International Journal of Computer Theory and Engineering Vol. 4, No. 2, April 2012

UAV Navigation Based on PIIFD/INS Method

Mostafa Kazemi, Majid Banaeyan, and Ali Pourmohammad, *Member, IACSIT*

***Abstract*—This paper investigates using a novel method for UAV guide navigation based on scene matching. In order to have accurate and reliable navigation system, inaccurate information of inertial navigation system and image coordination is used to improve the navigation parameters. The scene matching process is provided using PIIFD method. Interest point detectors are used in order to extract the most important interest point. Then, this point is compared by assigning an orientation to each one. The PIIFD method is invariant to brightness variation and rotation. The drawback of this method is being inaccurate relative to viewpoint change which is solved using inertial navigation parameters.**

***Index Terms*— UAV, PIIFD, scene matching, interest point.**

I. Introduction

Local features are mostly used in scene matching. Disadvantage of using these features is that change when scale and viewpoint angle change. Recently some solutions have been proposed for this problem. In all of them, first the image features are extracted and then based on these features, descriptors are defined. These descriptors should be invariant relative to affine transforms. Extracting features and assigning a descriptor to each feature need high processing time. The complexity decrease using some methods to eliminate transformation of viewpoint changes. In this paper, viewpoint changes transformation is eliminated by using the information of inertial navigation system. Then Partial Intensity Invariant Feature Descriptor, PIIFD, is applied on interest points to achieve invariant descriptor related to illumination and rotation.

Based on proposed method, first we transform UAV captured image to a top-down view using the information of inertial navigation system to eliminate the effects cause by tilt angle. We assume the reference image which is saved in UAV onboard computer, is in top-down view. Then the interest points are extracted from reference image and UAV captured image by using one of the interest point detectors. We must apply an invariant descriptor to illumination and rotation for each point and matching process. Hence, we apply PIIFD method which is invariant relative to transforms. The predefined flight path is stored in UAV onboard computer. Using this path and extracted point from capture image, the flight path is corrected and accuracy of navigation guide system is improved. This method is called justified coordinate of two images. In order to justified coordinate of images, coordinate axis of UAV camera must be transformed to Earth plate coordinate by considering flight status.

Manuscript received February 21, 2010; revised March 26, 2012.

The authors are with the Electrical Engineering Department, Amirkabir University of Technology, Hafez Ave., Tehran 15914, Iran (e-mails: mostafa.kazemi@gmail.com, majidbanayean@gmail.com, pourmohammad@aut.ac.ir; tel.: +98 21 64543392; fax: +98 21 66406469) .

Eliminating affine transformation caused by view angles changes, is the most time consuming process. The vision based navigation is more computationally efficient, if view angle changes could be eliminated by INS information. According to this purpose, UAV body coordinate system needs to transform to the motion geographic coordinate.

II. UAV Body Coordinate System and North-East-Down Coordinate System

The center of UAV Body Coordinate system is located at UAV gravity centre. If axis of this system is denoted by (*x, y, z*), the positive direction of *x*, *y* and *z* are located in direction of UAV motion, toward the right wing and toward the below respectively by the right hand rule [1]. This coordinate system is illustrated in Fig.1.

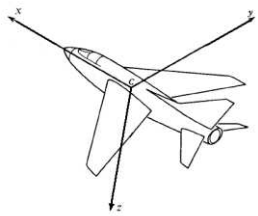

Fig. 1. UAV body coordinate

Also the center of North-East-Down Coordinate system is located at UAV gravity centre. If coordinate axis of this system is denoted by (x', y', z'), the position direction of x', y' and z' are located toward the North, East and Earth center respectively [1].

III. Transformation of UAV Body Coordinate to NED Coordinate

Usually, due to transformation of every coordinate to others, Euler angles are used which are shown in Fig.2.

