PARTICLE SWARM OPTIMIZATION BASED APPROACH TO ESTIMATE EPIPOLAR GEOMETRY FOR REMOTELY SENSED STEREO IMAGES

Manimala Mahato [1,*], Shirishkumar Gedam [1]

[1] Centre of Studies in Resources Engineering, Indian Institute of Technology Bombay, India – (manimala.mahato, shirish)@iitb.ac.in

Commission V, WG V/5

KEY WORDS: Fundamental Matrix, Particle Swarm Optimization, Epipolar Geometry, Remote Sensing, Random Sample Consensus, Stereo Vision

ABSTRACT:

A novel particle swarm optimization based approach for the estimation of epipolar geometry for remotely sensed images is proposed and implemented in this work. In stereo vision, epipolar geometry is described using 3 x 3 fundamental matrix and is used as a validation tool to assess the accuracy of the stereo correspondences. The validation is performed by enforcing the geometrical constraint of stereo images on the two perspective projections of a point in the scene for finding inliers. In the proposed method, the steps of particle swarm optimization such as the initialization of the position and velocity of the particles, the objective function to compute the best position found by the swarm as well as by each particle experienced so far, the updating rule of velocity for the improvement of the position of each particle, is designed and implemented to estimate the fundamental matrix. To demonstrate the effectiveness of the proposed approach, the results are obtained on a pair of remotely sensed stereo image. A comparison of the result obtained using the proposed algorithm with RANSAC algorithm is carried out. The comparison shows that, the proposed method is effective to estimate robust fundamental matrix by giving improved number of inliers than RANSAC.

1. INTRODUCTION

One of the most challenging problems in the area of computer vision and computer graphics is to find the geometrical constraint available between the two images of a stereo image pair irrespective of the specific objects in the scene. In stereo, the images capture different view of the same scene and are related by the epipolar constraint which is expressed mathematically by 3 X 3 fundamental matrix. The estimation of epipolar geometry from the stereo images has received a large attention and has become a core research area in the last two decades due to its enormous applications such as reconstruction, stereo analysis, camera self-calibration, motion segmentation, etc. The accurate fundamental matrix is computed using the parameters of the stereo camera. However, the complexity in estimation of fundamental matrix increases in case of remotely sensed images as the camera parameters are unknown. In this case, the most effective way of estimating the epipolar geometry is through the analysis of the stereo correspondence points (Longuet-Higgins 1981), (Xu and Zhang 1996). The computation of stereo correspondence points is extremely challenging due to the presence of noise, occlusion, and discontinuity, geometric and radiometric distortion in the stereo image pair. In case of remotely sensed images, the scenario becomes more complicated. The accuracy of the epipolar geometry of stereo image pair depends on the accuracy and density of the stereo correspondence points. Stereo correspondences are obtained using feature matching algorithm (Joglekar, Gedam, and Krishna Mohan 2014) which are divided into four steps: i) The detection of interest points in the left image and right image of the stereo image pair; ii) A feature descriptor is assigned by analyzing the neighbourhood pixels; iii) The matching of the conjugate feature points from left image to right image; iv) Pruning of correspondence points using consistency property such as left-right consistency. Feature matching algorithms estimates accurate but sparse correspondence points.

Some well known robust methods for fundamental matrix estimation are M-estimator, Least Median of Square regression LMedS, Random Sample Consensus (RANSAC) (Fischler and Bolles 1981) etc. In the literature, the estimation of fundamental matrix is optimized based on random sampling of stereo correspondence points. This is the basis of almost all highly robust estimators. These methods have in-built mechanism to reduce the influence of outliers. However, in case of remotely sensed images, the correspondence points may be inevitably corrupted by noise and outliers, such as false matches and badly located points due to occlusion, geometric and radiometric distortions. Moreover, if the correspondence points do not belong to different depth planes, the estimated fundamental matrix with the use of such correspondence points may not be able to represent the accurate epipolar geometry of the stereo image pair. However, RANSAC works poorly when outlier proportion is higher than half of the total number of stereo correspondence points used in the process of optimization. Hence, there is a need of a robust algorithm to estimate the fundamental matrix for which the performance of the algorithm should not be affected significantly due to outliers and noise. The aim of the proposed method is to use particle swarm optimization (PSO) algorithm to compute the epipolar geometry of the stereo image pair by solving the above mentioned issues. PSO is based on the behaviour of bird flocking for searching food. In this work, the estimation of fundamental matrix, is considered as an optimization problem and is solved using particle swarm optimization strategy by evolving the swarm through iterations.

