EVOLUTIONARY AGENTS FOR EPIPOLAR GEOMETRY ESTIMATION

Mingxing Hu [1], *Gordon Dodds*[1], *Baozong Yuan* [2]

1. Virtual Engineering Centre, Queen's University Belfast, Belfast, BT9 5HN, United Kingdom;
2. Institute of Information Science, Beijing Jiaotong University, Beijing, P. R. China, 100044
Email: m.hu@qub.ac.uk

ABSTRACT

This paper presents an evolutionary agent-based approach to epipolar geometry estimation. Each agent stands for a minimum subset for computing fundamental matrix, and evolves autonomously in the vast solution space to get the optimal result. In so doing, the agents rely on some reactive behaviors such as reproduction and diffusion, and collaborate with others with a subset template. Experimental results show that our approach performs better than other typical methods in terms of accuracy and computational efficiency, and is robust to noise and outliers.

1. INTRODUCTION

Matching two images of a single scene from different viewpoints is a fundamental problem in computer vision. In the last two decades, this problem has been intensively studied and many methods have been proposed. These methods almost all share the same geometrical constraint, i.e. the epipolar constraint, and the associated epipolar geometry consists of the minimal amount of information between two images and plays an important role in stereo vision.

The existing techniques for epipolar geometry estimation can be broadly classified into linear, iterative and robust approaches in [1]. The linear methods mainly include eight-point algorithm [1] and Hartley's eight-point algorithm [1]: the former is easy to use, but unstable and sensitive to noise; the latter improves the stability by normalizing the coordinates of image points before solving an eigenvalue problem. Iterative methods mainly minimize the epipolar distance between points and epipolar lines, or the Sampson distance based on the gradient. Chojnacki [2] recently addressed two iterative methods, FNS and CFNS, based on an approximate maximum likelihood estimate. The iterative methods offer a more accurate solution, but cannot cope with outliers. The outliers, however, are in gross disagreement with a specific postulated model and have to be handled by robust approaches. The best-known robust methods are probably Least Median of Squares (LMedS), Random Sample Consensus Paradigm (RANSAC), Maximum Likelihood Sample Consensus (MLESAC) and Maximum a Posterior Sample Consensus (MAPSAC). LMedS and RANSAC are quite similar. Both of them take randomly-selected subsets and repeat the process enough times to get the best solution. The difference lies in the way used to determine the result. The MLESAC and MAPSAC are the improvement upon RANSAC, and much more statistical models are applied for the final solution. We have already addressed an approach based on genetic algorithms (GAs) for robust estimation [3], which is especially fit for conditions where a significant number of outliers are expected. But GAs are compute-intensive, which is a defect of all GA-based applications.

Therefore, a new method is badly needed, which improves the accuracy and robustness of geometry estimation, and reduces the computational expense as well.

In this paper, we address an evolutionary agent-based [4] approach to epipolar geometry estimation, which considers both sides of the problem described above. In our method, the correspondences are viewed as a one-dimensional cellular environment in which the agents inhabit and evolve, and each agent stands for a minimum subset for computing a unique fundamental matrix. The agents execute some evolutionary behaviors such as breeding and diffusion, and collaborate with each other using a subset template updated after each generation. Hence, the offspring agents will likely find the optimal result in the vast search space. Experiments with both synthetic data and real images show that our method works as fast as iterative methods with loose constraints, i.e. CFNS, but obtains accurate results as well as or even better than the robust methods, i.e. LMedS, MAPSA and GAs.

The remainder of the paper is organized as follows. First, a brief introduction to epipolar geometry is presented. Then the new approach based on evolutionary agents is presented in detail, including agent definition, fitness function, evolutionary behaviors and subset template. Section 4 deals with the experimental results obtained from synthetic data and real images. Finally, the conclusions are drawn.