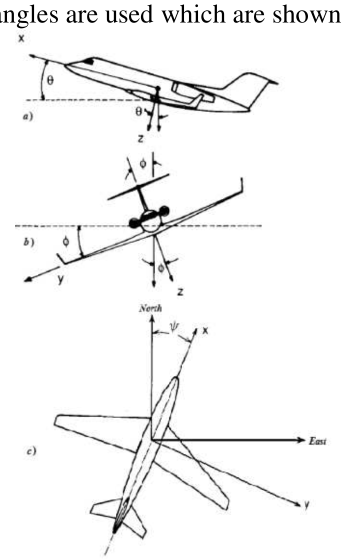

Fig. 2. Euler angles in INS.

Euler angles ψ, θ and φ are called yaw, pitch and roll angles respectively. The INS angular parameters are used to eliminate of view point changes [1]. The relation between NED coordinate and UAV body coordinate is computed as:

$$\begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} = T_{\psi} T_{\theta} T_{\varphi} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1)$$

where Tψ, Tθ and Tφ functions are computed as (2), (3) and (4) respectively.

$$T_{\psi} = \begin{bmatrix} \cos\psi & -\sin\psi & 0 & 0 \\ \sin\psi & \cos\psi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2)$$

$$T_{\theta} = \begin{bmatrix} \cos\theta & 0 & \sin\theta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\theta & 0 & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

$$T_{\varphi} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & \cos\varphi & -\sin\varphi & 0 \\ 0 & \sin\varphi & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4)$$

Altitude value is given to the UAV as others INS parameters. Suppose z_1 is altitude of reference image which is taken from with camera which has λ_1 as focal length. Also z_2 is UAV flight altitude and its camera has λ_2 as focal length. By using (5) we can eliminate scale changes [2].

$$P = \begin{bmatrix} k & 0 & 0 & 0 \\ 0 & k & 0 & 0 \\ 0 & 0 & k & 0 \\ 0 & 0 & -\frac{\lambda_1}{\lambda_2} & 1 \end{bmatrix} \quad (5)$$

where $k = z_2/z_1$. If vector (r_1, r_2, r_3) represents camera position respect to UAV gravity center, the transformation in (6) is usable for transformation of camera position to UAV gravity center [2].

$$C = \begin{bmatrix} 1 & 0 & 0 & -r_1 \\ 0 & 1 & 0 & -r_2 \\ 0 & 0 & 1 & -r_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (6)$$

As we explained, the descriptor represented in this paper is rotationally invariant. Hence, yaw angle can proposed as zero. Therefore, transform function of image is determined as:

$$\begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} = PCT_{\varphi} T_{\theta} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (7)$$

IV. PIIFD Method

First Step in scene matching is extracting information from and image. This information is regions of an image which have interest feature or specific templates. For instance, edges, blobs, distance between objects, corners and something else are features which are extracted from an image. The invariant descriptors are applied based on these features. In this paper, a descriptor which is invariant to the rotation and image intensity is used, and the others changed navigation parameters are eliminated by usage of inertial navigation system. Also the PIIFD method is applied for interest points such as corners and blobs [3]. The proposed PIIFD framework comprises the following four distinct steps:

- Detecting corner points using a Harris detector.
- Assigning a main orientation for each corner point.
- Extracting the PIIFD surrounding each corner point.
- Matching the PIIFD with bilateral matching and remove any incorrect matches.

First, corner points are extracted by Harris detector which is used as control point candidates. These points are selected due to that corner points are sufficient and uniformly distributed across the image domain. The Harris detector is one of the mostly usage of interest point detectors which is applied based on detecting image intensity changes and computing the second moment matrix surrounding each point. The second moment matrix is defined as follows [4]:

$$M = \begin{bmatrix} I_x^2 & I_x I_y \\ I_y I_x & I_y^2 \end{bmatrix} \quad (8)$$

where I_x and I_y are first order derivatives. Eigen values of the second moment matrix are used to specify type of features surrounding a selected point. Thus, the Harris detector can be algebraically express as:

$$R = \lambda_1 \lambda_2 - \alpha(\lambda_1 + \lambda_2)^2 = Det(M) - \alpha Tr^2(M) \quad (9)$$

where λ_1, λ_2 are eigen-values of M, α is a constant value (usually *0.04*), and *Det* and *Tr* are the determinant and trace of the matrix, respectively. Given a point p(x, y), it is considered as a corner point if and only if $R(p)>0$. More details about Harris detector can be found in [4].

