*MVA '98* IAPR Workshop on Machine Vision Applications, Nov. 17-19, 1998, Makuhari, Chiba, Japan

# 8—7 Localization Based on the Gradient Information for DEM Matching

Dong-Gyu Sim*
Dept. of Electronic Eng.
Sogang University

Rae-Hong Park†
Dept. of Electronic Eng.
Sogang University

## Abstract

This paper proposes the localization algorithm that estimates a ground position by comparing the recovered elevations estimated from aerial images with digital elevation model (DEM). The proposed algorithm consists of two stages: recovering the sampled elevations from multiple aerial images and matching them with DEM. While conventional algorithms estimate the elevation field, that is, recovered elevation map (REM) over a whole image, the proposed algorithm recovers the elevations only at finite number of sample points from a multiple image sequence and does not require rotation of REM. So, the proposed algorithm can estimate the ground position accurately by using a wide recovered area and can estimate the position much faster than conventional ones. Additionally, the proposed algorithm makes use of the gradient information of terrain at multiple sample points of multiple aerial images for considering global characteristics. Computer simulations with various images show the effectiveness of the proposed algorithm.

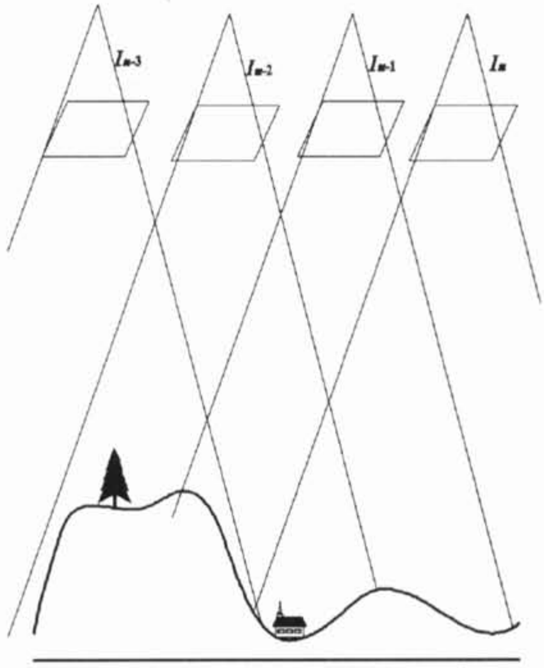



Figure 1: Area recovered by multiple images

## 1 Introduction

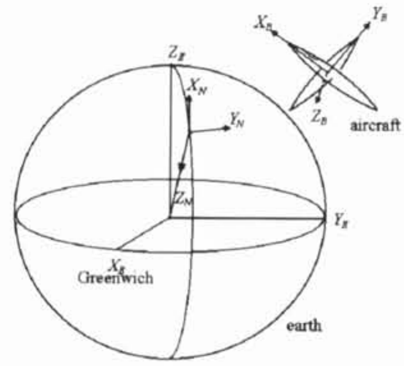
Aircraft navigation systems can be realized by a number of different approaches such as inertial navigation system (INS), global positioning system (GPS), and so on. INS has some drawbacks that the estimation error is increasing as navigation goes by. GPS can estimate an absolute position with more than three satellites, however, the signals from satellites can be disturbed and modified intentionally. Navigation parameters can also be estimated by using computer vision techniques [1][2]. A few algorithms have been proposed [1][2][3][4][5], however they are not practical for implementing real-time navigation systems because they require high computational complexity. An on-line vehicle motion estimation method from visual information was proposed based on Kalman filter estimation [1], with aerial images obtained by an expensive down-looking camera. Another algorithm with image data acquired by a down-looking camera was also proposed [6]. On the other hand, 3D reconstruction algorithms have been proposed by matching DEM and REM [2][3]. But these algorithms require high computational complexity. A fast, accurate, and cheap visual localization algorithm is necessary for practical navigation systems.

