Application of vision based techniques for position estimation

Seema. B. S[1], Sheetal N[2], Hamsa Rekha S D[3]

[1, 2, 3] *Assistant professor, Department of Electronic & Instrumentation Engineering, Dr. Ambedkar Institute of Technology, Bangalore, Karnataka, India*

Abstract- *The objective is to set up wireless communication for Unmanned Aerial Vehicles (UAVs) and to estimate the position of the UAV using vision-based techniques with an onboard camera, in a real-time, continuously interacting scenario. There are three types of data to be sent or received from the UAV to the Ground Control Station (GCS) – video, control and command and navigational data. Navigational data contains information such as azimuth, elevation, latitude, longitude, roll, pitch and yaw. Here, vision based navigation is explored with different algorithms (RANSAC feature detection and normalized cross correlation with prior edge detection) and the benefits and disadvantages of each are compared. Vision based navigation uses a camera onboard the UAV to continuously transmit an aerial video of the ground to the GCS and the position of the UAV is determined using geo-referenced images such as those obtained from Google Earth.*

***Keywords-* Unmanned Aerial Vehicle, Ground Control Station, Vision Based Technique image matching, Position Estimation.**

I. INTRODUCTION

Unmanned Aerial Vehicles (UAVs) have been developed and put into use since World War I, where they were used for military applications. The first UAV was a missile, Kettering Bug, named after its developer, David Kettering. Since then UAVs have found various uses in both commercial and military applications. They are especially useful for collecting information about different places where humans cannot venture. They can also be used to transport various materials to disaster-affected regions with minimum human interference. In this section, first, the previous methods are discussed and the need for vision based position estimation is explored. Then the objectives of this project are detailed.

Two methods were traditionally used for determining the position of the UAV to continuously transmit to the Ground Control Station – namely, GNSS (Global Navigation Satellite Systems) and INS (Inertial navigation Sensors). GPS uses time measurement of signals and triangulation techniques to estimate position and velocity whereas INS consists of a computer and sensors such as accelerometers, gyro, altimeters and magnetometers.

However, GNSS jamming devices are easily available and the signals can be jammed. For INS, the error in output grows with time. These are called unbound or drift errors, which can only be minimized not corrected. Thus, there is a need for a third technique to estimate position of the UAV in the event of a failure of GNSS or INS. Vision based position estimation can be used as a backup for either of the two traditional methods [1].

II. METHODOLOGY

For testing purposes, the camera and transmitter will consist of a PC, which mimics the UAV and sends the video to the ground control station. Inputs will be a geo-referenced image with latitude and longitude grid information and a UAV photo (after passing through the camera model). In real time, the camera must continuously transmit the video footage of the ground to the GCS where it is received and processed.

The ground station will consist of another PC, where the image received will be processed and the position will be determined using the aforementioned algorithms. In real-time scenario, the geo-referenced images will have to be downloaded according to the initial coordinates of the flight and approximate flight trajectory. The geo-referenced images should cover the entire trajectory of the flight for improved accuracy. The final position is estimated from the geo-referenced images' latitude longitude grid or by spatially referencing the geo-reference image [2, 3]. If the geo-reference database can be constructed or purchased, vision based techniques can be used by image matching with the images in the database though this may be time consuming. If the initial coordinates and approximate flight trajectory are known, the process can be speeded up. Here, however, the erroneous GPS coordinates are used to obtain a geo-reference image from Google Static Maps where the area depends on the zoom level and the input coordinates are at the center.

Two main matching techniques have been implemented – Normalized Cross Correlation with prior edge extraction [4, 5] and RANdom SAmple Consensus (RANSAC) feature detection algorithm [4]. The results such as effectiveness and time consumption, along with functionality for RANSAC algorithm for various levels of noise (determined by variance of Gaussian noise introduced) and for various levels of blur are compared, along with advantages and disadvantages of each. The comparison is essentially between feature detection methods and line detection methods. Position can be estimated after image matching, using spatially referenced objects [5, 6].