2. BACKGROUND

Consider the case of two images acquired from a 3D scene. An image point $\mathbf{m}_i$ in the left view corresponds to an image point $\mathbf{m}'_i$ in the right. According to the well-known epipolar constraint, we have [1]

$$\mathbf{m}_i'^T \mathbf{F} \mathbf{m}_i = 0$$

0-7803-8554-3/04/$20.00 ©2004 IEEE.

where $\mathbf{F}$ is a 3×3 matrix called fundamental matrix, $\mathbf{m}_i = (x_i,\ y_i,\ 1)^T$, $\mathbf{m}'_i = (x'_i,\ y'_i,\ 1)^T$ are the homogenous vectors for the corresponding points. $\mathbf{F}$ can be obtained by using only seven correspondences, which form the data matrix $\mathbf{Z} = (\mathbf{z}_1,\ \cdots,\ \mathbf{z}_7)$, and $\mathbf{z} = (x'_i x_i,\ x'_i y_i,\ x'_i,\ y'_i x_i,\ y'_i y_i,\ y'_i,\ x_i,\ y_i,\ 1)^T$, $i = 1,\ \cdots,\ 7$. The null space of the moment matrix $\mathbf{M} = \mathbf{Z}^T \mathbf{Z}$ is dimension 2, barring degeneracy ($\mathbf{Z}$ is 7×9). It defines a one-parameter family which exactly fits to the seven correspondences: $\alpha \mathbf{F}_1 + (1-\alpha)\mathbf{F}_2$. Introducing the constraint $\det|\mathbf{F}| = 0$ leads to a cubic in α

$$\det|\alpha \mathbf{F}_1 + (1-\alpha)\mathbf{F}_2| = 0$$

which has 1 or 3 real solutions for α . Ideally every possible subsample (seven correspondences) of all the correspondences n should be considered to get the optimal result, but this is usually computationally infeasible.

3. EPIPOLAR GEOMETRY ESTIMATION BASED ON EVOLUTIONARY AGENTS

According to the description in the last section, the result we prefer is the subset that contains only good data points (all inliers, that are slightly disturbed by image noise), which can be considered as an optimization problem of searching in a vast solution space. And evolutionary agents can be able to adjust their behaviors based on some learning mechanisms, which especially fits for the case of searching and optimization. So our work aims to apply evolutionary agents to epipolar geometry estimation and focuses on the computational aspect of the agents to obtain the optimal result.

3.1 Agent definition

Suppose that $\mathbf{S}$ is the dataset of all the correspondences $\{(\mathbf{m}_i,\ \mathbf{m}'_i) \mid i = 1,\ \cdots,\ n\}$ obtained from corner detector and matching algorithm. It may be viewed as a one-dimensional grid lattice of pairs of correspondences. The goal of the evolutionary agents in $\mathbf{S}$ is to search for the optimal subset we prefer. That is to say, the search space, $\mathbf{S}$, may be considered as an environment in which the agents inhabit and evolve.

Based on the above analysis, the evolutionary agent is defined as follows

$$Agent = \left\langle\ \mathbf{p},\ a,\ D_{fit},\ fml,\ \mathrm{Rep},\ Diff,\ Die\ \right\rangle$$

which includes seven parameters to denote its structure and evolutionary behaviors. $\mathbf{p}$ stands for the positions of an agent in $\mathbf{S}$, which is a seven-dimensional position vector, and the entry $p_k = \{i \mid i = 1,\ \cdots,\ n\}$, $k = 1,\ \cdots,\ 7$, is just the index number of correspondence lattice $\mathbf{S}$, as shown in Figure 1. In other words, $\mathbf{p}$ stands for the subset of correspondences for geometry estimation. a denotes the age of an agent, that is, the number of diffusion steps it has taken; D_{fit} symbolizes its fitness, which indicates the adaptability of an agent and can be computed using the epipolar constraint obtained from the correspondence of $\mathbf{p}$; fml represents the family index, which indicates where an agent comes from.

The above four parameters describe the internal state of an agent, while *Rep*, *Diff* and *Die* describe the external behaviors of it. *Rep* denotes the reproduction behavior; *Diff* represents the diffusion behavior; while *Die* indicates that an agent has a life span, it may die like a living thing.

