Geometrical properties between 2-D image plane and 3-D error space

Sung Shik KOH^{1a)}, Thi Thi Zin¹, and Hiromitsu Hama¹

¹ Graduate School of Engineering, Osaka City University

3–3–138, Sugimoto, Sumiyoshiku, Osaka 558–8585, Japan

a) koh@sys.info.eng.osaka-cu.ac.jp

Abstract: We study geometrical properties between 2-D image plane and 3-D error space under affine reconstruction. The purpose of our system is to contribute to more accurate 3-D reconstruction by analyzing geometrically 2-D to 3-D relationship. In situation for no missing feature points and no noise in the 2-D observation matrix, the accurate solution is known to be provided by Singular Value Decomposition. However, several feature points of the matrix have not been observed because of occlusions and low image resolution, and so on. In this case, there is no simple solution. To obtain accurate 3-D reconstruction by recovering missing feature points, we propose the analytic approach which can handle the error orientation and distance of missing feature points by the geometrical properties.

Keywords: geometrical properties, SVD factorization, 3-D error space, missing feature point

Classification: Science and engineering for electronics

References

- [1] M. Irani and P. Anandan, “Factorization with uncertainty,” *Proc. ECCV*, Dublin, Ireland, pp. 539–553, 2000.
- [2] R. Szeliski and S. B. Kang, “Shape ambiguities in structure-from-motion,” *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, pp. 506–512, 1997.
- [3] C. Tomasi and T. Kanade “Shape and motion from image streams under orthography: A factorization method,” *Int. J. Computer Vision*, vol. 9, no. 2, pp. 137–154, 1992.
- [4] D. Jacobs, “Linear fitting with missing data for structure-from-motion,” *Proc. CVIU*, vol. 82, pp. 57–81, 2001.
- [5] R. F. C. Guerreiro and P. M. Q. Aguiar, “Estimation of Rank Deficient Matrices from Partial Observations: Two-Step Iterative Algorithms,” *Proc. EMMCVPR*, pp. 450–466, 2003.
- [6] C. J. Harris and M. Stephens, “A Combined Corner and Edge Detector,” *Proc. Alvey Vision Conf.*, pp. 147–151, 1998.

1 Introduction

In computer vision, the matrix factorization methods have been widely used to reconstruct a 3-D shape and camera motion from 2-D image data matrices. In real video clips, these projections are not visible along the entire image sequence due to occlusion and a limited field of view [1, 2]. Thus, the observation matrix is incomplete. Tomasi and Kanade [3] proposed a heuristic approach according to growing by one row or by one column at a time after decomposing a full sub-matrix with missing elements by Singular Value Decomposition (SVD) factorization method. However, their method met a NP-hard problem of finding the largest full sub-matrix of a matrix with missing elements. Jacobs [4] improved their method by fitting an unknown matrix of a certain rank to an incomplete noisy matrix resulting from measurements in images. However, he presented the method of determining whether the incomplete matrix is stable or unstable. Guerreiro and Aguiar [5] proposed the expectation-maximization and two-step iterative algorithm. The algorithm converged to the global optimum in a very small number of iterations. However, the performance of them cannot fit missing feature points (FPs) for a few image sequences. The above algorithms are all to solve the missing FPs, however these make use of probability or linear approach. Against existing methods, we propose the new analytic approach so that it can handle directly the error orientation and distance of missing FP by geometrical property in order to apply to 3-D reconstruction fields of computer vision.

2 Geometrical properties

The main limitation for evaluation of the 3-D reconstruction is inability to compare against the real shape. The reason for that is the difficulty of obtaining the 3-D ground truth. But we need some criterions. What to do? The answer is using the 3-D error space newly introduced here instead of the distance between the true shape and the reconstructed one. Hence, we first define the 3-D error space and analyze some specific geometry between the perturbation of the missing FPs on 2-D image plane and the error of the 3-D reconstructed shape.

