Multiple Targets Geolocation using SIFT and Stereo Vision on Airborne Video Sequences

K. Han* and G. N. DeSouza**

Abstract—We propose a robust and accurate method for multi-target geo-localization from airborne video. The difference between our approach and other approaches in the literature is fourfold: 1) it does not require gimbal control of the camera or any particular path planning control for the UAV; 2) it can instantaneously geolocate multiple targets even if they were not previously observed by the camera; 3) it does not require a geo-referenced terrain database nor an altimeter for estimating the UAV's and the target's altitudes; and 4) it requires only one camera, but it employs a multi-stereo technique using the image sequence for increased accuracy in target geo-location. The only requirements for our approach are: that the intrinsic parameters of the camera be known; that the on board camera be equipped with global positioning system (GPS) and inertial measurement unit (IMU); and that enough feature points can be extracted from the surroundings of the target. Since the first two constraints are easily satisfied, the only real requirement is regarding the feature points. However, as we explain later, this last constraint can also be alleviated if the ground is approximately planar. The result is a method that can reach a few meters of accuracy for an UAV flying at a few hundred meters above the ground. Such performance is demonstrated by computer simulation, in-scale data using a model city, and real airborne video with ground truth.

Keywords: Unmanned aerial vehicle, geo-location, airborne video, multiple target.:

I. INTRODUCTION

Current research on target geo-location using passive sensors – e.g. monocular airborne camera – has achieved remarkable results [14], [3], [13], [4], [17], [16], [7]. In some cases, the uncertainty in the estimation is under 10 meters, with the UAV flying at an altitude of a few hundred meters. Unfortunately, these same achievements have been possible only by imposing some severe constraints to the systems.

In [16], [7], for example, the proposed system could track a moving target on the ground while controlling the vehicle and its gimbal camera. At the same time, the system could estimate the target geo-location within just 5m, when flying at an altitude of over 300m. However, such results were possible only if the velocity – magnitude and heading – of the target was constant. Also, in order to estimate the target altitude accurately, this and other systems in the literature had to rely on an equally accurate geo-referenced terrain database, a Geospatial Information System (GIS), or the Perspective View Nascent Technologies system (PVNT) [18].

In another system found in the literature [14], the authors proposed the use of a Recursive Least Squares filter to reduce the error in geo-location. However, this type of filter imposes

Department of Electrical and Computer Engineering, University of Missouri, 349 Eng. Building West, Columbia, MO, USA
*khx8d@mizzou.edu **DeSouzaG@missouri.edu

an even stronger constraint on the target: it had to remain stationary. Still, in the same work, the authors provided a quite useful analysis of the error, they determined its main sources and presented a study on the sensitivity and propagation of uncertainties in their method. Despite that, the errors reported were still quite large – 15m for a UAV at 60m high. Later, in a continuation of their previous work, [3], the authors extended even further their analysis of the error, they identified the sources of zero-mean noise versus constant bias, and derived an expression for the optimum altitude of the vehicle as a function of the path radius and the pixel area of the target object on the image plane. All that reduced the error to less than 10m at an optimum altitude slightly lower than 100m. However, the target still had to be stationary and the error in instantaneous estimates, i.e. without the RLS filter, could reach more than 40m, especially for non-optimum altitudes of the UAV.

In most cases, multiple frames were also required: whether they were used to estimate the target velocity [14], to reduce error using a RLS filter [3], Kalman filter [13], or to guarantee convergence and controllability of the gimbal cameras or the vehicle itself [16], [7]. In other cases, [15], even a specific path planning strategy for the UAV had to be outlined in order for the system to achieve good results.

In this work, we propose an efficient and accurate method to calculate the geo-location of multiple targets. Our method does away with most of the above constraints (e.g. elevation maps, stationary targets, etc.), while it provides a robust and instantaneous estimate of the target geo-location. By not employing camera gimbal control – where only one object can occupy the center of the image – we can handle multiple targets at the same time. Also, as in [14], we provide an analysis of the error in our method and a similar discussion regarding sensitivity and error propagation. Finally, to avoid the need for geo-referenced terrain databases or expensive altimeters, which can be impractical due to limitations in weight and power consumptions, we propose an algorithm to estimate the altitude of the UAV based on the detection of multiple feature points around each target object. At the end of the paper, we present the results of our method in three different testing scenarios: simulation, in-scale test using a model city, and real airborne data with ground truth.

