# QUALITY ANALYSIS OF VEHICLE-BASED SEQUENCE IMAGES RELATIVE ORIENTATION BASED ON COMPUTER VISION

YAN Li [a], XU Zhenliang [a,b] *

[a]School of Geodesy and Geomatics, Wuhan University, 129 Luoyu Road, Wuhan 430079, China
Email: lyan@sgg.whu.edu.cn;xuzhenliang@whu.edu.cn.
[b] School of Geomatics, Liaoning Technical University, 47 Zhonghua Road, Fuxin 123000, China

**Commission III, WG III/5**

**KEY WORDS:** Relative Orientation, Direct Geo-referencing, Computer vision, Sequence images, Mobile mapping, SIFT, RANSAC

**ABSTRACT:**

Analyzed the quality and its influence factors of relative orientation of vehicle-based sequence images by comparing with the result obtained from Position & Orientation System using Direct Geo-referencing. Studies have shown that, under normal circumstances, the image Relative Orientation based on Computer vision is more robust, and under special conditions, it is more practical for analyzing and increasing the quality of vision measurement. Besides these, the method can be borrowed by many surveying and mapping related fields, taking indoor robot mobile environment awareness and unmanned automobile for instance.

## 1. INTRODUCTIONS

Recent years, the using of POS (Position & Orientation System) in providing direct geo-referencing for imaging equipment makes it a breakthrough in converting MMS (Mobile Mapping System) to realistic productivity; it has been widely used in such space information service areas as digital city and intelligent transportation. Besides, it broadens the content of information mapping and surveying, improving the work efficiency and quality magnificently. But, in such areas as traditional photogrammetry, low altitude remote sensing based on unmanned airborne, close range photogrammetry on the ground, considering of the cost, safety and complexity, it is more practical to acquire the relative attitude of images by traditional vision methods, that is resection; Besides, in indoor robot environment sensing and unmanned driving field, it commonly uses vision to achieve autonomous navigation. So, RO (Relative Orientation) based on vision is still of great value in many areas.

In photogrammetry and CV (Computer Vision), RO between sequential images is achieved by the coplanar conditions of homologous points between two images and corresponding object points, in recent years, with the improvement of efficiency of matching algorithm, RO has realized automation in numerous of industrial photogrammetry software,（LI Jian et al 2010）, but this model mainly suites for conditions when the heading changes little, this model usually uses linearization and LS (Least Square) iterating to get the orientation parameters, but this model does not suits for MMS because of the great heading, the iterating is easily plunging into the local optimization and leads to failure if the initial value is not chose properly. Besides, the precision of traditional photogrammetry is just theoretic, objective precision of RO cannot be acquired, so one cannot analyze the factors that affect the precision thoroughly; the occurrence of high precision DG (Direct Geo-referencing) makes it possible to check the quality of RO based on traditional method. Based on this, this paper proposes a direct way to analysis the RO absolute error of sequential vehicle-based images based on DG and CV.

## 2. THE VISION RO PROCEDURES OF SEQUENTIAL IMAGES

The RO procedures of vehicle-based sequential images include the selection of stereo images, feature points extraction and matching, the estimation and evaluation of RO Euler angle; the basic processes are described as follows:

(1) The selection of sequential images. The MMS (mobile mapping system) is mainly aimed at acquiring city environment data, so when selecting sequential images, we can choose some building images with some degree of overlapping and with abundant feature points.

(2) Feature points extraction and matching of sequential images. The cameras on mobile mapping system usually belong to small frame non-metric camera, and the sequential images of MMS usually have large heading angle, the matching algorithms based on gray correlation in traditional photogrammetry can hardly insure the correctness of the RO (Schenk and Toth,1993). Fortunately, the high degree overlapping of MMS sequential images provides basic fundamentals for image matching and RO. Besides, SIFT (Lowe,2004), which is famous in matching algorithms in CV, makes it possible to solve this question. The features of SIFT are invariant to image scale and rotation, and are shown to provide robust matching across a substantial range of affine distortion, change in 3D viewpoint, addition of noise, and change in illumination. The features are highly distinctive, in the sense that a single feature can be correctly matched with

*YAN Li, professor, Ph.D supervisor. His research field includes digital photogrammetry,computer vision and Mobile Mapping Technology(MMT).

high probability against a large database of features from many images, so SIFT algorithm attracts many attentions in photogrammetry especially in MMS. In the experiment, firstly, extract the feature points of sequential images and coarse matching, then reject the outliers to get the inliers by RANSAC method (Fischler and Bolles, 1981), namely precise matching. The experiment result shows that, in conditions without precise initial value and hypothesis test, this method is still very effective in extracting feature points and matching (McGlone et al., 2004 ; Nister 2003).