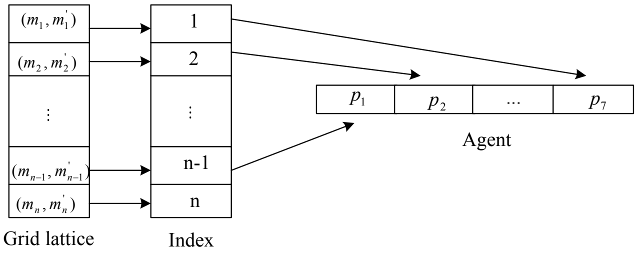

Figure 1. Agent representation

3.2 Fitness function

In order to compare the fundamental matrix obtained, we compute Sampson distance of correspondences between images. For more information about Sampson distance, readers are referred to the works [1, 5]. Sampson Distance can be expressed as:

$$D_S = \sum d_S^2 = \sum r^2 w_S = \sum r^2 \left(\frac{1}{r_x^2 + r_y^2 + r_x^{'2} + r_y^{'2}}\right)$$

where

$$r_x = f_1 x' + f_4 y' + f_7\ , \quad r_y = f_2 x' + f_5 y' + f_8$$

$$r_x^{'} = f_1 x + f_2 y + f_3\ , \quad r_y^{'} = f_4 x + f_5 y + f_6\ .$$

Then the fitness function of an agent

$$D_{fit} = \sum_{i=1}^{n} \omega_i d_i + \sum_{i=1}^{n} \beta(1-\omega_i)$$

where ω_i satisfies the following equation

$$\omega_i = \begin{cases} 1 & if\ \ d_i \le \beta \quad (inlier) \\ 0 & otherwise \quad (outlier) \end{cases}$$

and $\beta = 1.96\delta = 1.96 \times 1.4828(1 + \frac{5}{n-p})\sqrt{med_i\,|\,d_i\,|}$ [5], where p the dimensionality of the parameter. That is to say, we only consider the Sampson distance of inliers and the outliers make no contribution to the fitness function.

3.3 Evolutionary behavior [4]

Evolutionary agents adapt to their environment mainly by way of switching between two behavioral responses, namely, reproduction and diffusion. And we employ $A^{(g)}$ to stand for the set of all active agents in generation g .

(1) Reproduction: In the reproduction process, each active agent $\alpha^{(g)}$ (g denotes the generation of the evolutionary process) will breed a finite number, m , of offspring agents, $\alpha^{(g+1)}$. The larger m is, the more offspring

agents will be created to search in the vast solution space. However, the computational time will increase dramatically, if no parallel algorithm is applied. Thus, in the experiments, m is set to 2. The differences between $\alpha^{(g)}$ and $\alpha^{(g+1)}$ are mainly in the position vectors $\mathbf{p}^{(g)}$, $\mathbf{p}^{(g+1)}$. Two elements of $\mathbf{p}^{(g)}$ are selected and changed into: (1) an element from the subset template (which will be described in the following section) or (2) a random index number of $\mathbf{S}$. The probability of the former case is P_c, and that of the latter is $1-P_c$. The larger P_c is, the fewer new positions will appear in the offspring, which affects the explorative ability of agents. So P_c is fixed to 0.3.

In other words, $\alpha^{(g+1)}$ may be viewed as the mutation of his parent $\alpha^{(g)}$, and new values are introduced to $\mathbf{p}^{(g+1)}$ from highly potential good data (subset template) or from random generator. In GAs [3], however, only a part of chromosomes will be selected for reproduction from the gene pool by means of roulette wheel.

(2) Diffusion: The diffusion behavior is important for an agent to search for new positions in correspondence lattice. After the reproduction process, each agent of generation $(g+1)$, $\alpha^{(g+1)}$, will compare its fitness with that of its parent $\alpha^{(g)}$. Only those agents which have better fitness will survive and be appended to the agent set $A^{(g+1)}$. Their parents, however, will become inactive and be removed from the agent set. If $\alpha^{(g+1)}$ are worse than their parents, they will be deleted at once and the ages of their parents $\alpha^{(g)}$ will be increased by one.

If the age of an agent exceeds its life span, it will be removed from the environment, which avoids endless trial-and-error and thus reduces useless computations. If there is no active agent in the evolutionary environment, the whole process halts.

3.4 Subset template

In order to improve the collaborative ability of agents, we apply a subset template in the evolutionary process. After each generation (reproduction and diffusion), we compute the fitness of every active agent, and the position vector of the agent with maximum fitness is selected as the subset template of this generation. So in the next generation the agents can select the elements from the template for reproduction by probability P_c, which have higher possibility to be good data points. This scheme enables evolutionary agents to inherit the most effective reproduction and diffusion from their successful parents.

