Aerial Image Registration for Tracking

Miss. Ketaki Narendra Patil[1], Miss. Vaishali Londhe[2]

[1]*P.G. Scholar Y.T.I.E.T., Karjat, Mumbai, India*
Ketakipatil39@gmail.com.
[2]*Assistant Professor Y.T.I.E.T., Karjat, Mumbai, India*
*Vaishali.*londhe@tasgaonkartech.com

***

Abstract: *For detecting the tracking of moving target in video or in image, the relation between video and scene is fixed by registering the video frames at ground level. When camera capturing video is fixed with respect to scene is fixed then detected motion is target motion. Detection and tracking of objects of interest has traditionally been an area of interest in the computer vision literature. This paper describes a framework for feature evaluation and fusion in aerial image tracking. In these we describe a set of efficient descriptors suitable for small sized targets in aerial video or in image based on intensity, feature extraction. Image registration is basic step used in tracking applications. To separate object motion from camera motion in an aerial video, the corresponding frames are registered in planar background and feature point registration is used for tracking the object.*

Keywords-Aerial, video, Image registration.

1. INTRODUCTION

1.1Introduction and Motivation

Image Registration is a process, which finds the location where optimal matching is obtained by matching a template image called the reference image over the searching region of an input image using a suitable similarity measures. The method although computationally intensive, is simple, straightforward and robust and requires no apriori information about the two images. In these the video images are obtained from an aircraft, which is an Unmanned Aerial Vehicle (UAV) with on-board guidance and navigation system. The aircraft carries a camera in a gimbal which acquires images of the territory and sends the information to a Ground Control Station (GCS) in real time. During flight, the pilot in the Ground Control Station may identify a region of interest with a mouse click on the real time video where a target is present. The target which appears on a small window could be tracked by engaging track mode. The position of the target at all subsequent frames are identified by the window and the information is transmitted to the onboard gyro stabilization platform (platform on which the camera is mounted) to correct the azimuth and elevation of the camera so that the target appears at the center of the frame throughout tracking.

To Detect and track moving targets in a video captured by a moving camera, frames in the video are registered to keep the relation between the camera and the scene fixed. Registration of images from a moving camera is challenging particularly when 3-D structures, such as buildings, are present in the scene. Scene points visible in one view may be occluded from another view by the structures. Image registration is the process of estimating a transformation to overlay and align two images of the same scene taken at different times, from different viewpoints, and by different sensors. Image registration is a crucial step in diverse fields such as weather forecasting, integrating information into geographic information systems (GIS), medicine, cartography, computer vision, etc. The general automatic registration process consists of four steps: feature detection, feature description, feature matching, and transformation estimation. The first two steps are to extract a certain type of feature from each image (feature detection), and to describe each feature in a quantitative manner (feature description). Then, using the detected features (control points, CPs) and their descriptions, the correspondence matching (control point pairs, CPPs) information is extracted by comparing how similar the feature descriptions are (feature matching). Therefore, a good feature description will retain the same values after a registration transformation is applied. Finally, transformation can be calculated from the correspondence matching information. Since the matching information (CPPs) usually includes outliers, many methodologies use additional information such as spatial constraints to filter those outliers before calculating the transformation.

1.2 Steps involved in Image Registration

- **Feature Detection**: Salient and distinctive objects (closed-boundary regions, edges, contours, line intersections, corners, etc) in both reference and sensed images are detected.

- **Feature Matching**: The correspondence between the features in the reference and sensed image established.

- **Transform Model Estimation**: The type and parameters of the so-called mapping functions, aligning the sensed image with the reference image, are estimated.

- **Image Resampling and Transformation**: The sensed image is transformed by means of the mapping functions.

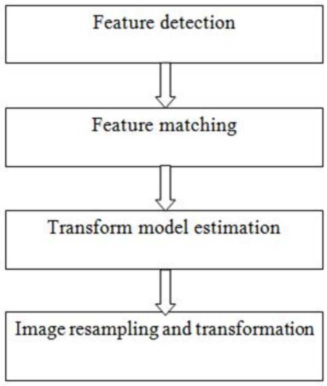

Fig 1.1: Steps involved in Image Registration

2. PROPOSED SYSTEM

In the proposed image registration First, a global homography is found to approximately align the images. This is achieved by reducing the resolution of the images sufficiently so that local geometric differences between the images become small and negligible. After registering the images at the lowest resolution, the resolution of the images is increased by a factor of two while dividing each image into four equal blocks. A local homography is then calculated to register corresponding blocks in the images. The resolution of the images is increased by a factor of two in this manner while dividing each image block into four new blocks until the blocks in the highest resolution Images are obtained. From the global homography obtained at the lowest resolution and local homographies obtained at the intermediate and highest resolutions, an overall transformation is calculated to register the images.