(3) The estimation of sequential images altitude matrix based on CV. The RO in photogrammetry and CV is the same in theoretic, just there are some differences in describing and solving the question. In traditional photogrammetry, the RO between images mainly accomplished by liberalizing the model and combining the LS iterating, but in CV, it is though fundamental matrix to get the relationship of homologous points in two images, and get altitude matrix by matrix decomposition of fundamental matrix (ZHANG Zuxun，2004).

(4) The acquiring and quality estimation of relative altitude angle. The procedure is as described in Figure 1.

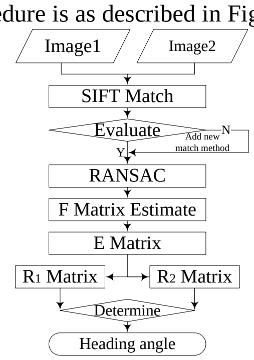



Figure1. The RO procedures of sequential images based on CV

## 3. THE MATHEMATICAL MODEL OF SEQUENTIAL IMAGES RO BASED ON CV

### 3.1 The coordinate system of IMU and Camera in MMS

The coordinates related in this paper mainly consist of IMU and Camera coordinate based on CV. the CV coordinate, the origin of the coordinate is the camera optic center, and $Z_C$ points from the image plain to the object. It is just as described in Figure 2.

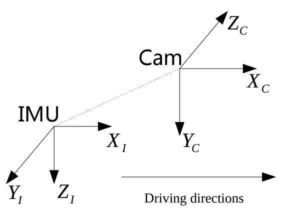



Figure2. The IMU and Camera coordinate system of LandMark MMS

The coordinate axes $X_I$-$Y_I$-$Z_I$ of IMU and the coordinate axes $X_C$-$Y_C$-$Z_C$ of camera are nearly parallel or perpendicular, the relationship between them can be expressed as:

$$\boldsymbol{R}_{IMU}^{Cam} = \Delta\boldsymbol{R}\cdot\boldsymbol{R}_0 = \begin{bmatrix} 1 & -\gamma & -\beta \\ \gamma & 1 & -\alpha \\ \beta & \alpha & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

Where $\Delta\boldsymbol{R}$ is the extrinsic calibration matrix of the camera, $\alpha\beta\gamma$ are misalignment of IMU coordinate system and camera coordinate system, they are all small, in this experiment, $\Delta\boldsymbol{R}$ is not rigorous calibrated.

### 3.2 The model of images RO based on CV

The procedures of getting orientation parameters directly by matrix decomposition are described below:

(1) Acquire the fundamental matrix $\boldsymbol{F}$ by the homologous points of two sequential stereo images model combing the 8 points fitting method. And $\boldsymbol{F}$ must meet:

$$\boldsymbol{xFx'} = \boldsymbol{0}$$

(2) Acquire the essential matrix $\boldsymbol{E}$ of two normalized images by the calibrated internal parameters matrix $\boldsymbol{K}$ of the camera.

$$\boldsymbol{E} = \boldsymbol{K}^T\boldsymbol{FK} \text{ or } \hat{\boldsymbol{x}}\boldsymbol{E}\hat{\boldsymbol{x}}' = \boldsymbol{0}$$

(3) Acquire the altitude matrix $\boldsymbol{R}$ of sequential images by SVD decomposition of essential matrix $\boldsymbol{E}$, in CV, it usually makes the coordinate system of the first camera as world coordinate system, and other cameras are relative to it. Taking the RO of sequential images into consideration, we make coordinate of the former image as world coordinate system. According to the practical situation of MMS, generally, there are two results $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ of essential matrix decomposition (Hartley,R and Zisserman,A.,2003)

$$\boldsymbol{R}_1 = \boldsymbol{UWV}^T \text{ or } \boldsymbol{R}_2 = \boldsymbol{UW}^T\boldsymbol{V}^T$$

where,

$$\boldsymbol{W} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathrm{SVD}(\boldsymbol{E}) = \boldsymbol{USV}^T$$

finally, keep one of $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ according to practical situation.