---

[*] Corresponding Author

PSO algorithm was invented by Eberhart and Kennedy (Eberhart and Kennedy 1995) as part of a sociocognitive study while investigating the notion of collective intelligence in the graceful motion of swarm of birds. There are several reasons due to which particle swarm optimization (PSO) is one of the most popular swarm intelligence techniques for continuous optimization problems (Gong et al. 2014). PSO converges very fast toward optimal solution, and is simple and efficient. PSO has less number of tuning parameters which makes it easy to implement. PSO algorithm simulates the intelligence and the ability of flocks of birds, schools of fish and herds of animals to adapt to their environment by finding the rich sources of food and avoiding the predators using the "information sharing" mechanism. The set of randomly generated solutions which is represented as initial swarm moves in the design space and converges to the optimal solution through a number of iteration based on the large amount of information about the design space which is shared by all the members of the swarm.

2. PROPOSED METHODOLOGY

2.1 Epipolar Geometry of stereo image pair

Epipolar line is the intersection line of two geometrical planes. The first plane is determined by the optical center of stereo camera and the point in the scene. The other plane is the photographic plane of camera. Hence, the stereo correspondence points are positioned on the epipolar line of the stereo image pair. Therefore, the epipolar line is very useful to assess the accuracy of the correspondence points. The mathematical equation of the epipolar constraint consists of the fundamental matrix and the homogeneous coordinate of the correspondence points. Considering, (u_1, u_2) is a correspondence matching pair in which u_1 is a point in the left image and u_2 is the correspondence of u_1 in the right image. Therefore, the epipolar equation is

$$u_2^T \ F \ u_1 = 0 \qquad (1)$$

where F is the fundamental matrix.

2.2 Fundamental matrix estimation using Particle Swarm Optimization algorithm

First, a brief introduction of meanings and details of PSO algorithm is given. Each particle in the swarm has two properties, position and velocity. During the process of optimization, velocity and position of each particle is updated through the swarm's search experience and its own search experience. The swarm's search experience is represented by the best position of the swarm found so far and denoted as *Gbest*, and each particle's search experience is the best position of that particle found so far and denoted as *Pbest*. In order to define *Gbest* and *Pbest*, the position of each particle is evaluated by a fitness function.

In the proposed method, the steps of particle swarm optimization such as the initialization of the position and velocity of the particles, the objective function to compute the best position found by the swarm as well as by each particle experienced so far, the updation of velocity for the improvement of the position of each particle, is designed and implemented to estimate the fundamental matrix. The position of each particle in the swarm has a 2D structure, of size 3 x 3, representing the

fundamental matrix for the input stereo image pair. To solve the complex real-world optimization problem of fundamental matrix estimation, improved initialization of swarm is used instead of randomly selected swarm, because it leads to more accurate and faster convergence. The proposed approach is divided into two parts. First part is the initialization of the swarm needed for the particle swarm algorithm. The second part is the improvement of the velocity and position of each particle through iterations using the whole swarm's search experience and the particle's search experience for finding the optimal solution.

2.3 Particle Initialization

To initialize the swarm, the set of stereo correspondence points are found using feature matching algorithm. In order to remove the false matches the correspondences are pruned by enforcing bidirectional constraints. Image features are with special properties and structural significance. Some examples of image features are edges, corners, image gradients etc. The important properties of any image feature are invariance, detectability, interpretability, and accuracy. The invariance property of feature extractor ensures that the same feature will get detected in pair of stereo images even under different transformations (geometric and radiometric). Detectors are used to locate the image features known as interest point. Further, a descriptor is assigned to each interest point using the information about neighbouring pixels.

Given a stereo image pair I_L and I_R of size M X N, in which each pixel $X = (x, y)$ is within the image domain Ω. 128D SIFT feature descriptor (Lowe 2004) is extracted to characterize local image structures and encode contextual information by analyzing each pixel with respect to the neighbourhood pixels in terms of intensity variation, gradient variation, histogram of magnitude, gradient, and direction as per the steps of assigning a SIFT feature descriptor (Lowe 2004). With various measures it is demonstrated that SIFT descriptors outperform other feature descriptors (Mikolajczyk and Schmid 2005).