II. PROPOSED METHOD

As Figure 1 illustrates, our method can be divided into three major steps, which are executed concurrently: object tracking; target geo-location; and altitude estimation. In the first step, object tracking, multiple objects can be tracked simultaneously
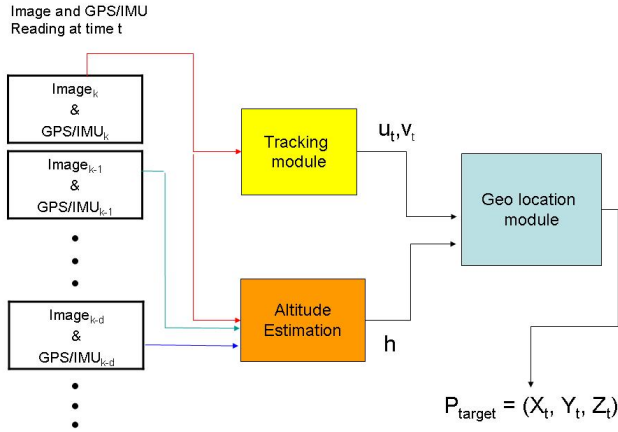
Figure 1. Modules in our framework

using a combination of differential optical flow and the KLT feature tracker [12], [2]. Due to space limitation, we will not describe this step of our framework. It should suffice to say that for each object detected and tracked by this algorithm, their pixel coordinates are made available to the other modules of the system.

In the second step, target geo-location, we start the discussion assuming that the altitude of the UAV with respect to the target is known. As in most systems today, [3], [14], [16], [7], [13], this information is required in order to calculate the distance and angle between the camera and the target. However, unlike those same systems, our method does not require the assumption of a flat terrain [13], a geo-registered terrain map [16], [7], or any other constraint on how the target is allowed move [3], [14].

In the final step of the method, that same altitude that was assumed to be known in the second step is actually estimated and fed back into the previous step. Since each tracked object may be on a different ground plane, various estimates are in fact computed: each one representing the height of the camera/UAV with respect to an specific tracked object. Together with the pixel coordinates of the targets, this information is also made available to the other modules of the system. Without loss of generality, in the next sections we will explain the method assuming that one single object is being tracked. The extension of the method for multiple targets can be easily inferred from such explanation. As we will also explain later, the only constraint for the method to work is the existence of enough feature points around the object – i.e. on the same ground plane surrounding the object. If that constraint is not satisfied, the system can assume that all targets are on the same plane and estimate a single altitude for all targets. But before we explain our method in more detail, we need first to establish some notations.

A. Coordinate System Convention

In order to determine the geo-location of an object, we must assume the existence of a world reference frame. Conceptually, however, it is not really important what reference frame is used. That is, let the coordinate frames with subscripts I, C, and U represent, respectively, the inertial (or world), the camera, and the UAV coordinate frames. Now, if the 3D pose (position and orientation) of an object can be determined with respect to, say, the camera coordinate frame, then the same object pose with respect to the inertial reference frame (that is, its geo-location) can be easily derived by a simple transformation from the camera coordinate frame to the inertial coordinate frame – which is given by the GPS/IMU data. In other words, since the relative pose of the camera frame with respect to inertial frame can be expressed by a homogeneous transformation matrix:

$$^IH_C = \begin{bmatrix} ^IR_C & ^IT_C \\ \mathbf{0}_{1\times3} & 1 \end{bmatrix} \tag{1}$$

where ^IR_C represents the 3x3 rotational component, and ^IT_C is the 3x1 translational vector. Then, the pose of an object with respect to the camera coordinate frame can be converted into its pose with respect to the inertial frame by a simple multiplication:

$$^IP = {}^IH_C * {}^CP$$

Actually, the transformation from the camera frame into the inertial frame can be further specified by

$$^IH_C = {}^IH_U * {}^UH_C$$

where U is the UAV coordinate frame standing between the camera and the inertial frames.