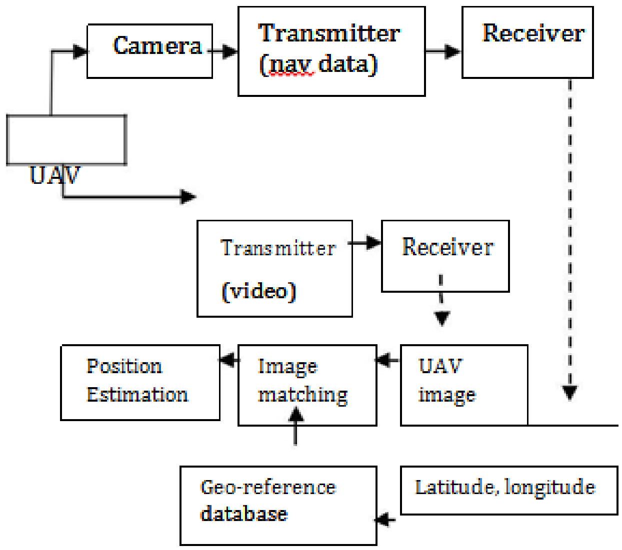


Figure 1. Overall block diagram

III. IMAGE MATCHING TECHNIQUE

The two algorithms are tested and compared as follows:

A. Normalized Cross Correlation:

If there are two images, A and B, the cross correlation image is defined as follows:

$$C(u,v)\ \sum_{x,y\ 0}^{m} A(x,y)B(xu,\ y\ \ v) \quad (1)$$

This formula is not ideal as it is invariant to energy changes and thus may lead to matching errors. Therefore, an improved formula is used as given as follows:

$$\sum_{x,y\ 0}^{m,n} \frac{1}{A} \frac{A(x,y)B(x\ \ u,y\ v)}{B} \quad (2)$$

The input image and geo-reference images are taken and pre-processing techniques such as noise filtering and de-blurring are performed. Then the edges are detected, the normalized cross correlation matrix is computed. The location of the peak in noted and the correlation offset is extracted. If there are several peaks, the largest of them is to be extracted. The image can then be recovered. The UAV image is overlaid on the geo-reference one using the peak calculated to note its position, this implementation is as shown in Figure 2.

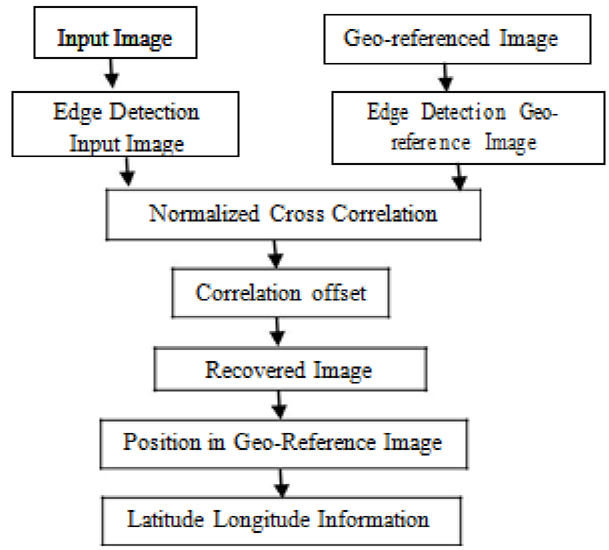


Figure 2. Normalized Cross Correlation Algorithm

B. RANSAC Algorithm:

The RANdom SAmple Consensus (RANSAC) algorithm is a feature based method of image matching that is more reliable than the edge based methods as they compute features which are rotation invariant and robust

[6] These features are classified into two i.e. inliers (points that match) and outliers (points that do not match). The RANdom SAmple Consensus (RANSAC) algorithm is used to find the features of the image and eliminate those that do not match (outliers) [7]. Features have to be robust, rotation invariant, invariant to distortions and obstructions and independent. The features can be selected manually for comparison but the UAV that has a camera on board cannot hover over a particular area. It has to be in constant motion. Therefore, the extraction and comparison will be automated and continuous. RANSAC algorithm gives accurate results even though there is scaling and rotation between the geo-referenced image and camera image at the cost of execution time [8].

The input image and geo-reference images are taken and converted to grayscale. The features are then extracted from the images and compared. A minimum number of 15 features are required without which the program will need to run again, which is one of the disadvantages of RANSAC. The image is recovered by recovering the transform applied on it using the features detected. The recovered image is overlaid on the geo- reference one to note its position. Then, the position of the UAV in the geo-reference image is determined. Thus the latitude and longitude information can be extracted as shown in Figure 3.