The conventional algorithms make use of only two images to reconstruct REM for localization. Ground area covered by a camera is not wide and the recovered area with two images is narrow. In this paper, the localization algorithm using a multiple aerial image sequence for DEM matching is proposed. Fig. 1 shows the area that can be recovered with multiple images. The proposed algorithm estimates the aircraft position by matching the multiple gradients obtained from multiple aerial images and DEM, where the robust estimation scheme is used.

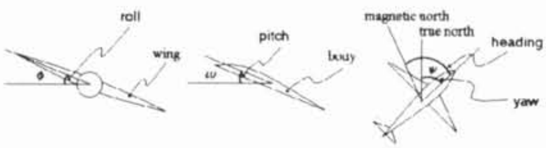
The rest of the paper is structured as follows. In Section 2, the localization algorithm using multiple images is proposed. Experimental results show the

*Address: Department of Electronic Engineering, Sogang University, C.P.O. Box 1142, Seoul 100-611, Korea. E-mail: sdg@eevision1.sogang.ac.kr.

†Address: Department of Electronic Engineering, Sogang University, C.P.O. Box 1142, Seoul 100-611, Korea. E-mail: rhpark@ccs.sogang.ac.kr.




(a) Navigation and body coordinates




(b) Attitude of an aircraft

Figure 2: Coordinates and attitude of an aircraft

effectiveness of the proposed algorithm in Section 3. Finally, conclusions are summarized in Section 4.

# 2 Proposed Localization Algorithm with DEM Information

## 2.1 Coordinate system and attitude parameters of an aircraft

In a navigation system, it is important to define a reference coordinate system. Two coordinates are considered: navigation coordinate and body coordinate [7]. The navigation coordinate $(X_N, Y_N, Z_N)$ represents a reference coordinate system as shown in Fig. 2(a), where universal transverse Mercator (UTM) as an orthogonal coordinate is used in our system. A transformation from the navigation coordinate $(X_N, Y_N, Z_N)$ into the aircraft body axis coordinate $(X_B, Y_B, Z_B)$, and its inverse transformation are necessary for navigation parameter estimation. The transformation is represented by a rotation matrix expressed as a function of $\phi$, $\omega$, and $\psi$ that denote roll, pitch, and yaw parameters of an aircraft, respectively. $\phi$ and $\omega$ denote angles of a wing and a body of an aircraft with respect to the surface of earth, respectively. $\psi$ represents an angle of a body with respect to the truth north. Note that the true north is different from the magnetic north, for example, by about $6.2^o$ as shown in Fig. 2(b), at Deajon and Kongju regions, over which one of test aerial image sequences was acquired.