2.1 3-D error space

In order to evaluate the precision of the 3-D shape reconstruction from an observation matrix, the 3-D error space is introduced as follows:

- (a) For recovering a missing FP ($\mathbf{p}_m$), its initial position ($\mathbf{p}_t + \Delta e$) is first fit roughly, and three FPs ($\mathbf{p}_{bA}$, $\mathbf{p}_{bB}$, and $\mathbf{p}_{bC}$) are randomly selected, which are called bias FPs being neighbors of the missing FP ($\mathbf{p}_m$), on the same 2-D image plane. Next, new FPs ($\mathbf{q}_i$) are added, which are called Reference Points (RPs), on a circular pattern ($r = \Delta c_i$) centering on the missing FP ($\mathbf{p}_m$). The aspects are shown in Fig. 1 (a), where Π_2 is a reference plane composed of $\mathbf{q}_i$ on 2-D image plane and $\overrightarrow{\mathbf{p}_m \mathbf{q}_i}$ is a reference vector (RV) composed of $\mathbf{p}_m$ and $\mathbf{q}_i$ on 2-D image plane.

- (b) Using affine SVD factorization, the roughly fitted FP ($\mathbf{p}_m = \mathbf{p}_t + \Delta e$), three bias FPs ($\mathbf{p}_{bA}, \mathbf{p}_{bB}$, and $\mathbf{p}_{bC}$), and circular RPs ($\mathbf{q}_i$) are reconstructed to $\mathbf{P}_m^*$, ($\mathbf{P}_{bA}^*, \mathbf{P}_{bB}^*$, and $\mathbf{P}_{bC}^*$), and $\mathbf{Q}_i^*$ on the 3-D reconstruction space, respectively (see Fig. 1 (b)). The symbol “*” means a perturbation.
- (c) A 3-D error space is defined as the coordinates of 3-D point vectors which are transformed from three distances between three bias FPs ($\mathbf{P}_{bA}^*, \mathbf{P}_{bB}^*$, and $\mathbf{P}_{bC}^*$) and a FP ($\mathbf{P}^*$) on 3-D reconstruction space. For example, a missing FP ($\mathbf{P}_m^*$) on 3-D reconstruction space can be represented to $\mathbf{P}_m$ on 3-D error space from three Euclidean distances between three bias FPs and the missing FP as follows:

$$\mathbf{P}_m = (L_A, L_B, L_C), \quad (1)$$

where $L_A = \|\overrightarrow{\mathbf{P}_m^* \mathbf{P}_{bA}^*}\|$, $L_B = \|\overrightarrow{\mathbf{P}_m^* \mathbf{P}_{bB}^*}\|$, and $L_C = \|\overrightarrow{\mathbf{P}_m^* \mathbf{P}_{bC}^*}\|$.

Hence, circular RPs ($\mathbf{Q}_i^*$) can be also expressed as $\mathbf{Q}_i$ on the 3-D error space as follows:

$$\mathbf{Q}_i = (L_{Ai}, L_{Bi}, L_{Ci}), \quad (2)$$

where $L_{Ai} = \|\overrightarrow{\mathbf{P}_m^* \mathbf{P}_{bA}^*}\|$, $L_{Bi} = \|\overrightarrow{\mathbf{P}_m^* \mathbf{P}_{bB}^*}\|$, and $L_{Ci} = \|\overrightarrow{\mathbf{P}_m^* \mathbf{P}_{bC}^*}\|$, $i = \{1, 2, \dots, Z\}$.

Fig. 1. Comparison of some parameters on 2-D image plane, 3-D reconstruction space, and 3-D error space. (a) missing FP, three bias FPs and circular RPs on 2-D image plane. (b) parameters reconstructed from (a) on the 3-D reconstruction space. (c) parameters transformed from (b) on the 3-D error space.

This aspects is shown in Fig. 1 (c), where Π_3 is a reference plane composed of $\mathbf{Q}_i$ on the 3-D error space and $\overrightarrow{P_m \mathbf{Q}_i}$ is a Reference Vector (RV) composed of $\mathbf{P}_m$ and $\mathbf{Q}_i$ on the 3-D error space.