A. Assigning Main Orientation to Each Corner Point

A main orientation is assigned to each corner point before extracting the PIIFD. Using this orientation causes achieving invariance to image rotation. In this paper, a continuous method, averaging squared gradients is used to assign the main orientation [5] [6]. This method uses the averaged perpendicular direction of gradient which is limited within $[0, \pi)$ to represent a control point candidate's main orientation. For image I, the new gradient $[G_x \quad G_y]^T$ is expressed as follow:

$$\begin{bmatrix} G_x \\ G_y \end{bmatrix} = \mathrm{sgn}(I_y) \begin{bmatrix} I_x \\ I_y \end{bmatrix} \quad (10)$$

In order to compute the main orientation, the image gradients should be averaged or accumulated within an image window. Opposite gradients are canceled each other if they are directly averaged or accumulated. But they are supposed to reinforce each other because they indicate the same main orientation. Squaring the gradient vector is a solution to this problem in complex domain before averaging. The squared gradient vector $[G_{s,x} \ \ G_{s,y}]$ is given by:

$$\begin{bmatrix} G_{s,x} \\ G_{s,y} \end{bmatrix} = \begin{bmatrix} G_x^2 - G_y^2 \\ 2G_x G_y \end{bmatrix} \tag{11}$$

Then, the average squared $\begin{bmatrix} \overline{G_{s,x}} & \overline{G_{s,y}} \end{bmatrix}^T$ gradient is computed using a Gaussian-weighted circular window

$$\begin{bmatrix} \overline{G_{s,x}} \\ \overline{G_{s,y}} \end{bmatrix} = \begin{bmatrix} G_{s,x} * h_\sigma \\ G_{s,y} * h_\sigma \end{bmatrix} \tag{12}$$

where h_σ is the Gaussian-weighted kernel, and the * is the convolution operator. The variance (σ) of the Gaussian window isn't too small or too big. Computation of the average orientation is sensitive to noise in a small window and cannot represent the local orientation in a large window. The main orientation φ of each neighborhood with $0<\varphi<\pi$ is given by:

$$\varphi = \begin{cases} \tan^{-1}(\overline{G_{s,y}} / \overline{G_{s,x}}) + \pi, & \overline{G_{s,x}} \geq 0 \\ \tan^{-1}(\overline{G_{s,y}} / \overline{G_{s,x}}) + 2\pi, & \overline{G_{s,x}} < 0 \cap \overline{G_{s,y}} \geq 0 \\ \tan^{-1}(\overline{G_{s,y}} / \overline{G_{s,x}}), & \overline{G_{s,x}} < 0 \cap \overline{G_{s,y}} < 0 \end{cases} \tag{13}$$

Hence, the main orientation φ(*x, y*) is assigned for each control point candidate p(*x, y*).

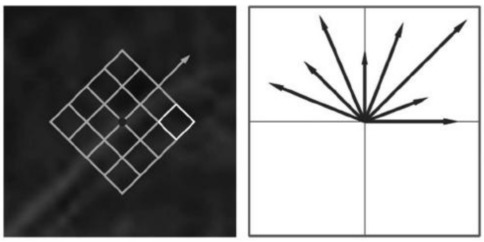

Fig. 3. Extracting PIIFD relative to main orientation of control point candidate. (a) Neighborhood surrounding the control point. (b) Orientation histogram extracted from the highlighted small square in (a).