In order to detect SIFT feature point a pyramid of images with different scales of Gaussian function is constructed. Difference of Gaussian images is computed at different scales of Gaussian smoothed images. Further, the interest points which are invariant to scale and orientation in scale-space are detected. For every interest points in the image, its 16 x 16 neighbourhood is divided into 4 x 4 pixels array. Gradient magnitude and orientation are computed for each feature points as per SIFT feature descriptor algorithm (Lowe 2004). SIFT feature descriptor of vector size 4 x 4 x 8=128 is obtained. SIFT features can describe both the shape and the orientation of every pixel; thus, rotation variances between two images can be taken care. The SIFT feature points in the left image are matched with the SIFT feature points in the right image using nearest neighbour feature matching algorithm. Therefore, a set of stereo correspondence points are obtained. The correspondence points are pruned using the bidirectional left-right consistent property and are used to initialize each particle in the swarm.

Each particle is represented by its position in the swarm. The encoding scheme of the particles is application dependent and fundamental to PSO algorithm. The position of particles is encoded directly by the fundamental matrix and represented as a 3 X 3 matrix. The initial swarm is filled by computing the position of each particle as the fundamental matrix estimated from the set of pruned correspondence points using RANSAC

algorithm (Fischler and Bolles 1981). The same set of stereo correspondence points generates different fundamental matrices by the several use of RANSAC algorithm due to random sampling.

2.4 Particle Updation

The movement of each particle during the optimization in PSO is determined by a directional operator which is the velocity of that particle. The velocity of each particle is important in the process of convergence as it determines the direction in which a particle needs to move. In PSO, the velocity of each particle is determined based on its personal best position and the global best position of the swarm found so far. The angle of potential direction of any particle will be small if the direction of the personal best and global best are similar. However, a larger angle provides the exploration of the search space in a more effective way. The searching diagram of PSO algorithm is shown in Fig. 1.

Let $X_i(t)$ and $V_i(t)$ be the position and velocity of particle p_i at iteration t. The position $X_i(t)$ is updated at each iteration based on the velocity $V_i(t)$ as in (2).

$$X_i(t+1) = X_i(t) + V_i(t) \qquad (2)$$

The velocity vector reflects the socially exchanged information and is defined as in (3):

$$V_i(t+1) = W V_i(t) + C_1 r_1 (Pbest_i - X_i(t)) \qquad (3)$$
$$+ C_2 r_2 (Gbest - X_i(t))$$

Where $W, C_1, C_2 \geq 0$, W is the inertia weight, C_1 and C_2 are the acceleration coefficients, $r_1, r_2 \in [0,1]$ are random numbers, $Pbest_i$ is the personal best position of particle i, $Gbest$ is the global best position of the swarm found so far. The personal best position, global best position, inertia weight, and acceleration coefficients are the important parameters which influences the flight of the swarm towards the global optima. In case of inertia weight which is introduced in (Shi and Eberhart 1998), a larger value is better for global search while a smaller value is better for local search. In this work, $C_1 = C_2 = 2.0$ and inertia weight W is linearly decreased from 0.9 to 0.4.

The best position of each particle as well as of the swarm experienced so far are chosen based on the evaluated parameter values using the designed objective function considering the constraints related to the epipolar geometry of the stereo image pair. In this work, the objective function is the maximization of number of inliers pruned by the fundamental matrices which are represented as the particles. The algorithm is terminated as the swarm converges to the optimal solution.

In this work, matching error is computed for each estimated correspondence pair by applying the epipolar geometry encoded in the fundamental matrix (Trucco and Verri 1998), respecting the epipolar constraint of stereo matching as in equation

$$u_2^T \; F \; u_1 = \varepsilon_u \qquad (4)$$

where ε_u is the matching error, F is the fundamental matrix, (u_1, u_2) is the correspondence pair between the left image and the right image. Ideally, ε_u must be zero, but practically, the value ε_u does not come exactly zero. Therefore, we have decided a threshold value ε_0 and the corresponding pairs, which give matching error less than ε_0 are considered as inliers i.e.

$$u_2^T \; F \; u_1 \leq \varepsilon_0 \qquad (5)$$

Hence, each particle is evaluated by the number of inliers computed by that particle.
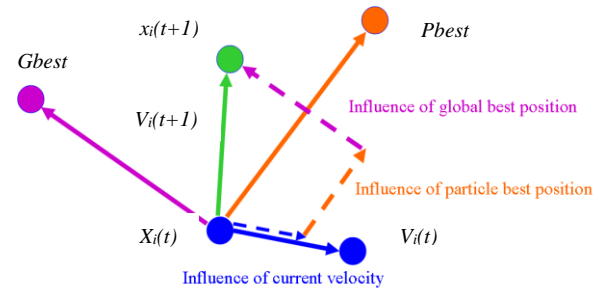