Once again, the first transformation, ^IH_U, represents the pose of UAV coordinate frame, and it can be obtained from the GPS/IMU on-board the UAV. Similarly, the second matrix, ^UH_C, represents the pose of the camera inside the UAV, which is some times controlled by a gimbal device. However, whether a gimbal device is present or not, the challenges in determining the above homogeneous transformations are quite similar to those known in the robotics circles by *hand-eye* and *head-eye* calibration. In this paper, we will assume that such calibration was carried out by an accurate procedure similar to the one developed in [5], [9].

B. Target Geo-Location

The first step of our algorithm consists of estimating the target geo-location using a single image. This process becomes quite straightforward, if we assume that the altitude (or height) of the camera with respect to the target is known. Figure 2 illustrates this idea, which we further detail next. Unless specified otherwise, all 3D coordinates below (points P and p) are with respect to the inertial coordinate frame. Also, as we mentioned earlier, the transformation between camera and inertial coordinate frames can be achieved by a simple matrix multiplication using ^IH_C (or its inverse ^CH_I).
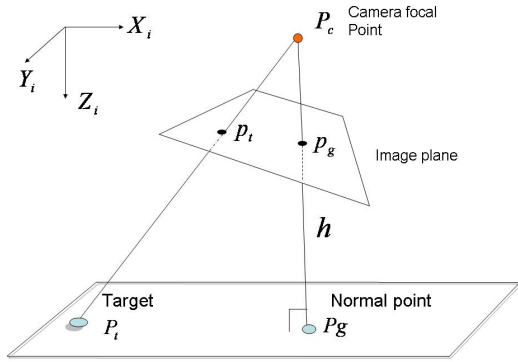
Figure 2. Estimation of the target pose from a single image.

Let us call P_c the focal point of the camera – which is also the origin of the camera coordinate frame. Since the 3D coordinate of this point is obtained from the camera hand-eye calibration and the GPS/IMU readings, we can calculate the homogeneous coordinates of the closest point on the ground P_g by simply adding h to the z-coordinate of P_c, i.e. $P_g = P_c + [0, 0, h, 0]^t$.

Also, the homogeneous coordinate of a point in space and its projected pixel coordinates are related by the following expression [8]:

$$\begin{bmatrix} s.u \\ s.v \\ s \end{bmatrix} = A_{int} * {}^{C}H_I * \begin{bmatrix} {}^{I}x \\ {}^{I}y \\ {}^{I}z \\ 1 \end{bmatrix} = A_{int} * \begin{bmatrix} {}^{C}x \\ {}^{C}y \\ {}^{C}z \\ 1 \end{bmatrix} \qquad (2)$$

where, ${}^{C}H_I$ is usually referred to as the extrinsic parameters of the camera, while A_{int} contains its intrinsic parameters and is given by:

$$A_{int} = \begin{bmatrix} \alpha_u & \gamma & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \qquad (3)$$

Where, α_u and α_v are associated with the focal length of the lens; γ is the axes skew factor; and (u_0, v_0) represents the principle point in the image coordinate frame. Now, since P_t, the target point, is on the same plane as P_g, its z coordinate with respect to the inertial coordinate frame is also h. Also, since the point p_t represents the target as observed by the camera, its pixel coordinates are provided by the tracking algorithm. Using eq. (2) and the information above, we can solve for the remaining two coordinates of P_t with respect to the inertial frame, i.e. the ${}^{I}x_t$ and ${}^{I}y_t$ coordinates. In order to do that, let:

$$C = A_{int} * {}^{C}H_I$$

then,

$$p_t = C * P_t$$

or

$$\begin{bmatrix} su_t \\ sv_t \\ s \end{bmatrix} = C * \begin{bmatrix} {}^{I}x \\ {}^{I}y \\ h \\ 1 \end{bmatrix}$$

which, after multiplying out, becomes,

$$\begin{bmatrix} (uc_{31} - c_{11}) & (uc_{32} - c_{12}) \\ (vc_{31} - c_{21}) & (vc_{32} - c_{22}) \end{bmatrix} \begin{bmatrix} {}^{I}x \\ {}^{I}y \end{bmatrix} =$$

$$= \begin{bmatrix} hc_{13} + c_{14} - uhc_{33} - uc_{34} \\ hc_{23} + c_{24} - hvc_{33} - vc_{34} \end{bmatrix}$$

and the solution for ${}^{I}x_t$ and ${}^{I}y_t$ is given by:

$$\begin{bmatrix} {}^{I}x \\ {}^{I}y \end{bmatrix} = \begin{bmatrix} (uc_{31} - c_{11}) & (uc_{32} - c_{12}) \\ (vc_{31} - c_{21}) & (vc_{32} - c_{22}) \end{bmatrix}^{-1} *$$

$$* \begin{bmatrix} hc_{13} + c_{14} - uhc_{33} - uc_{34} \\ hc_{23} + c_{24} - hvc_{33} - vc_{34} \end{bmatrix}$$