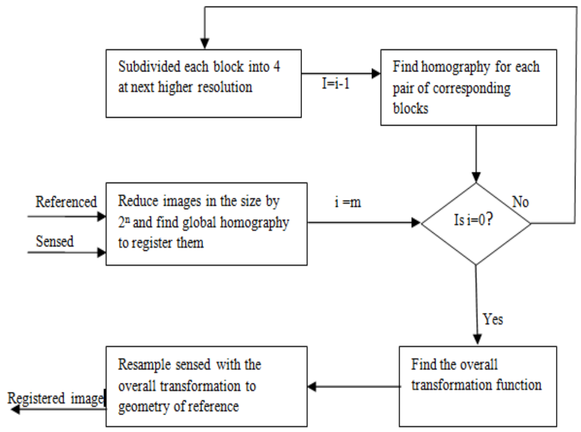

Fig 2.1: Flow of computations in image registration

3. METHDOLOGY

3.1 The methods used for Aerial Image Registration

3.1.1. Finding the Global Homography

To register images from different views of a 3-D scene, a combination of global and local homographies is used. The global homography account for the viewangle difference between the camera capturing the images and is independent of the scene content. The local homographies are scene dependent and account for local geometric differences between the images captured from different views of the scene. To estimate the global homography, local geometric differences between the images are reduced by reducing the resolution (size) of the images. We take the aerial images of urban and suburban scenes of 30-cm/pixel resolution suggest that blocks of 128 × 128 pixels contain sufficient corresponding landmarks to estimate homography parameters to register corresponding blocks in the images. Therefore, if D is the diameter of the reference image, the images must be reduced in size by a factor of 2m, with m to satisfy. If four or more corresponding landmarks in the images to find them. To have a fully automatic method, it is required that four or more corresponding landmarks be determined in the images. Landmark correspondence is achieved in two steps. First, a set of landmarks is detected in each image, and then, correspondence is established between the two sets of landmarks it can be uses the steps are:

1) Landmark Detection

Landmarks are locally unique points in images, facilitating the search for corresponding points in the images. A great number of landmark detection methods have been developed throughout the years. We are interested in detectors that are independent of the orientation of an image. Two widely used detectors are Harris corners and the scale-invariant feature transform (SIFT) points. The former detects corners, and the latter detects centers of blobs. Depending on the type of images available, one may use one or the other detector. If the images contain buildings and other man-made structures where corners are abundant, the corner detector is more appropriate, while if the images represent terrain scenes where man-made structures are absent, the blob detector is more appropriate. If no information about the images is available, one may use one detector as the default and, if that fails, use the alternate detector. In our work, since the focus is on images of urban and suburban scenes, the Harris detector is used as the default, and the SIFT detector is used as the alternate.

2) Landmark Features

Once unique landmarks are detected in each image, there is a need to find correspondence between the landmarks. To expedite the correspondence process, a number of features are extracted from the surroundings of each landmark. The vector of features associated with a landmark is known as a descriptor. The SIFT landmark detector, as well as many other detectors, has associating descriptors that characterize various properties of the neighbourhood of a landmark. The SIFT detector associates a vector of 128 features, describing the intensity gradient properties of the neighbourhood of a landmark. If the Harris corner detector is used, features must be calculated for each corner. The features used are as follows: 1) a Huinvariant moment of order two 2) a Huinvariant moment of order three 3) a complex moment of order two 4) The Laplacian of Gaussian of standard deviation of 3 pixels 5) the center contrast feature at the landmark. Therefore, in our work, the landmarks detected by both Harris and SIFT detectors use these invariant features when determining the correspondence between the landmarks. The algorithm to establish correspondence between two sets of landmarks using these features.

3) Correspondence

As Landmark the resolution of the images is reduced, local geometric differences between the images decrease while the global relation between the images remains the same. The coarse-to-fine strategy is therefore, to reduce local geometric differences between the images sufficiently so that the images can be globally registered with a homography. Once this global homography is determined, the resolution of the images can be gradually increased, and the neighborhoods can be tracked from low to high resolution.The correspondences obtained at a resolution are used to refine the transformations obtained at one level lower resolution.