B. Extracting the PIIFD Surrounding Each Corner Point

The local features are extracted in a manner invariant to image intensity and partially invariant to image intensity using the main orientation of each control point candidate. As shown in Fig. 3, assumed the centered point is a control point candidate and the big square which consists of 4×4 small squares is the local neighborhood surrounding his control point candidate. The main orientation of this control point candidate is illustrated by the arrow. A tradeoff between distinctiveness and computational efficiency needs to consider in order selecting the size of neighborhood. The image gradient magnitudes and orientations are sampled in this local neighborhood in order to extract the PIIFD. Also the gradient orientations are rotated relative to the main orientation in order to achieve orientation invariance. For a given small square in this neighborhood e.g., the highlighted small square shown in Fig. 3(a), an orientation histogram, which evenly covers 0-360 with 16 bins (0°, 22.5°, 45°… 337.5°) is formed. The gradient magnitude of each pixel that falls into this small square is accumulated to the corresponding histogram entry. It is important to avoid the boundary affects in which the descriptor abruptly changes as a sample shifts smoothly from being within one histogram to another or from one orientation to another. Therefore, bilinear interpolation is used to distribute the value of each gradient sample into adjacent histogram bins. The processes between extracting PIIFD and SIFT are almost the same, therefore, PIIFD and SIFT have some common characteristics. For example, both PIIFD and SIFT are partially invariant to affine transformation [7].

An outline is a line marking the multiple contours or boundaries of an object or a figure in an image. The basic idea of achieving partial intensity invariance involves extracting the descriptor from the image outlines. In this paper, image outline extraction is simplified to extract the constrained image gradients. The gradient orientations at corresponding locations in multimodal images may possibly point to opposite directions and the gradient magnitudes usually change. Two operations are applied on the image gradients in order to achieving partial intensity invariance. First, the gradient magnitudes are normalized piecewise to reduce the influence of change of gradient magnitude.

In a neighborhood surrounding each control point candidate, the first 20% strongest gradient magnitudes are normalized to 1, second 20% to 0.75, and by parity of reasoning the last 20% to 0. Second, the orientation histogram with 16 bins is converted to a degraded orientation histogram with only 8 bins (0°, 22.5°, 45°, …, 157.5°) by calculating the sum of the opposite directions [see Fig. 3(b)]. If the intensities of this local neighborhood change between two image modalities (for instance, some dark vessels become bright), then the gradients in this area will also change. However, the outlines of this area will almost remain unchanged. The degraded orientation histogram constrains the gradient orientation from 0 to π, and then the histogram achieves invariance when the gradient orientation rotates by 180. Consequently, the descriptor achieves partial invariance to the aforementioned intensity change. The second operation is based on the assumption that the gradient orientations at corresponding locations in multimodal images point to the same direction or opposite directions. It is difficult to mathematically prove this assumption as "multimodal image" is not a well-defined notation, although for intensity inverse images (an ideal situation), this assumption is absolutely sustainable. Actually, the degraded orientation histogram is not as distinctive as the original one, but this degradation at the cost of distinctiveness is acceptable for achieving partial invariance to image intensity. For the case shown in Fig. 3, there are in total 4×4 = 16 orientation histograms (one for each small square). All these histograms can be denoted by:

$$H = \begin{bmatrix} H_{11} & H_{12} & H_{13} & H_{14} \\ H_{21} & H_{22} & H_{23} & H_{24} \\ H_{31} & H_{32} & H_{33} & H_{34} \\ H_{41} & H_{42} & H_{43} & H_{44} \end{bmatrix} \tag{14}$$

where H_{ij} denotes an orientation histogram with eight bins. The main orientations of corresponding control points may point to the opposite directions in multimodal image pair.