2.2 Geometrical properties between two reference planes Π_2 and Π_3

Assume that the error vectors of a 2-D FP and its 3-D reconstructed FP are perpendicular to the camera optic axis and have approximately the same orientation and the nearly proportional size to each other. According to our considerations, the relationship between two reference planes Π_2 and Π_3 are analyzed on the 2-D image plane and the 3-D error space. So, we made the hypothesis as follow:

Plane: Since 2-D RVs are on a plane Π_2 , the 3-D RVs are also approximately located on a plane Π_3 .

$$\Pi_2 = \{ \overrightarrow{p_m \mathbf{q}_i}, i = 1, 2, \dots, Z \}, \quad (3)$$

$$\Pi_3 \cong \{ \overrightarrow{P_m \mathbf{Q}_i}, i = 1, 2, \dots, Z \}. \quad (4)$$

Pattern: If the RPs on Π_2 are distributed on a circular pattern, then the RPs on Π_3 are distributed on an ellipse and are very close to be circular.

$$\{ \mathbf{q}_i \} : \text{circular pattern on } \Pi_2, \quad (5)$$

$$\{ \mathbf{Q}_i \} : \text{ellipse pattern on } \Pi_3. \quad (6)$$

Symmetry: If two of the RVs on Π_2 exist on symmetrical positions with $\mathbf{p}_m$, then their positions on Π_3 are nearly symmetric.

$$\overrightarrow{p_m \mathbf{q}_{i_a}} \cong -\overrightarrow{p_m \mathbf{q}_{i_b}} \Leftrightarrow \overrightarrow{P_m \mathbf{Q}_{i_a}} \cong -\overrightarrow{P_m \mathbf{Q}_{i_b}}. \quad (7)$$

Size: If two of the RVs on Π_2 are in the same direction with different sizes, then the RVs on Π_3 keep their magnitude relationships and ratios relative to the size.

$$\| \overrightarrow{p_m \mathbf{q}_{i_a}} \| < \| \overrightarrow{p_m \mathbf{q}_{i_b}} \| \Leftrightarrow \| \overrightarrow{P_m \mathbf{Q}_{i_a}} \| < \| \overrightarrow{P_m \mathbf{Q}_{i_b}} \|, \quad (8)$$

$$\frac{\| \overrightarrow{p_m \mathbf{q}_{i_a}} \|}{\| \overrightarrow{p_m \mathbf{q}_{i_b}} \|} \cong \frac{\| \overrightarrow{P_m \mathbf{Q}_{i_a}} \|}{\| \overrightarrow{P_m \mathbf{Q}_{i_b}} \|}. \quad (9)$$

Angle: If three RVs on Π_2 are arranged in some angles, then the RVs on Π_3 are also arranged similarly, while keeping the relationship of magnitude and ratio around the angle.

$$\angle \mathbf{q}_i \mathbf{p}_m \mathbf{q}_{i+1} < \angle \mathbf{q}_i \mathbf{p}_m \mathbf{q}_{i+2} \Leftrightarrow \angle \mathbf{Q}_i \mathbf{P}_m \mathbf{Q}_{i+1} < \angle \mathbf{Q}_i \mathbf{P}_m \mathbf{Q}_{i+2}. \quad (10)$$

These facts have been confirmed through the experimental results from synthetic cube and real data are described in Section 3.

3 Experimental results

In this section, we describe the experiments that illustrate the behavior of our algorithm using synthetic and real video sequence. For synthetic cube, we preset camera configuration as shown in Fig. 2 (a), a cube is set on a 3-D world with a set of cameras. The size of the cube is $1 \times 1 \times 1$ [unit], and the cube contains eight 3-D corner points. All points are tracked from twenty image frames taken to cover 180 degrees with the same intrinsic parameters. The cube is placed with its centroid at 2.5 [unit] from the first camera. The cameras are pointed towards the centroid of the cube. In order to evaluate the geometrical properties, we preset some parameters as follows: (i) the 3rd FP is missing on the 12th image. (ii) the initial position ($\mathbf{p}_t + \Delta \mathbf{e}$) is located at $(x + 0.3, y - 0.3)$ [unit]. (iii) the radius of circular 20 RPs is set up to $\Delta c_i = 0.3$ [unit]. The number of RPs has no limitation because the proposed algorithm is able to use the geometrical properties if the number of reference points is greater than three (see Fig. 1). The geometrical relationship of the missing FP on the 12th image frame is shown in Fig. 2 (b) and (c), which illustrate one missing FP and twenty RPs on Π_2 and Π_3 , respectively.

Fig. 2. Camera configuration and geometrical relationship on the synthetic cube. (a) 3-D cube and camera configuration (diamond markers). (b) one missing FP and twenty RPs on Π_2 . (c) one missing FP and twenty RPs on Π_3 .