(4) Altitude matrix decomposition based on Inertial Navigation System. As we all know, the coordinate of IMU ( $\boldsymbol{O}_I$ ) is defined based on Local horizontal coordinate ($\boldsymbol{O}_{LH}$).Transform matrix from $\boldsymbol{O}_I$ to $\boldsymbol{O}_{LH}$ can be expressed as $\boldsymbol{R}_I^{LH}$ .So in the test, the altitude matrix between former and later images can be expressed as $\boldsymbol{R}$ :

$$\boldsymbol{R} = \boldsymbol{R}_I^{LH}$$

$$= \begin{bmatrix} \cos y & 0 & \sin y \\ 0 & 1 & 0 \\ -\sin y & 0 & \cos y \end{bmatrix} \begin{bmatrix} \cos p & -\sin p & 0 \\ \sin p & \cos p & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos r & -\sin r \\ 0 & \sin r & \cos r \end{bmatrix}$$

$$= \begin{bmatrix} \cos y\cos p & -\cos y\sin p\cos r + \sin y\sin r & \cos y\sin p\sin r + \sin y\cos r \\ \sin p & \cos p\cos r & -\cos p\sin r \\ -\sin y\cos p & \sin y\sin p\cos r + \cos y\sin r & -\sin y\sin p\sin r + \cos y\cos r \end{bmatrix}$$

$$= \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$$

where, $y$ , $p$ , $r$ represent the difference of yaw, pitch and roll of two sequential images, so:

$$roll = \arctan\left(-\frac{b_3}{b_2}\right)$$

$$pitch = \arcsin(b_1)$$

$$yaw = \arctan\left(-\frac{c_3}{a_1}\right)$$

## 4. EXPERIMENT AND ANALYSIS OF VEHICLE-BASED SEQUENTIAL IMAGES RO

### 4.1 The experiment data and method

In this paper, the vehicle-based sequential images are provided by Beijing maggroup group (www.maggroup.org), the size of the images is 1600×1200, and the size of the pixels is 4.4μm × 4.4μm, the focal length of the camera is 7.5mm; In city with flat area, the heading angle of the sequential images changes greatly(generally 20°), when the MMS moves smoothly, the pitch and roll change little, which is the feature of the sequential images in such area, so we ignore this in the experiment. The data showed in table 1 is the measurement value of IMU of 11 sequential images taken around a supermarket, the biggest heading angle is 17°, and the smallest heading angle is 3°.

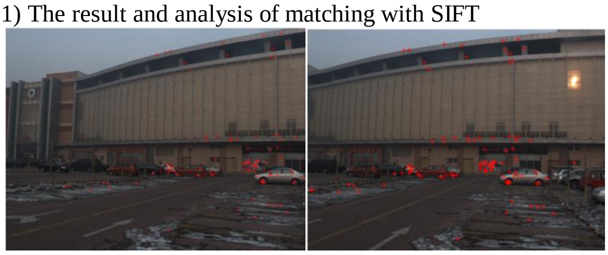
Table.1 The Heading IMU measurement value of sequential images

| Image No. | 70 | 71 | 72 | 73 |
|---|---|---|---|---|
| IMU meas value(°) | 317.547 | 16.965 | 346.878 | 3.884 |
| Image No. | 74 | 75 | 76 | 77 |
| IMU meas value(°) | 17.608 | 27.608 | 34.193 | 38.051 |
| Image No. | 78 | 79 | 80 | |
| IMU meas value(°) | 41.139 | 43.705 | 46.675 | |

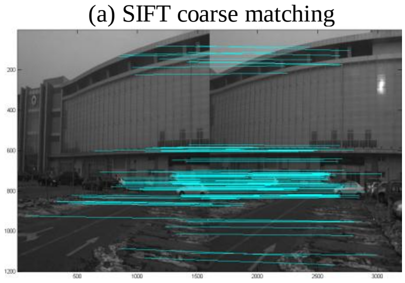
In the experiment, we use Matlab as software, and we use Intel (R) Core (TM) i3 CPU with four processors, the memory is 4G.