4. EXPERIMENTAL RESULTS

The above section has provided a computational model of evolutionary agent, and the complete algorithm is given in Figure 2.

In this part, our approach is compared with several typical methods, including LMedS, RANSAC, MAPSAC, CFNS and GAs. And the source codes for comparison are partly provided by X. Armangue [6].

```
begin
   Initialize the agents $\{\alpha^{(0)}\}$ by selecting the elements of
position vector $\mathbf{p}$ from the index of $\mathbf{S}$ in numerical
order ($1, 2, 3, \cdots, n$),
   Compute the fitness of $\{\alpha^{(0)}\}$,
   Select the subset template from $\{\alpha^{(0)}\}$,
   While $A \neq \phi$ do
      Reproduction process,
      Compute the fitness of newly reproduced agents,
      Diffusion process,
      Vanishing process,
      Update the subset template,
   end while
```

Figure 2 The algorithm for evolutionary agent computation

4.1 Experiments with synthetic data

In the experiments with synthetic data, the correspondences are randomly generated by space points in the region of $\Re^3$ visible to two different positions of a synthetic camera: $\mathbf{P}_1 = \mathbf{C}[\mathbf{I} \mid \mathbf{R}]$ ($\mathbf{C}$ stands for camera intrinsic matrix) and $\mathbf{P}_2 = \mathbf{C}[\mathbf{R} \mid \mathbf{t}]$, where the camera makes a rotation $\mathbf{R}$ and a translation $\mathbf{t}$. Here the total number of correspondences is 100, and there are only 10 agents in $\{\alpha^{(0)}\}$. The number of agents for initialization may be larger than 10, but it will take more time for computation and ten agents are also good enough for real applications. The experiments are divided into two groups:

(G1): Six different groups of Gaussian noise are added to the projective correspondences, whose means are 0 and variances vary from 0.5 to 3.0 (in steps of 0.5) in Table 1.

(G2): The means and variances of Gaussian noise are fixed to 0, 1, respectively; the percentage of outliers disturbed by the noise and false matches varies from 10% to 50% (in steps of 10%) in Table 2.

Table 1 Sampson distance under variable variance of noise

	LMedS	RANSAC	MAPSAC	CFNS	EA	GA
0.5	2.40	2.98	0.32	0.44	0.24	0.23
1.0	3.43	4.52	0.48	0.88	0.50	0.53
1.5	2.25	5.71	0.93	1.52	0.99	0.94
2.0	3.69	4.04	1.28	2.08	1.21	1.07
2.5	3.28	5.24	1.49	2.70	1.65	1.62
3.0	3.37	5.07	1.86	2.70	1.75	1.79

Table 2 Sampson distance under different percentage of outliers

	LMedS	RANSAC	MAPSAC	CFNS	EA	GA
10%	2.35	4.73	1.21	2.67	0.81	0.69
20%	2.39	4.83	1.72	2.82	1.16	1.03
30%	2.46	6.85	1.42	2.99	1.18	1.30
40%	1.75	8.25	2.10	2.89	1.71	1.57
50%	3.06	9.26	2.87	2.95	1.84	1.81

Table 3 Average computing times for two groups (s)

Group	LMedS	RANSAC	MAPSAC	CFNS	EA	GA
G1	1.89	0.93	4.04	0.60	0.82	2.47
G2	1.86	0.94	4.09	0.63	0.86	2.49

Table 1 and 2 show the experimental results of (G1) and (G2), respectively. And Table 3 shows the average computing time spent by the methods in (G1) and (G2). From these tables, we can notice that EA (Evolutionary Agent) method performs almost as well as the GA approach, but the computing time of GA is three times as that of EA. MAPSAC is also a well-known robust method against noise and outliers. But the results are slightly worse than those of EA, and the computational time of MAPSAC is very expensive. CFNS turns out to be the quickest method in the six algorithms, but the Sampson distances increase greatly when the variance of noise is great or a large number of outliers are involved.

4.2 Experiments with medical images

We also employ medical images for comparison. Three different pairs of images are taken from a laparoscopic operation. Figure 3 illustrates the first pair of images we use and the white circles denote the feature points obtained by the corner detection and matching.