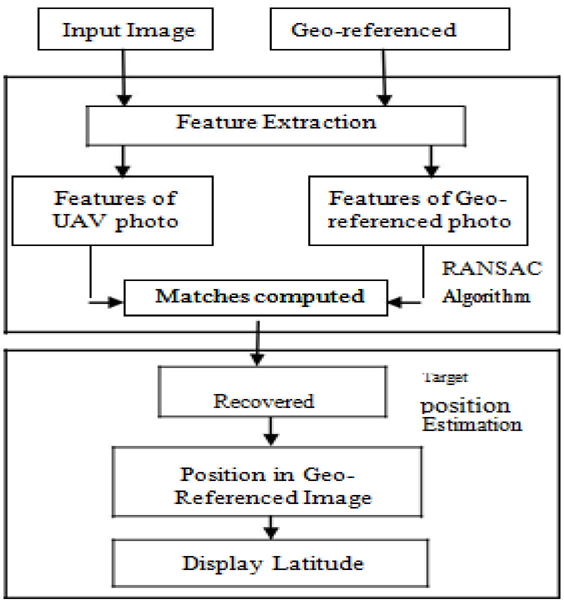


Figure 3. RANSAC Algorithm

IV. POSITION ESTIMATION

The UAV image is overlaid on the geo-reference image and the position of the UAV image matrix is stored in indices. In this section, the exact position estimation from the position of the UAV image in the geo-referenced one is discussed. The geo-referenced image contains information such as latitude and longitude pair at the center. First, the bounds of the geo-referenced image must be calculated. The pixel values at the center of the geo-referenced image can be calculated by converting the latitude and longitude values into meters and then from meters into pixels [9]. Latitude and longitude obtained from the geo-reference image will be in decimal degrees i.e., WGS84 (World Geodetic System) which are converted into meters using the following formulae,

$$x \quad \frac{lon\ OS}{180} \tag{3}$$

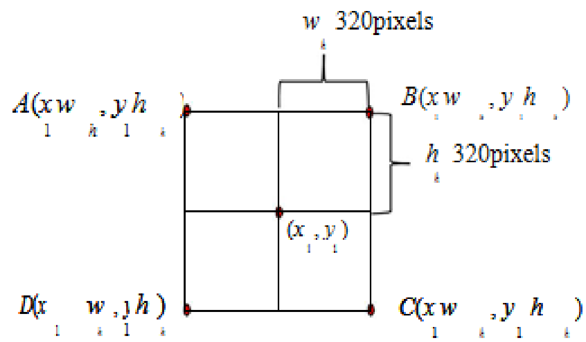
Where

$$y \quad \frac{\log \frac{\tan(90 lat)}{360}}{180} \tag{4}$$

$$OS \quad \frac{R_e\ 2}{2}$$

TS = Tile size (320 X 320 used here).
ZL = Zoom level (17-19 for Google maps)

Thus, the pixel values at the center are determined.
The corner pixel location can be calculated as follows:



Where width = height = 640 pixels
W_h= width/2 = 320 pixels
h_h =height/2 = 320 pixels

Longitude is represented in the x direction and latitude in represented in the y direction. Then, the corner pixels must be converted into meters and then converted to latitude and longitude using the following formulae:

$$A_{WGS}\ (lon,lat)\ \frac{(x_1\ w_h)res\ OS}{OS}180, \frac{(y_1\ h_h)res\ OS}{OS}180 \tag{7a}$$

$$B_{WGS}\ (lon,lat)\ \frac{(x_1\ w_h)res\ OS}{OS}180, \frac{(y_1 h_h)res\ OS}{OS}180 \tag{7b}$$

$$C_{WGS}\ (lon,lat)\ \frac{(x_1\ w_h)res\ OS}{OS}180, \frac{(y_1 h_h)res\ OS}{OS}180 \tag{7c}$$

$$D_{WGS}\ (lon,lat)\ \frac{(x_1\ w_h)res\ OS}{OS}180, \frac{(y_1\ h_h)res\ OS}{OS}180 \tag{7d}$$

Thus, the bounding values of the latitude longitude for the image are calculated. A spatial referencing object is created and the bounding values calculated in the previous step are assigned to it [10]. The size of the image (640x640) is also assigned. The position of the UAV image within the geo-referenced image is calculated. In feature detection method, the position of the recovered image after applying the geometric transform is stored [11, 12]. Using the spatial referencing object and the known pixel values of the position as determined, the latitude longitude values can be computed using the pix2latlon () function in MATLAB. Once the position of the target is determined then the location is displayed on the click of the mouse pointer.

