selection in aerial images for vision navigation," *International Journal of Optics*, vol. 2014, 2014.

[6] Z. Jin, X. Wang, M. Morelande, B. Moran, Q. Pan, and C. Zhao, "Landmark selection for scene matching with knowledge of color histogram," in *Proc. 17th International Conference on Information Fusion*, Salamanca, Spain, July 7-10 2014.

[7] Z. Jin, X. Wang, B. Moran, Q. Pan, and C. Zhao, "Efficient scene matching using salient regions under spatial constraints," in *Proc. 17th International Conference on Information Fusion*, Salamanca, Spain, July 7-10 2014.

[8] Yaojun Li, Ying Yu, Quan Pan, and Chunhui Zhao, "Scene matching based on spatial relation constraint in suitable-matching area," in *Intelligent Computing and Intelligent Systems, 2009. ICIS 2009. IEEE International Conference on*. IEEE, 2009, vol. 4, pp. 598–603.

[9] Jan Kybic, "Fast no ground truth image registration accuracy evaluation: comparison of bootstrap and hessian approaches," in *The Fifth IEEE International Symposium on Biomedical Imaging (ISBI 2008)*. IEEE, 2008, vol. I, pp. 792–795.

[10] Jan Kybic, "Bootstrap resampling for image registration uncertainty estimation without ground truth," *IEEE trans. on Image Processing*, vol. 19, no. 1, pp. 291–294, January 2010.

[11] C. D. McGillem and M. Svedlow, "Image registration error variance as a measure of overlay quality," *IEEE trans. on Geoscience Electronics*, vol. GE-14, no. 1, pp. 44–49, January 1976.

[12] I.S. Yetik and Arye Nehorai, "Performance bounds on image registration," *Signal Processing, IEEE Transactions on*, vol. 54, no. 5, pp. 1737–1749, May 2006.

[13] Elizabeth Million, "The hadamard product," *Course Notes*, vol. 3, 2007.