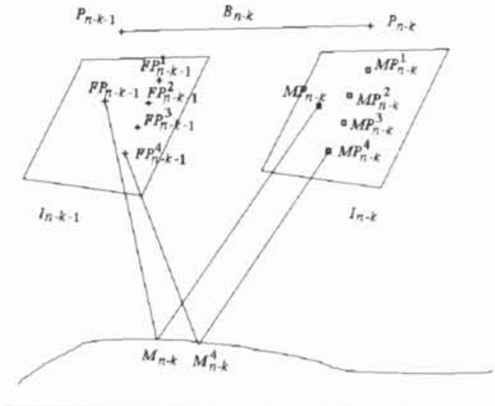



Figure 3: Features and matching points

## 2.2 Proposed localization algorithm

The proposed algorithm consists of two stages: recovering the elevations at sample points and matching them with DEM [2][3]. At the first stage, the elevations at the sample points of multiple images are calculated based on the stereo matching scheme, with correspondences between consecutive aerial image pairs. At first, the feature points that have the maximum local variance are extracted at uniform intervals in the previous image. The feature correspondence is obtained by block-based matching, with the normalized correlation coefficient (NCC) criterion measure. In Fig. 3, $FP_{n-k-1}$ denotes the feature point having a maximum variance at the $(n-k-1)$th frame in the aerial image sequence. $MP_{n-k}$ represents the point in the $(n-k)$th frame, corresponding to the feature point $FP_{n-k-1}$, where $P_{n-k-1}$ and $P_{n-k}$ denote 3D ground positions of the aircraft in the $(n-k-1)$th and $(n-k)$th frames, respectively. The difference of these two positions is denoted by $B_{n-k}$. $FP^j_{n-k-1} (j = 1, ..., M)$ is the $j$th feature in the $(n-k-1)$th frame. $MP^j_{n-k}$ of the $(n-k)$th frame $I_{n-k}$ is a corresponding point detected by the NCC measure. These matching points are detected with $N$ image pairs ($k = N - 1$ to 0). A sample REM is recovered with the consecutive feature points and matching points. That is, $M^j_{n-k}$ is calculated with $\phi$, $\omega$, $\psi$, $FP^j_{n-k-1}$, and $MP^j_{n-k}$ by the stereo matching scheme [3][4].

In the second stage, the recovered REM is normalized by dividing the REM value by that of $M_{n-k}$ which is estimated with the maximum variance point $FP_{n-k-1}$ and its matching point $MP_{n-k}$. The proposed algorithm searches the area with 5 pixels × 5 pixels interval. Assuming that the resolution of DEM is 2 $m$, the real search interval is 10 $m$×10 $m$. The DEM matching algorithm is described as follows.

$M_{in}$=Infinite
For $dx = x_0$ to $x_1$ step $\delta x$
For $dy = d_0$ to $y_1$ step $\delta y$

```
E = 0
For k = N - 1 to 0 step -1
For j = 1 to M step 1
   Grad_REM = M^j_{z,n-k} - M_{z,n-k}
   DEM^j = DEM(M^j_{x,n-k}+dx, M^j_{y,n-k}+dy)
   DEM^0 = DEM(M_{x,n-k}+dx, M_{y,n-k}+dy)
   Grad_DEM = DEM^j - DEM^0
   E+ = ρ(Grad_REM - Grad_DEM)
IF (M_in ≥ E)
   M_in = E
   Store dx, dy
X_n+ = dx
Y_n+ = dy
```

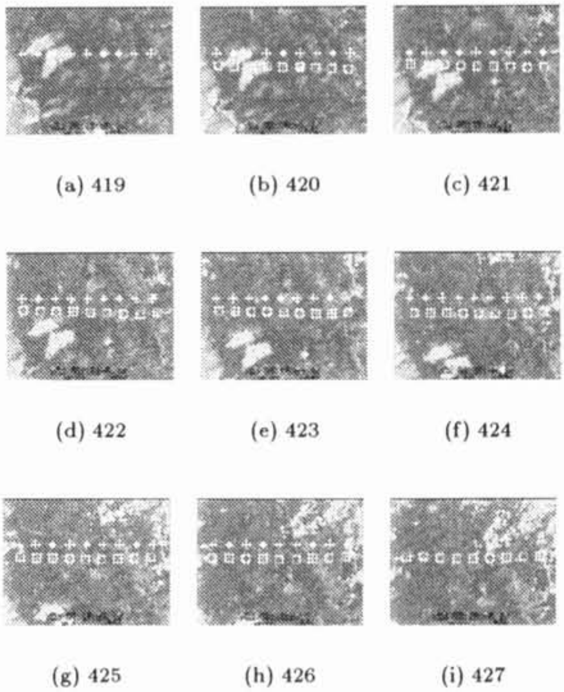
where $\delta x$ and $\delta y$ denote the search intervals along the $x$ and $y$ directions, respectively. $DEM(x, y)$ represents an elevation at 3D real position $(x, y)$. The proposed algorithm does not use the mean square error (MSE) but a robust cost function $\rho$ based on M-estimation that can effectively reject outliers. The M-estimator can estimate a stable solution of input data that contain some severe errors. Whereas the conventional algorithm based on the MSE estimator might yield wrong position because of a few wrong matching points, the proposed algorithm based on the robust estimator can estimate a correct result for severely deteriorated input data.