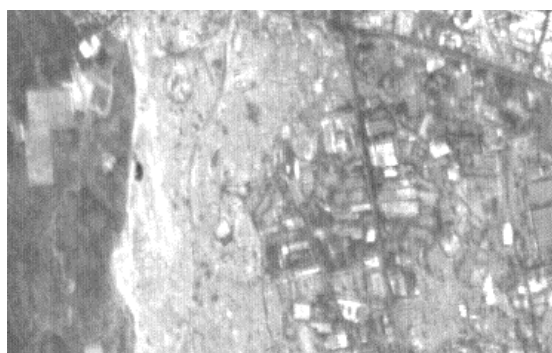Figure 1. Searching diagram of PSO algorithm

3. RESULT AND DISCUSSION

To demonstrate the effectiveness of the proposed approach, the results are obtained by applying the proposed method on a remotely sensed stereo image pair. A comparison of this method with RANSAC algorithm is carried out. The comparison shows that, the proposed method is effective to estimate robust fundamental matrix by giving improved number of inliers than RANSAC.

The test image pair is a subsection of the remotely sensed stereo image of Mumbai suburb area. The stereo image pair is taken by cartosat-1 in 2007. Its size is 235 X 370 pixels. Substantial illumination difference and geometric distortion among the image pair is present. This stereo image pair is used to measure the effectiveness and performance of the proposed method. The left image and right image of the dataset is shown in Figure 2a and 2b respectively.

In order to initialize the position of each particle in the swarm, SIFT feature points are extracted from the left image as well as from the right image of the test pair in Figure 2. The SIFT feature points are superimposed on the left image using red * points as shown in Figure 3. The feature points in the left image are matched with the feature points in the right image using SIFT feature matching algorithm and hence, the correspondence pairs are obtained. Applying RANSAC algorithm multiple times on the set of correspondence points, multiple fundamental matrices are obtained. These fundamental matrices are considered as the initial position of each particle in the swarm. It is clearly seen from the Figure 3 that, the feature points are distributed all over the image. Therefore, the fundamental matrices obtained from these correspondence points as the initial swarm increases the possibility to converge to global optima using the proposed optimization PSO based algorithm and hence, be able to estimate the accurate epipolar geometry of the given test stereo pair.
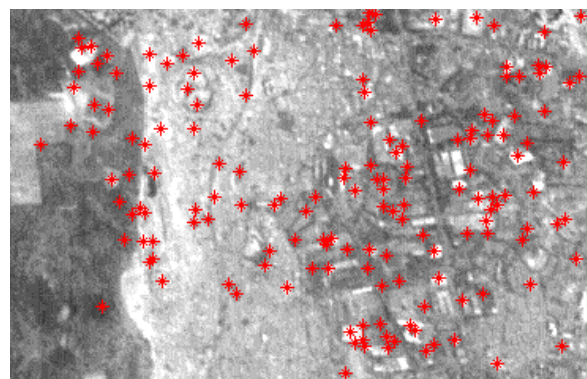
(a)

Figure 3: SIFT feature points (in red *) are superimposed on the left image

The particles in the swarm are moved through a number of iterations according to the velocity estimated using the information sharing mechanism. The algorithm is terminated as all the particles in the swarm converge to same solution and no further improvement is possible. Figure 4 is showing the correspondence points which are pruned by the fundamental matrix computed using the proposed particle swarm based optimization algorithm. In order to demonstrate the effectiveness of the proposed optimization approach, the number of inliers pruned by the fundamental matrix estimated using the proposed method and is compared with the number of inliers pruned by the fundamental matrix estimated using RANSAC algorithm. Table 1 shows the comparative result. The number of correspondence points extracted using SIFT feature matching algorithm is 162. Among them the number of inliers pruned by RANSAC algorithm is 40 whereas the number of inliers pruned by the proposed method is 53. The comparison shows that the fundamental matrix estimated using the proposed method is more robust by estimating the epipolar geometry compared with the RANSAC method.