This situation will still occur even we have already constrained the gradient orientations to the range [0° 180°], and break the rotation invariance. For example, the main orientations of corresponding control points extracted from an image and its rotated version by 180° always point to the opposite directions. In this paper, a linear combination of two subdescriptors is proposed to solve this problem. One subdescriptor is the matrix H computed by (14). The other subdescriptor is a rotated version of H: H: $Q = \mathrm{rot}(H, 180^{\circ})$. The combined descriptor, PIIFD, can be calculated as follows:

$$des = \begin{bmatrix} (H_1 + Q_1) \\ (H_2 + Q_2) \\ c\left|H_3 - Q_3\right| \\ c\left|H_4 - Q_4\right| \end{bmatrix} \tag{15}$$

$$H_i = \begin{bmatrix} H_{i1} & H_{i2} & H_{i3} & H_{i3} \end{bmatrix} \tag{16}$$

$$Q_i = \begin{bmatrix} Q_{i1} & Q_{i2} & Q_{i3} & Q_{i3} \end{bmatrix} \tag{17}$$

where c is used to tuning the proportion of magnitude in this local descriptor. It is obvious that PIIFD is a 4×4×8 matrix. For the convenience of matching, it is quantized to a vector with 128 elements. Finally, the PIIFD is normalized to a unit length.

C. Matching the PIIFD with Bilateral Matching and Remove Any Incorrect Matches

The best-bin-first (BBF) algorithm [8] is used to matching the correspondences between two images. This algorithm identifies the approximation of closest neighbors of points in high dimensional spaces. This is approximation in the sense that it returns the closest neighbor with the highest probability.

Even the bilateral BBF algorithm cannot guarantee that all matches are correct. Fortunately, it is easy to excluding the incorrect matches using the control point candidates' main orientations of matched control point candidates. Suppose there are k bilateral matches in total, $m(p_{11}, p_{21})$, . . . , $m(p_{1k}, p_{2k})$,where p_{1i} denotes the control point candidate for extracting feature descriptor f_{1i} in image I_1, and p_{2i} denotes the corresponding control point candidate in I_2. It is obvious that all differences between any p'_{1i} is and p'_{2i} is orientations are almost the same. If one of these differences of orientations is much bigger or smaller than the others, then this matching is definitely incorrect. Most incorrect matches are actually excluded according to this criterion.

V. Experimental Results

More methods are used to guidance of UAVs. In this section, we evaluate efficiency of PIIFD method and compare it with SIFT to propose a novel method for navigation guide systems. The PIIFD descriptor which is applied in this paper is not invariant relative to wide range in scale's changes. However, The PIIFD descriptor is considered invariant in scale range between 0 and 1.8 which illustrated in Fig. 4.

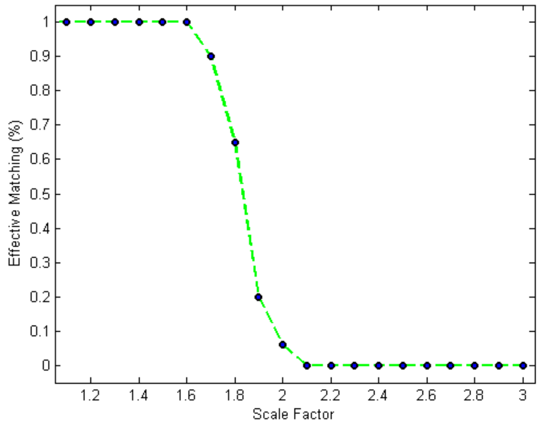

Fig. 4. Effective Matching relative to scale factor

We change the φ and θ angels simultaneously in 0 to 60 degrees interval as viewpoint change in order to evaluate the method stability. The original and rotated replica of scene with different φ and θ angles are illustrated in Fig. 4. We propose the percent of matched point in this method and SIFT method as a comparing criterion. The difference between Proposed method and SIFT algorithm is illustrated in Fig. 6. By increasing the viewpoint angle over 30 degree, the number of matching points which are detected with PIIFD are not reliable.