Whereas the conventional algorithm needs the elevation calculations as many as image size for dense REM [2][3], the proposed algorithm requires only $M \times N$ elevations. For example, assuming that image size is $320 \times 240$ and $M \times N$ is $9 \times 8 = 72$, the conventional algorithm requires elevation calculations 1067 times greater than that of the proposed algorithm. The number of additions for matching sampled REM with DEM is equal to $NM(x_1 - x_0)(y_1 - y_0)/(\delta x \delta y)$. According to computer simulations on Pentium 200, the proposed algorithm takes 5 seconds while the conventional algorithm requires 1800 seconds.

## 3 Experimental Results and Discussions

The effectiveness of the proposed algorithm is shown by the computer simulation with two aerial image sequences ($320 \times 240$) that are acquired with a CCD camera attached to a helicopter and a light airplane. The aerial image sequences are recorded with a $\beta$-cam video camera and a Hi-8m camcoder. The field of view (FOV) of the former camera is set to $42.7^o \times 54.7^o$, while that of the latter camera is set to $34.7^o \times 25.7^o$. Each of them is digitized at 1 frame/sec and quantized to eight bits.

The proposed algorithm makes use of $280 \times 200$ search area with $10 \times 10$ interval, in order to detect a feature point $FP_{n-k-1}$ that has the largest variance in a frame. The search area is set to $100 \times 100$ to detect a matching point $MP_{n-k}$ corresponding to $FP_{n-k-1}$. The proposed algorithm searches $9 \times 9$ area to detect $MP^j_{n-k}$ corresponding to $FP^j_{n-k-1}$ with the initial disparity $(MP_{n-k} - FP_{n-k-1})$. In the experiment, the window size for matching is set to $11 \times 19$, the number of features for a pair of images is 9, the number of image pairs is 8, and the resolution for search DEM is 7.5m $\times$ 7.5m. The two consecutive images are assumed to be overlapped by 50 %.



(a) 419 (b) 420 (c) 421 (d) 422 (e) 423 (f) 424 (g) 425 (h) 426 (i) 427

Figure 4: Test aerial image sequence ($320 \times 240$)

Figs. 4(a)-(i) show nine consecutive images acquired with one second interval, where the white cross (+) marks represent the feature points $FP^j_{n-k-1}$ to estimate the elevations, and the box (□) marks denote the matching points $MP^j_{n-k}$ corresponding to the feature points $FP^j_{n-k-1}$. In the case of this image sequence, the overlapped ground area with two images is about 300m $\times$ 300m. The REM recovered over such a small overlapped area does not contain many features, resulting in a large probability that the DEM matching fails. A plain surface is usually estimated over a small overlapped area, thus it is difficult to correctly match the plain surface with DEM. However, the recovered area by the proposed algorithm is wide enough to match DEM because the multiple images are used. In this experiment, 654 $\times$ 896 DEM is used, and the matching result is shown in Fig. 5, where the black dots denote the feature points $FP^j_{n-k-1}$ in Fig. 3, while the white boxs represent the real positions of the feature points. The result of the proposed algorithm is satisfactory, and real position error is 26.1m. The ground resolution of DEM is 2.6m per pixel, and ground search area is 1700.4m $\times$ 2329.6m. The DEM is gen-