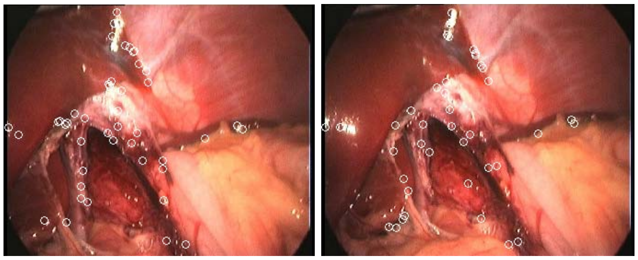

Figure 3 The medical images from two viewpoints

Table 4 Sampson distance of different pairs of medical images

Group	LMedS	RANSAC	MAPSAC	CFNS	EA	GA
MG1	6.03	17.68	3.53	13.16	3.10	2.96
MG2	2.71	16.55	4.85	9.67	4.64	4.04
MG3	4.60	25.73	4.47	10.91	4.15	3.78

Table 5 Average computing time for three groups (Second)

	LMedS	RANSAC	MAPSAC	CFNS	EA	GA
Time	1.56	0.79	3.88	0.54	0.80	2.45

Table 4 and 5 show the Sampson distances and computing time of the experiments, respectively. We can see that the accuracy of EA is a little worse than that of GA, but better than those of other four methods. The mean Sampson distances of LMedS, RANSAC, MAPSAC, CFNS and GA are 1.12, 5.04, 1.08, 2.83, 0.91 times as much as that of EA. As to the computational efficiency, EA works so fast that the computing time for EA is 0.33 and 0.21 as much as those of GA and MAPSAC, respectively. Here it should be pointed out that LMedS gives good results in terms of accuracy (especially in MG2), but it does not always model the epipolar geometry properly. For example, we detect 79 correspondences in the second image pair (MG2), and do the experiments with LMedS and EA, respectively. The average number of outliers detected by LMedS is 48.8 (more than 60% of all the correspondences), and as to EA it is only 28.2. In other words, LMedS removes too many corresponding points, even some good ones, to make the Sampson distances smaller. Generally speaking, evolutionary agent-based approach performs best in terms of accuracy and computational efficiency.

5. CONCLUSION

In this paper, we describe an evolutionary agent-based approach to epipolar geometry estimation. Each agent stands for a minimum subset of corresponding points, and inhabits and evolves in the correspondence lattice. The progressively reproduced agents, being computational entities, directly operate in the given environment, execute a number of reactive behavioral responses, and collaborate with each other with a subset template. Experiments with both synthetic data and medical images show that our approach performs better than other typical methods in terms of accuracy and computational efficiency. It can obtain an optimal (or near optimal) result in the vast solution space and is robust to the outliers disturbed by bad locations and false matches.

6. ACKNOWLEDGEMENTS

Thanks to Royal Belfast Hospital for provding the laparoscopic images. The research is partly supported by VEC, National Natural Science Foundations and Doctoral Foundations of China.

7. REFERENCES

[1] X. Armangue and J. Salvi, “Overall view regarding fundamental matrix estimation”, Image and Vision Computing, vol.21, pp. 205-220, 2003

[2] W. Chojnacki, M.J. Brooks, A. Van Den Hengel and D. Gawley, “FNS and HEIV: relating two vision parameter estimation frameworks”, 12th International Conference on Image Analysis and Processing, pp. 152-157, 2003.

[3] M.X Hu, B.Z. Yuan and X.F. Tang, “A new robust fundamental matrix estimation based on genetic algorithm”, Computing and Informatics, vol. 20, no. 5, pp. 471-486, 2001.

[4] J.M Liu, Y.Y. Tang and Y.C. Cao, “An evolutionary autonomous agents approach to image feature extraction”, IEEE Trans. On Evolutionary Computation, vol. 1, no. 2, pp. 141-158, 1997.

[5] P.H.S. Torr and D.W. Murray, “The development and comparison of robust methods for estimating the fundamental matrix”, International Journal of Computer Vision, vol. 24, no. 3, pp. 271-300, 1997.

[6] Matlab Toolkit: Http://eia.udg.es/~armangue/research