In summary, P_t is given by $\begin{bmatrix} {}^{I}x, & {}^{I}y, & h \end{bmatrix}^t$ as above.

C. Altitude Estimation

In Section II-B, we assumed that the height of the UAV – and therefore, the height of the camera – was known. In fact, many systems reported to date ([14], [13]) use an altimeter of some sort (passive or active) to provide this information. However, passive altimeters – relying for example on barometric sensor – are neither accurate nor reliable. Usually, they provide 20ft accuracy; they require adjustment before taking off; and their readings vary considerably under different weather conditions [1]. On the other hand, active sensors consume more power from the UAV and can be easily detected. Therefore, in this work we propose the use of computer vision to estimate the altitude of the UAV.

In order to estimate the altitude of the UAV, we assume that at every instant t, the system stores an image frame as well as the current pose of the UAV and camera – that is, GPS/INS data and gimbal control, if one is present. Those images are processed by a feature detection algorithm (SIFT) [11][10] and salient feature points Ps are extracted from the images. Whenever the altitude of the UAV needs to be estimated, a pair of images taken at the current instant t and at a previous instant $t - d$ are selected for the altitude estimation algorithm.

First, salient points from the two images are matched and the 3D coordinates, XYZ, of these points are reconstructed. Figure 3 illustrates this idea. Points in the vicinity of the target are regarded as being on the same plane, and therefore, having the same height as the target. Those points are used to estimate the altitude of the UAV with respect to that particular target, and consequently, the geo-location of the same target as explained in sec. II-B.
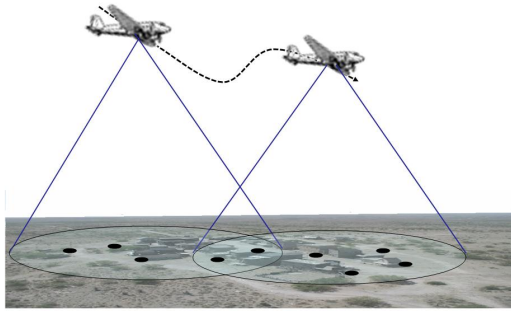
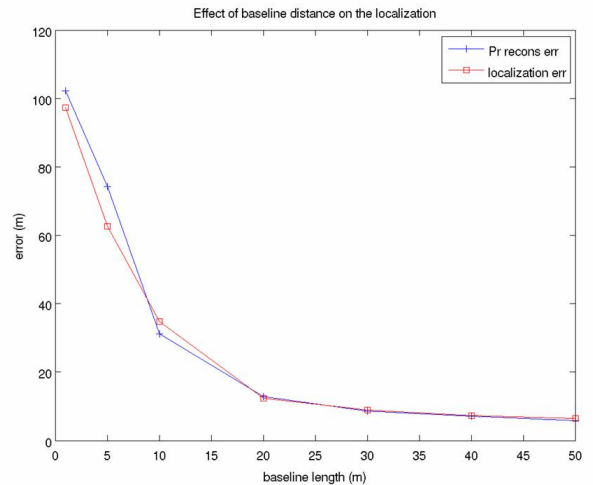Figure 3.   Overlap of the camera FOVs at two time instants

Figure 4.   effect of base line distance on the 3D reconstruction (Blue line represents the reconstruction error of salient point, and red line represents the resultant localization error of the target)

As it is well known [8], the longer the baseline (the distance between stereo cameras) the more accurate is the 3D reconstruction from the stereo vision. However, in our case, since the longer the baseline, the more likely for a feature point to disappear from the second frames, we must select a size of baseline that maximizes both constraints. That is: it maximizes the accuracy in 3D reconstruction, while it keeps a reasonably large overlap between two frames (Figure 3).