Fig. 5. Same with different view point

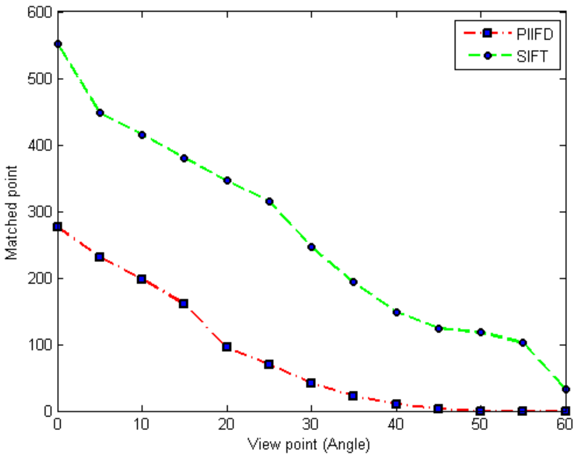

Fig. 6. Number of matched points relative to viewpoint

However these changes in φ and θ angles results 30% error in parameters of inertial navigation system. But, it is obviously that the nominal error in inertial navigation system is about 10% to 15% which is indicate the changes in angels is not more than 20 degree. Although the SIFT method has better results than the PIFT method, but it consumes more time than PIFD method which is because of using spatial-scale hybrid domain to define a descriptor and giving a direction to each point in this domain.

VI. Conclusion

In this paper a PIIFD method is used to navigate UAVs which has very good computational results. This descriptor covers a vast area of brightness variation and using the inertial navigation system information introduces a useful navigation system based on visual features. The proposed method not only has less computational complexity in comparison to other defined methods for navigating UAVs but also introduces better results in comparison to them.

References

[1] Jan Roskam. "Airplane Flight Dynamics and Automated Flight Controls," *Darcorporation,* 2001.

[2] R. C. Gonzalez and R. E. Woods "Digital Image Processing." *Massachusetts: Addison-Wesley*, 1992.

[3] J. Chen, J. Tian, N. Lee, J. Zheng, and T. Smith, "A Partial Intensity Invariant Feature Descriptor for Multimodal Retinal Image Registration," IEEE *Trans. on Biomedical Engineering,* vol. 57, no. 7, pp. 1707 – 1718, June 2010.

[4] C. Harris and M. Stephens. "A Combined Corner and Edge Detector," *In Proceedings of 4th Alvey Vision Conference*, pp. 147-151, August 1988.

[5] A. M. Bazen and S. H. Gerez, "Systematic methods for the computation of the directional fields and singular points of fingerprints," *IEEE Trans. on Pattern Anal. Mach. Intell*., vol. 24, no. 7, pp. 905–919, Jul. 2002.

[6] L. Hong, Y. Wan, and A. Jain, "Fingerprint image enhancement: Algorithm and performance evaluation," *IEEE Trans. on Pattern Anal. Mach. Intell.* vol. 20, no. 8, pp. 777–789, Aug. 1998.

[7] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis*., vol. 60, no. 2, pp. 91–110, 2004.

[8] J. Beis and D.G. Lowe, "Shape indexing using approximate nearest neighbor search in high dimensional spaces," in *Proc. Conf.Comput.Vis. Pattern Recognit.* Puerto Rico, pp. 1000–1006, 1997.

Ali Pourmohammad was born in Azerbaijan. He has a Ph.D. in Electrical Engineering (Signal Processing) from the Electrical Engineering Department, Amirkabir University of Technology, Tehran, Iran. He also teaches several courses (C++ programming, multimedia systems, Microprocessor Systems, digital audio processing, digital image processing and digital signal processing I & II). His research interests include digital signal processing and applications (audio and speech processing and applications, digital image processing and applications, sound source localization, sound source separation (determined and under-determined blind source separation (BSS), Audio, Image and Video Coding, Scene Matching and ...) and multimedia and applications.