VI. RESULTS AND DISCUSSION

Algorithm uses geo-reference images taken from Google Earth and the UAV shot which is obtained from the onboard camera .The latitude and longitude pair used for all the testing purposes is 12.9653471, 77.651822703 respectively. The camera model is applied to obtain the UAV image (Figure 4) is used as the UAV photo for evaluating the algorithms.

Figure 4.UAV image obtained from camera model

The image matching done by NCC showing the peak determined based on edge matching and the recovered images are shown in Figure 5 and features detected by RANSAC algorithm, matched and overlaid pictures are as shown in Figure 6.

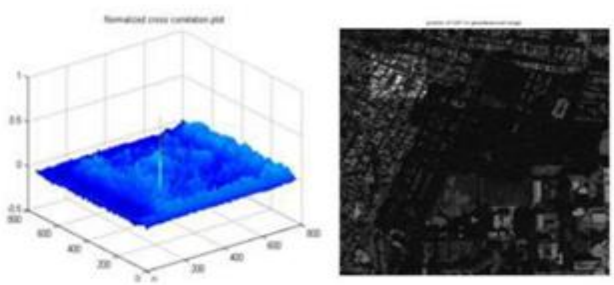


Figure 5. Peak determined using edge matching and the recovered image (NCC)

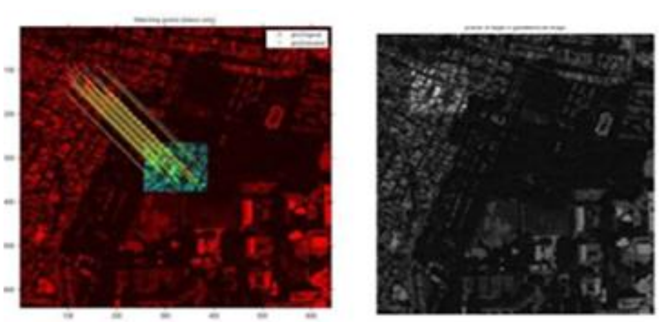


Figure 6. Detected Inliers and the Recovered Image (RANSAC)

The errors due to added noise and computational inaccuracies are tabulated in this section. The ideal latitude longitude values are calculated prior to testing by manually checking the latitude longitude with Google maps. The ideal latitude longitude pair is (12.9663, 77.6508). Table 1 represents the distance per degree change of latitude and longitude is used to calculate the error in meters.

Table.1. Surface distance per degree change in latitude

Latitude (degrees)	Surface distance per degree change in latitude (km)	Surface distance per degree change in longitude (km)
0	110.874	111.320
15	110.649	107.551
30	110.852	96.486
45	111.132	78.847
60	111.412	55.800
75	111.618	28.902
90	111.694	0.000

Table.2 represents, the average time taken by RANSAC is more than twice the amount of time taken by normalized cross correlation.

Table.2. Computational time taken by the two methods

Method	Average Time (s)
RANSAC	8.023
Normalized Cross Correlation	4.964

Even though NCC is faster, it fails in the event of noise, scaling, rotation or blur. Thus, the camera would have to be positioned perpendicular to the ground and the navigation commands would have to include separate measures to ensure that the necessary opposite rotation be taken for the camera to oppose the UAV's rotation. We can say that the accuracy of RANSAC under ideal circumstances (minimum noise and least blur) is approximately 10m as seen in Table 3. The algorithm fails in the event of variance of Gaussian noise being equal to or greater than 0.1.Whereas, NCC has better accuracy after noise has been removed, but the noise removal process adds to the computational time Table. 4.