In order to find the best compromise between these two constraints, we tested the performance of the localization by changing the time interval d between the pair of images, and therefore, changing the length of the baseline. Figure 4 shows the error in reconstruction as a function of the baseline. The test was performed assuming typical uncertainties of the sensors, i.e. 1 meter in the GPS reading, 1 degree in IMU reading, and 5 pixels in target detection and tracking. In the next section, we will explain these uncertainties in more detail.

At this point, the reader may be wondering why the system does not use the same stereo-vision approach to directly calculate the position of the target. In many applications involving airborne surveillance, target detection, etc., an object of interest must be geolocated as soon as it appears in front of the camera. A typical GPS sensor has a update frequency of only 4Hz, which means that the two stereo images have to be at least that same time apart (i.e. 250ms). As proposed, our system can instantaneously geolocate any target without the need to wait for the next frame to be acquired. Besides, as we explained above, in order to increase the accuracy of the system, the *"distance"* between the two images used for 3D reconstruction must be maximized – usually requiring a longer distance than what would correspond a 250ms interval. Moreover, our method allows for instantaneous geolocation of the target even if the target itself is not present in one of the images.

III. RESULTS

The proposed algorithm has been tested for three types of data: synthetic data; in-scale data; and airborne video taken from a helicopter. As for the synthetic data, we performed a computer simulation using a typical scenario, where noise

was added to all steps of the method: image processing, sensor reading, etc. Next, the algorithm was tested with real images obtained by moving an industrial robot arm over a in-scale model city. Finally, we tested the method using an airborne image sequence, where a GPS/IMU device was attached to the camera as if on a real UAV, and another GPS device was attached to a vehicle (target) to provide ground truth for our test.

A. Simulation data and Error Analysis

In order to validate our method, we first calculated the error sensitivity to each parameter measured or calculated by the algorithm. The testing conditions were set up as in [16], [7]. That is, we assumed an UAV flying on a circular trajectory with a 100m radius, while the altitude of UAV was kept at 60 m from the ground target. In addition, as discussed in Section II-C, the camera baseline was set to 15m, or 25% of the altitude of the UAV. Finally, we added noise to each simulated reading from the sensors – the noise was made equal to the maximum nominal error of each device according to the manufacturer's specifications. The importance in this step of the analysis of our method is to establish to which sensor reading or which parameter of the method the estimated geo-location is most sensitive. That will allow us to identify which part of the algorithm should be further improved in the future.

Table I shows the localization error at the nominal flying conditions described earlier when a certain amount of noise – indicated by "Uncertainty" in the third column of the table – is added to each of the parameters separately. In other words, each row under column D_{ji} represents the total error in target geo-location as a function of the uncertainty of the ith sensor alone. By doing so, we contrasted the sensitivity to parameter i^{th} in our method, D_{2i}, with the one reported in [14] – shown in column D_{1i}. As one would expect, our method is mostly

i	Parameter	Uncertainty	D_{1i}	D_{2i}
1	Azimuthα_{az}	.5°	0.55m	N/A
2	Elevationα_{el}	.5°	0.85m	N/A
3	yaw_{uav}	5°	4m	2.635m
4	$pitch_{uav}$	5°	5.5m	7.472m
5	$roll_{uav}$	5°	8.5m	7.228m
6	X_{uav}	5m	5m	5.016m
7	Y_{uav}	5m	5m	5.001m
8	h_{uav}	5m	6.4m	3.575m
9	u	5 pix	0.75m	0.388m
10	v	5 pix	0.95m	0.374m
11	$\sqrt{\sum(D_i)^2}$		14.9m	13.351m

Table I

ERROR SENSITIVITY OF THE ALGORITHM IN [14], D_{1i}, AND THE PROPOSED METHOD, D_{2i}

Error mean	Error std
5.66 cm	4.78 cm

Table II

MEAN AND STANDARD DEVIATION IN GEO-LOCATION USING THE MODEL CITY SEQUENCE

(a)
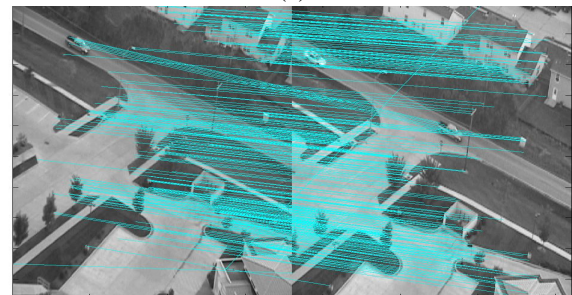
(b)

Figure 6.   a) targets located by the tracking module; and b) SIFT features matched between two images, 60 frames apart

Figure 5.   Model city used for simulation

sensitive to pitch and roll of the UAV as these parameters affect the most the calculation of the relative altitude with respect to the target, as explained in section II-C.