(b)

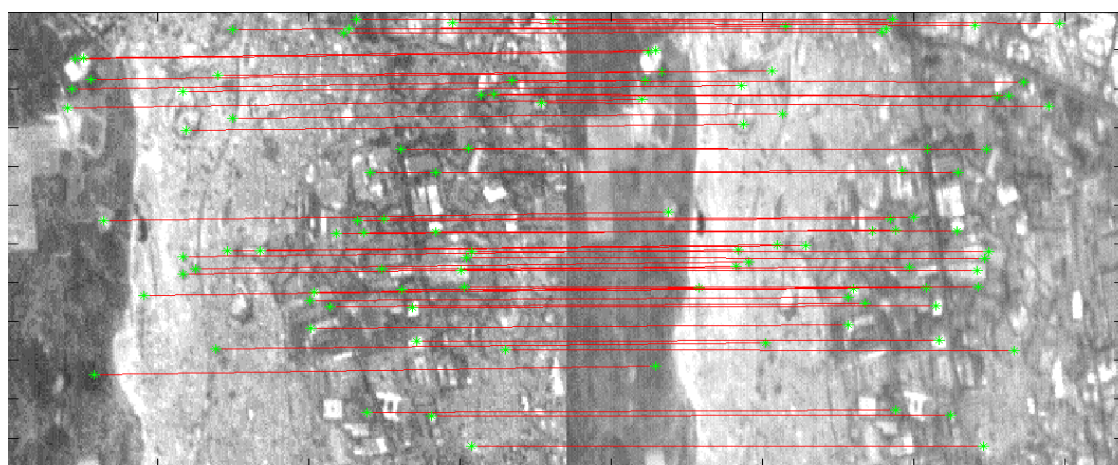Figure 2. (a) Left Image (b) Right Image of test image pair

Figure 4: The correspondence points pruned by the fundamental matrix estimated using the proposed method

Table 1. Comparative result of the proposed method with
RANSAC algorithm

Image	Number of correspondence points extracted using SIFT feature matching algorithm	Number of Inliers pruned by RANSAC	Number of Inliers pruned by the proposed method
Test Image Pair	162	40	53

4. CONCLUSION

The proposed particle swarm based optimization algorithm is effective to estimate the epipolar geometry of the stereo image pair. The estimation of fundamental matrix is a challenging problem due to the noise, occlusion, geometric and radiometric distortion present in the stereo image. The proposed approach is robust to the proportion of outliers in the stereo correspondences. The fundamental matrix is used as a constraint for finding inliers in many computer vision and photogrammetry applications. The obtained inliers are a useful input as ground control points for remotely sensed images. The initialization of the swarm is as effective in improving the convergence accuracy as well as the convergence time.

ACKNOWLEDGEMENT

REFERENCES

Eberhart, Russell, and James Kennedy. 1995. "A New Optimizer Using Particle Swarm Theory." In *Micro Machine and Human Science, 1995. MHS'95., Proceedings of the Sixth International Symposium On*, 39–43.

Fischler, Martin A, and Robert C Bolles. 1981. "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography." *Communications of the ACM* 24 (6): 381–95.

Gong, Maoguo, Qing Cai, Xiaowei Chen, and Lijia Ma. 2014. "Complex Network Clustering by Multiobjective Discrete Particle Swarm Optimization Based on Decomposition." *IEEE Transactions on Evolutionary Computation* 18 (1): 82–97.

Joglekar, Jyoti, Shirish S. Gedam, and B. Krishna Mohan. 2014. "Image Matching Using SIFT Features and Relaxation Labeling Technique -A Constraint Initializing Method for Dense Stereo Matching." *IEEE Transactions on Geoscience and Remote Sensing* 52 (9): 5643–52.

Longuet-Higgins, H. C. 1981. "A Computer Algorithm for Reconstruction a Scene from Two Projection." *Nature* 293: 133–35.

Lowe, David G. 2004. "Distinctive Image Features from Scale-Invariant Keypoints." *International Journal of Computer Vision* 60 (2): 91–110.

Mikolajczyk, Krystian, and Cordelia Schmid. 2005. "A Performance Evaluation of Local Descriptors." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27 (10): 1615–30.

Shi, Yuhui, and Russell Eberhart. 1998. "A Modified Particle Swarm Optimizer." In *Evolutionary Computation Proceedings, 1998. IEEE World Congress on Computational Intelligence., The 1998 IEEE International Conference On*, 69–73.

Trucco, Emanuele, and Alessandro Verri. 1998. "Introductory Techniques for 3-D Computer Vision". *Prentice Hall Englewood Cliffs*.

Xu, G., and Z. Zhang. 1996. "Epipolar Geometry in Stereo, Motion and Object Recognition". *Kluwer Academic Publishers*.