**4.2 The result and the analysis of the experiment**

(1) The result and analysis of matching with SIFT



(a) SIFT coarse matching



(b)RANSAC precise matching

Figure3. Result of coarse and precise matching (Images 70#,71#)

Taking image 70# and 71# for example, 722 feature points and 776 feature points were found in the two images with algorithm SIFT, it can be seen from the picture that, the feature points founded are well distributed, meanwhile, there are some wrong matches in areas with similar textures, after rejecting the outliers by RANSAC algorithm, 115 inliers can be acquired, after identification, the incorrect matching rate is within 2%, the results are nearly the same with different sequential images, so it can be concluded that the use of SIFT algorithm combining RANSAC in the feature points extraction and matching of vehicle-based sequential images is stable.

Besides, the RO computing time of per pair of images is about 10s, after analysis, we find that matching takes about 90% of total time, of course, the efficiency of data processing have a lot to do with the SIFT algorithm and the experiment platform, and SIFT algorithm extracts feature points in per image separately and matches with other images, and this is the bottleneck of quick matching by SIFT.

(2) The result and analysis of sequential images RO based on CV

The data showed in table 2 are the IMU measurement value and CV Calculation value of RO of 11 sequential images selected.

| Image pairs | 70-71 | 71-72 | 72-73 | 73-74 | 74-75 |
|---|---|---|---|---|---|
| IMU Value(°) | 12.366 | 16.965 | 17.006 | 13.724 | 10.000 |
| CV Value(°) | 13.217 | 19.296 | 17.451 | 16.120 | 12.561 |
| Difference(°) | -0.851 | -2.331 | -0.445 | -2.396 | -2.562 |
| Image pairs | 75-76 | 76-77 | 77-78 | 78-79 | 79-80 |
| IMU Value(°) | 6.585 | 3.858 | 3.088 | 2.566 | 2.970 |
| CV Value(°) | 8.551 | 6.770 | 4.873 | 5.226 | 5.508 |
| Difference(°) | -1.966 | -2.912 | -1.784 | -2.660 | -2.538 |

Table.2 The heading angle of IMU measurement and CV Calculation

Because this MMS uses high precision POS/LV made by Appplanix and the position and attitude are acquired by tightly value of IMU, combining the IMU measurement value and GPS measurement value, the attitude can be seen as truth value within a short time, We compare the attitude acquired by CV to the measurement.

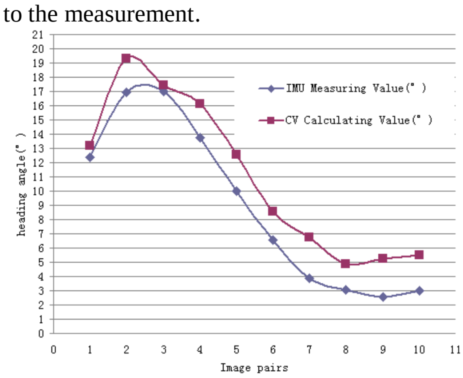



Figure4.The comparison of heading value from IMU and CV

From Figure4 we find that, the heading value acquired from these two methods changes in the same changing tendency, the smallest difference is 0.445°and the biggest difference is 2.912°, and the RMS(Root Mean Square) is 0.768°. This result is nearly the same with the result in paper (DING Meng and CAO Yunfeng, 2007), this is to say, the absolute error of RO of vehicle-based sequential images based on CV is within 3°.This mainly relates to the quality of the images and matching algorithm. The quality of the images affected by noise, brightness, occlusion, foreshortening and image pixel dissimilarity caused by resampling effects, and matching algorithm directly affects the precision and distribution of the feature points extracted.