B. Model City Data

Before the validation with real data, the algorithm was also tested on a scaled model city by having a toy train moving on a regular path. Figure 5 shows a top view of the settings used for this test. We attached two cameras to the end effector of a KAWASAKI UX-150 industrial robot and programmed the robot to follow a certain trajectory on top of the model city, simulating an UAV flying over the target.

We then took more than 200 stereo snapshots of the moving target, whose left and right image coordinates were identified manually and used to reconstruct the precise location of the target in space. Since accurate camera and hand-eye calibrations [19], [9], [6] together with simple 3D reconstruction can be used to establish the position of the target within less than a millimeter, this data was used as ground truth for our tests. It would be very hard to scale up the results from this test to the sizes of real objects since we cannot proportionally scale the optics involved. However, just as a point of reference, for the 1:160 scale model used in this tests, we employed a baseline, according to Section II-C, of 1.6cm and a simulated altitude of the UAV (robot) of 1.2m. This is equivalent to a real UAV flying at 192m of altitude and using a baseline of 2.5m.

Finally, we applied our complete algorithm (image tracking and geo-location) to all images and compared our estimates for the target location with the ground truth obtained as above. Table II shows the results from this test: an average error in target localization of 5.66cm and standard deviation of 4.78cm. Once again, if the real errors could be inferred by simply applying the scale factor, that corresponds to an error in geo-location of under 8m, or a little over 4.5% of the altitude of the UAV.

C. Real Data - Airborne Video

In our last test, a person flew with a camera and GPS/IMU device installed in a helicopter. Meanwhile, another GPS device was set up on a ground vehicle. We estimated the GPS position of the ground vehicle captured using the helicopter video, and this positions were compared with the real GPS positions of the ground vehicle. This image sequence constitutes, as far as we know, the only database of airborne video that contains not only metadata concerning the air-vehicle, but also the target on the ground.

As a result of this experiment, the mean localization error and the standard deviation were found to be 25m and 11m, and the target distance from the observer was approximately 115m. It turned out that the performance of the proposed

Experiment	Distance to Target (m)	Baseline (m)	Error mean (m)	Error std (m)
Helicopter	115	53	25	11
Airplane	2297	900	68	35
Airplane	2174	1410	26	8

Table III
MEAN AND STANDARD DEVIATION IN GEO-LOCATION USING THE
HELICOPTER AND AIRPLANE EXPERIMENTS

algorithm with real data was well below our expectations. After further investigations of the possible causes, we noticed that the roll angle from the IMU presented a much larger (2-5 times) uncertainty than the specifications by the manufacturer. So, we run the algorithm against another dataset, this time from an airplane at an altitude of over 1000m. Since the IMU sensor inside the airplane had a much better sensor, with specifications well within our expectations, the error in geo-location and the standard deviation were also much better: 26m and 8m respectively. Those values were achieved for a distance between two images of about 840 frames (at 30fps), which for the velocity of the airplane represented 1410m of baseline. Table III summarizes these and other results obtained with real data.

IV. CONCLUSION AND FUTURE WORK

A novel algorithm capable of estimating target geo-location was developed. Not only the proposed method does away with most usual constraints – geo-referenced terrain database, altimeter, arbitrary path planning or gimbal control – but it also outperforms some other current methods in the literature, e.g. [13], [14]. We have also shown that the sensitivity of our algorithm to the various parameters is comparable, if not a little better than in [14]. As the same sensitivity analysis predicted, a key element of estimating the target position is the correct measurement of camera pose and orientation with respect to the inertial frame, in special the orientation (IMU). Another advantage of our method is in the possibility of adjusting the accuracy of the geolocation by adjusting the distance between the two images used by the algorihtm, as the last two rows in Table III demonstrate.