Moreover, in order to confirm the effectiveness of the proposed method, real data image sequences are treated. The FPs are tracked over the 36-frame ‘dinosaur’ sequence. The 3rd image frame (720×576) is shown in Fig. 3 (a),

where the tracked FPs are denoted by symbol ‘+’ and extracted by the Harris interest operator [6]. Also for real data, we preset some parameters as follows: (i) the 11th FP is one of three missing FPs on the 3rd image. (ii) the initial position ($\mathbf{p}_t + \Delta \mathbf{e}$) is located at $(x + 20, y - 20)$ [pixel]. (iii) the radius of circular $\Delta c_i = 20$ [pixel]. To illustrate the geometrical relationship, one missing FP and twenty RPs on Π_2 and Π_3 are described in Fig. 3 (b) and (c), respectively.

Fig. 3. The 3rd image frame and geometrical relationship on the real data. (a) the tracked FPs (‘+’) on the 3rd image frame. (b) one missing FP (yellow square) and twenty RPs (blue points) on Π_2 . (c) one missing FP and twenty RPs on Π_3 .

Eq. (11) to (16) represent the evaluations for all geometrical properties. In order to compare with synthetic and real data, the scale of real data is normalized to 1 [unit].

$$e_1 = \sum_{i=1}^Z \min_{Q \in \Pi_3} \|Q_i - Q\|, \quad (11)$$

$$e_2 = \sum_{i=1}^Z \min_{Q \in \xi} \|Q_i - Q\|, \quad \xi = \text{ellipse}, \quad (12)$$

$$e_{31} = \sum_{i_a=1}^{Z/2} \left\| \frac{\overrightarrow{P_m Q_{i_a}}}{\overrightarrow{P_m Q_{i_b}}} - 1 \right\|, \quad (13)$$

$$e_{32} = \sum_{i_a=1}^{Z/2} \left(\angle \mathbf{Q}_{i_a} \mathbf{P}_m \mathbf{Q}_{i_b} - \pi \right), i_b = i_a + \frac{Z}{2}, \quad (14)$$

$$e_4 = \sum_{i_a=1}^{Z/2} \left\| \frac{\|\overrightarrow{\mathbf{p}_m \mathbf{q}_{i_a}}\|}{\|\overrightarrow{\mathbf{p}_m \mathbf{q}_{i_b}}\|} / \frac{\|\overrightarrow{\mathbf{P}_m \mathbf{Q}_{i_a}}\|}{\|\overrightarrow{\mathbf{P}_m \mathbf{Q}_{i_b}}\|} - 1 \right\|, \quad (15)$$

$$e_5 = \sum_{i=1}^{Z-1} \left(\angle \mathbf{Q}_i \mathbf{P}_m \mathbf{Q}_{i+1} - \frac{2\pi}{Z} \right). \quad (16)$$

The results of accuracy for the geometrical properties as follows: For synthetic data, the mean errors of e_1 , e_2 , e_{31} , e_{32} , e_4 , and e_5 are 0.0027, 0.0022, 0.0010, 3.19°, 0.0028, and 0.13°, respectively. For real data, they are 0.0105, 0.0034, 0.0017, 3.47°, 0.0049, and 0.17°, respectively. According to these experimental results, it can be verified that the proposed geometrical properties are inevitably right because there are within a very few error rates distributed for plane, pattern, symmetry, size and angle between Π_2 and Π_3 .

4 Conclusion

We have presented the interesting geometrical properties between 2-D image plane and 3-D error space under affine projection. In order to contribute to more accurate 3-D reconstruction, we focused on analyzing 2-D to 3-D relationship geometrically. The accuracy of the proposed approach is evaluated for synthetic and real video sequence, respectively. As the results, it can be verified that our geometrical properties decrease to very low error rates. The result means that the proposed approach can provide very well to recover the missing FPs.

In the near future, we apply the proposed geometrical properties to recover missing FPs. Furthermore, we will extend this proposed system to real world applications, under noise distribution, 2-D video sequence to 3-D TV conversion techniques based on structure from motion approaches which can be used to convert already recorded monoscopic video material to 3D content.

Acknowledgments

This work was supported by the Korea Research Foundation Grant funded by Korea Government (MOEHRD, Basic Research Promotion Fund) (M01-2004-000-20052-0).