After research, it can be seen that the difference of heading calculated by CV and IMU measuring value is always negative or positive, it shows systematic, and preliminary judgment it is caused by the misalignment of the camera coordinate and the IMU coordinate. Besides The greater the heading is, the smaller the compute error, and the smaller the heading is, the greater the compute error, it can be concluded that small heading is easy to lead to instability of computing in some degree.

## 5. CONCLUSIONS

Applying the matching and estimation method of CV in MMS, broaden the research area of photogrammetry, it can be concluded from the experiment that have conducted:
(1) Compare to traditional photogrammetry, the CV RO based on matrix have a more concise form, and is easier to programming, besides, this method suits for general heading angle.
(2) Applying the excellent extraction and matching algorithm SIFT and RANSAC robust estimation method in vehicle-based sequential images processing, the experiments show this method has high stability.
(3) Now the precision of RO based on CV is limited, but the precision is relative high.
The quality of RO mainly relates to the quality of the images and matching algorithm. With the development of computer hardware and the popularize of parallel arithmetic, the time of extraction and matching will be shorter, this method is expected to be used in indoor visual navigation or provide initial RO value in position field, which has little require on time and precision. Moreover, after rigorous calibration, the quality can be greatly improved.
RO is a fundamental question in photogrammetry and CV. RO based on vision is a method with low cost and relatively high reliability, so in traditional photogrammetry and CV, RO based on vision privilege over other method. But, the precision of the result by this method is not very high, and there are many factors that affect the result by vision method, including the quality of the image, the precision of camera calibration parameters, the quality of image matching, data processing and its optimization etc. Under the condition that the quality of the image is unchangeable, the quality of the result can be improved purely by improving the data processing method and algorithm, but the extent is limited. For example, the homologous points acquired by SIFT were not identified by correlation coefficient, the precision cannot be guaranteed. It is hard to guarantee large overlap (the overlap over 60%) (LI Jinwen and ZHAN Zongqian,2010), in order to meet the demand of multiple view matching and bundle adjustment. Besides, there are many uncertain factors in environment; these are important factors that influence the RO based on vision, with the improvement of relative technology, the quality of RO will be improved.

## Acknowledgements

I would like to thank Beijing maggroup group for the vehicle based sequential image data of Landmark MMS. Also I appreciate the enthusiastic help of Cheng Xu, manager of the company, Doc.Wang Yanzheng and Liu Hua, postgraduate of School of Geodesy and Geomatics, Wuhan University.

## References

[1] LI Jian et al., Best feature point matching method research and application in automatic relative orientation [J]. Science of Surveying and Mapping .2010. 4
[2] Schenk, T., Toth, C., 1993. Towards a fully automated aerotrian-gulation system. Proc. ACSMrASPRS Annual Spring Convention, New Orleans, vol. 3, pp. 340–347.
[3] D G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints [J]. International Journal of Computer Vision, 2004. 60(2): 91-110
[4]FISCHLER M A,BOLLES R C.Random sample consensus:A Paradigm formodel fitting with applications to image analysis and automated cartography[J].Communications ACM,1981,24(6): 381-395.
[5] McGlone, J.C., Mikhail, E.M., Bethel, J., Mullen, R. (Eds.), 2004.Manual of Photogrammetry, fifth ed. American Society of Photogrammetry and Remote Sensing. ISBN: 1-57083-071-1.
[6] Nistér, D., 2003. Preemptive RANSAC for live structure and motion estimation. Proc. of the IEEE International Conference on Computer Vision, pp. 199–206.
[7] ZHANG Zuxun. Digital Photogrammetry and Computer Vision [J]. Editorial Board of Geomatics and Information Science of Wuhan University.2004.12
[8] Hartley, R; Zisserman, A. "Multiple View Geometry in Computer Vision." Cambridge University Press, 2003.
[9] DING Meng; CAO Yun-feng. Method to obtain UAV's flight attitude based on computer vision [J]. Transducer and Microsystem Technologies.2007.26(11).
[10] LI Jinwen,ZHAN Zongqian. Wide relief automated image stitching [J]. Chinese Museum.2010.01