In the future, we plan to combine the proposed method with a Bayesian estimator, such as particle filter, to reduce the uncertainty in sensor readings and improve even further the geo-location.

REFERENCES

[1] "Kollsman servoed altimeter with automatic pressure standby – type a41322 10 015 (aau 34/a)." [Online]. Available: http://www.kollsman.com

[2] S. Baker and I. Matthews, "Lucas-kanade 20 years on: A unifying framework," *International Journal of Computer Vision*, vol. 56, pp. 221–255, 2004.

[3] D. B. Barber, J. Redding, T. W. McLain, R. W. Beard, and C. Taylor, "Vision-based target geo-location using a fixed-wing miniature air vehicle," *Journal of Intelligent Robotics Systems*, vol. 47, pp. 361–382, December 2006.

[4] P. DeLima, G. York, and D. Pack, "Localization of ground targets using a flying sensor network," in *Proceedings of the IEEE international Conference on Sensor Networks, Ubiquitous, and Trusworthy Computing*, June 2006, pp. 194–199, taichung, Taiwan.

[5] G. N. DeSouza, A. H. Jones, and A. C. Kak, "An world-independent approach for the calibration of mobile robotics active stereo heads," in *Proceedings of 2002 IEEE International Conference on Robotics and Automation*, May 2002, washington DC, USA.

[6] G. N. DeSouza and A. C. Kak, "A subsumptive, hierarchical, and distributed vision-based architecture for smart robotics," *IEEE Transactions on Systems, Man and Cybernetics - Part B*, vol. 34, no. 5, Oct. 2004.

[7] V. N. Dobrokhodov, I. I. Kaminer, K. D. Jones, and R. Ghabcheloo, "Vision-based tracking and motion estimation for moving targets using small uavs," in *Proceedings of 2006 American Control Conference*, June 2006, minneapolis, MN, USA.

[8] O. D. Faugeras, *Three-Dimensional Computer Vision : A Geometric Viewpoint*. MIT Press,, November 1993.

[9] R. Hirsh, G. N. DeSouza, and A. C. Kak, "An iterative approach to the hand-eye and base-world calibration problem," in *Proceedings of 2001 IEEE International Conference on Robotics and Automation*, vol. 1, May 2001, pp. 2171–2176, seoul, Korea.

[10] T. Kadir and M. Brady, "Sailency, scale and image description," *International Journal of Computer Vision*, vol. 45, pp. 83–105, November 2001.

[11] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, 2004.

[12] B. Lucas and T. Kanade, "An iterative image registration technique with an apllication to stereo vision," in *Proc. of 7th International Joint Conference on Artificial Intelligence(IJCAI)*, 1981, pp. 674–679.

[13] M. J. Monda, C. A. Woolsey, and C. K. Reddy, "Ground target localization and tracking in a riverine environment from a uav with a gimbaled camera," in *Proceedings of AIAA Guidance, Navigation and Control Conference*, August 2007, pp. 6747–6750, hilton Head, SC, USA.

[14] J. Redding, T. W. McLain, R. W. Beard, and C. Taylor, "Vision-based target localization from a fixed-wing miniature air vehicle," in *Proceedings of 2006 American Control Conference*, June 2006, pp. 2862–2867, minneapolis, MN, USA.

[15] R. Rysdyk, "Uav path following for constant line-of-sight," in *Proceedings of the 2nd AIAA Unmanned Unlimited Systems, Technologies and Operations Aerospace, Land and Sea conference*, September 2003, san Diego, California, USA.

[16] I. H. Whang, V. N. Dobrokhodov, I. I. Kaminer, and K. D. Jones, "On vision-based target tracking and range estimation for small uavs," in *Proceedings of AIAA Guidance, Navigation and Control Conference*, August 2005, san Franscisco, CA, USA.

[17] W. Whitacre, M. Campbell, M. Wheeler, and D. Stevenson, "Flight results from tracking ground targets using seascan uavs with gimballing cameras," in *Proceedings of 2007 American Control Conference*, July 2007, new York, NY, USA.

[18] K. C. Yap, "Incorporating target mensuration system for target motion estimation along a road using asynchronous filter," *Thesis disertation for Master of Science in Engineering Science, Naval Postgraduate School*, December 2006.

[19] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, November 2000.