Table.3. Functionality of RANSAC under Noise Variance

Noise Variance (σ)	Latitude (degree)	Longitude (degree)	Error due to	
			Latitude. (m)	Longitude (m)
0.005	12.9662	77.6507	11.0649	10.7551
0.01	12.9664	77.6512	11.1298	10.0204
0.015	12.9661	77.6512	22.1298	23.0204
0.02	12.9665	77.6511	22.1113	32.4533
0.05	12.9666	77.6511	33.1923	32.2653
0.1	-	-	-	-

Table 4: Functionality of NCC under Noise Variation

Noise Variance (σ)	Latitude (degree)	Longitude (degree)	Error due to	
			Latitude (m)	Longitude (m)
0.01	12.9664	77.6509	11.1298	21.5102
0.02	12.9664	77.6509	11.1298	21.5102
0.03	12.9664	77.6509	11.1298	21.5102
0.04	12.9664	77.6509	11.1298	21.5102
0.05	12.9664	77.6509	11.1298	21.5102

VII. CONCLUSIONS

For position determination using vision based techniques, edge detection methods and feature detection methods have been compared. The NCC algorithm can be equipped for real time if the time taken by the pre-processing mechanisms can be minimized and the issues of scaling and rotation can be addressed. It is advantageous as the error in latitude longitude is almost constant and hence can be easily corrected as opposed to feature based methods where the noise plays a major role in determining the error. But, the RANSAC is more attractive for real time usage as the accuracy is approximately 10m.

RANSAC fails in the event of extreme distortion or noise but is relatively invariant to rotation; noise and scaling as feature detection methods are more efficient than edge detection and brute force comparison methods. RANSAC has no upper bound on the time it takes to compute the features. One of the main drawbacks of this algorithm is if minimum numbers of features are not detected then the algorithm has to be executed again.

ACKNOWLEDGMENT

We would like to thank Mr. N Santha kumar, FMCD Department, CSIR-National Aerospace Laboratory and our institution for their constant encouragement and support to carry out this work.

REFERENCE

[1] Markus Olgemar, "Camera based navigation: Matching between sensor reference and video image",Master Thesis, Linkopings University, 2008.

[2] Carol Martinez, Ivan F. Mondragon, Miguel A.Olivares-Mendez, Pascual Campoy, "Onboard and ground vision pose estimation techniques for UAV control", Master Thesis, Universiad Politechnecia de Madrid.

[3] Dong-Gyu Sim, Rae-Hong Park, Rin-Chul Kim, Sang-Uk Lee, Ihn-Cheol Kim," Navigation Parameter Estimation from sequential Aertial images", Master Thesis, Seoul National University, Korea.

[4] Dong-Gyu Sim, Rae-Hong Park, Rin-Chul Kim, Sang-Uk Lee, Ihn-Cheol Kim, "Integrated Position Estimation Using Aerial Image sequences", IEEE, Transactions on pattern Analysis and Machine Intelligence, Vol. 24, 2002.

[5] Gianopaolo Conte, Patrick Doherty," Vision based unmanned Aerial Vehicle navigation using geo-referenced information" ,Master Thesis, Linkopings University.

[6] Serge Belongie, Jitendra Malik, Ian Puzicha , "Shape matching and object recognition using shape contexts", IEEE, Transactions on pattern Analysis and Machine Intelligence, Vol. 24, 2002.

[7] V. Venkateshwar, Rama Chellappa, "Extraction of straight lines in aerial images", IEEE, Transactions on pattern Analysis and Machine Intelligence, Vol. 14, pp. 1111-1114, 1992.

[8] Mansoor Ahsan, Suhail Akhtar, Adnan Ali, Farrukh Mazhar, Muddssar Khalid, "An algorithm for autonomous aerial navigation using MATLAB Mapping toolbox", World Academy of Science, Technology and Engineering 66, pp. 858-861, 2012.

[9] Ryan Schaafsma, "UAV imaging position errors", Master Thesis, California State university, 2012. [10]Stuart Ness, "Geo-registration of Remotely Sensed Imagery", Department of Computer Science and Engineering, University of Minnesota.

[10] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, Estimation with Applications to Tracking and Navigation: Theory, Algorithms, and Software, Wiley, New York, July 2001.

[11] Y. C. Chen, F. Y. Hsiao, J. F. Shen, F. C. Hung, S. Y. Lin, "Application of MATLAB to Vision based navigation of UAVs",8th IEEE International Conference on Control and Automation, pp. 877-882, June 2010.