3.2 Feature based Registration

The feature based techniques begin with extraction of feature points on the images. These points can be found and aligned on the two images. This yields a points sets on one image which is aligned to point set S' on the other image. By fitting a transformation to these points the transformation matrix can be estimated. Popular feature point detectors are the Harris corner point detector. SIFT detector and the SURF detector. In case of aerial images the transformation cannot be restricted to translation or rotation, thus the more general affine or perspective transformation has to be used. Therefore, we use the feature based method to find the homography matrix of the perspective transformation.

3.2.1 Feature Point

We use the Harris corner detector which is more suitable in manmade environments where corners are abundant. It is also computationally less expensive than the other feature point detectors such as SIFT or SURF. Note that if there are no feature points, like in large homogenous areas, the registration fails.

3.2.2 Point Alignment

Corresponding points are searched on two consecutive frames by the Lucas-Kanade pyramidal optical flow algorithm. This step yields the positions of the feature points on the next image, thus the transformation between the frames can be calculated. The applied optical-flow algorithm assumes small displacement and nearly constant pixel intensity across the consecutive frames.

3.3 Target Tracking

The goal of tracking is to track all the detected foreground regions as long as they remain visible in the field of view of the camera. The output of this module consists of trajectories that depict the motion of the target objects. In case of UAV videos several tracking options are available. One can perform tracking in a global mosaic or opt for tracking using geographical locations of the objects. Tracking in geographical locations is often called geo-spatial tracking and requires sensor modeling.

4.CONCLUSION

In these we register images at the ground level using the homography constraint in landmark correspondences, thereby avoiding landmarks that belong to buildings and moving targets when determining the registration parameters. It defines an overall transformation in terms of a global transformation that depends on the camera positions with respect to the scene and a number of local transformations that characterize the contents of the scene. It registers images captured with different camera orientations by using rotationally invariant landmark features. It registers images at the ground level even when the ground changes in slope and elevation. This is achieved by using different homographies to register different image neighborhoods/blocks. As long as the images cover a continuous and smooth ground and there are sufficient correspondences at ground level, it can find a continuous and smooth transformation to register the images at ground level. The ability to register images at ground level makes it possible to remove camera motion and keep only target motion in registered images. Limited ground areas visible to both images limit usage of the method. If sufficient ground areas do not appear in the images due to occlusion or lack of contrast, sufficient correspondences required determining the global homography may not be possible to find, failing the registration.

REFERENCES

[1] P. Cheng and C.-H.Menq, “Real-time continuous image registration enabling ultraprecise 2-D motion tracking,” IEEE Trans. Image Process., vol. 22, no.5, pp. 2081–2090, May 2013.

[2] K. Jackovitz, V. Asari, E. Balster, J. Vasquez, and P. Hytla, “Registration of region of interest for object tracking applications in wide area motion imagery,” in Proc. AIPR, 2012.

[3] Z. Liu, J. An, and Y. Jing, “A simple and robust feature point matching algorithm based on restricted spatial order constraints for aerial image registration,” IEEE Trans. Geosci. Remote Sens., vol. 50, no. 2, pp. 514 527, Feb.2012.

[4] H. Gonc¸alves, L. Corte-Real, and J. A. Gonc¸alves, “Automatic image registration through image segmentation and SIFT,” IEEE Trans. Geosci. Remote Sens., vol. 49, no. 7, pp. 2589–2600, Jul. 2011.

[5] T. Kim and Y.-J. Im, “Automatic satellite image registration by combination of matching and random sample consensus,” IEEE Trans. Geosci. Remote Sens., vol. 41, no. 5, pp. 1111–1117, May 2011.

[6] X. Fan, H. Rhody, and E. Saber, “A spatial-feature-enhanced MMI algorithm for multimodal airborne image registration,” IEEE Trans. Geosci. Remote Sens., vol. 48, no. 6, pp. 2580–2589, Jun. 2010.

[7] J.Ma, J. C.-W. Chan, and F. Canters, “Fully automatic subpixel image registration of multiangle CHRIS/Proba data,” IEEE Trans. Geosci. Remote Sens., vol. 48, no. 7, pp. 2829–2839, Jul. 2010.

[8] [8] X. Mei and F. Porikli, “Joint tracking and video registration by factorial hidden Markov models,” in Proc. IEEE Int. Conf. Acoust., Speech Signal Process., 2008.

[9] A.Wong and D. A. Clausi, “ARRSI: Automatic registration of remote sensing images,” IEEE Trans. Geosci. Remote Sens., vol. 45, no. 5, pp. 1483–1493, May 2007.

Miss. Ketaki N. Patil Received the B.E. Degree from the Department of Computer Engineering affiliated to Pune Univeristy, in 2013. She is currently pursuing M.E. in Computer Engineering degree under the guidance of Prof. Vaishali Londhe.