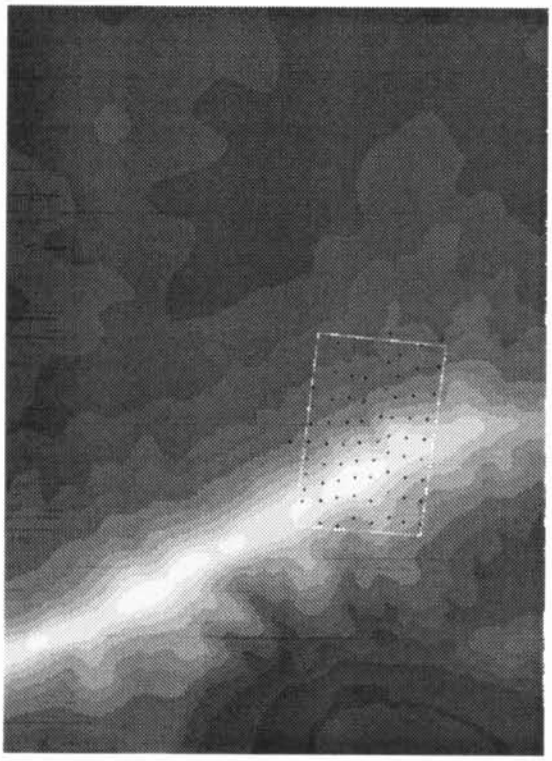


Figure 5: Matching result

erated by interpolating the contours that are manually drawn with 25m interval.

Table I shows the errors estimated by the conventional and proposed algorithms, where the estimated error is calculated in terms of the Euclidean distance between the real and estimated positions. Nine sequential images at each experiment are used with the image sequences acquired by a helicopter and a light airplane. In the cases of the first spot of the helicopter sequence and the third spot of the light airplane sequence, the conventional algorithm yields large errors because mountains are very high and large in this area. Thus, the area recovered by only two images does not contain enough features. But, the proposed algorithm makes use of nine images, and the recovered area is wide enough to detect reliable matching points. Even though these images have a few mismatch points, the proposed algorithm can estimate the correct matching points because of the robust M-estimator.

## 4 Conclusions

In this paper, a localization algorithm is proposed by comparing the elevations recovered from multiple images with DEM. The proposed algorithm uses the sampled elevations to dramatically reduce the computational complexity. The proposed algorithm can estimate wide REM with multiple images. The effectiveness of the proposed algorithm is shown via various computer simulations. Further research will focus on the development of the real-time algorithm

**TABLE I. Errors estimated by the conventional and proposed algorithms (m).**

| aircraft | Helicopter | | Light airplane | | | |
|---|---|---|---|---|---|---|
| method\spot | 1 | 2 | 1 | 2 | 3 | 4 |
| conventional | 351 | 34 | 37 | 56 | 196 | 27 |
| proposed | 26 | 29 | 23 | 37 | 36 | 31 |

and its application to a practical visual navigation system.

## 5 Acknowledgements

This work was supported in part by Agency for Defence Development and by ACRC (Automatic Control Research Center), Seoul National University.

## References

[1] Y. Bresler and S. J. Merhav, "On-line vehicle motion estimation from visual terrain information Part II: Ground velocity and position estimation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, no. 5, pp. 588-604, Sept. 1986.

[2] J. J. Rodríguez and J. K. Aggarwal, "Matching aerial images to 3-D terrain maps," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-12, no. 12, pp. 1138-1149, Dec. 1990.

[3] M. S. Kang, R.-H. Park, and K.-H. Lee, "Recovering an elevation map by stereo modeling of the aerial image sequence," *Optical Engineering*, vol. 33, no. 11, pp. 3793-3802, Nov. 1994.

[4] D.-G. Sim, S. Y. Jeong, R.-H. Park, R.-C. Kim, S.-U. Lee, and I. C. Kim, "Navigation parameter estimation from sequential aerial images," in *Proc. 1996 IEEE Int. Conf. Image Processing*, vol. II, pp. 629-632, Lausanne, Switzerland, Sept. 1996.

[5] D.-G. Sim, S.-Y. Jeong, D.-H. Lee, R.-H. Park, R.-C. Kim, S. U. Lee, and I. C. Kim, "Hybrid estimation of navigation parameters from aerial image sequence" *IEEE Trans. Image Processing*, accepted, 1998.

[6] Q. Zheng and R. Chellappa, "A computational vision approach to image registration," *IEEE Trans. Image Processing*, vol. IP-2, no. 3, pp. 311-326, July 1993.

[7] C.-F. Lin, *Advanced Control Systems Design*. Englewood Cliffs, NJ: Prentice-Hall, 1994.