Conditions for Segmentation of 2D Translations of 3D Objects

Shafriza Nisha Basah[1], Alireza Bab-Hadiashar[1], and Reza Hoseinnezhad[2]

[1] Faculty of Engineering and Industrial Sciences, Swinburne University of Technology
[2] Melbourne School of Engineering, The University of Melbourne,
Victoria, Australia
{sbasah,abab-hadiashar}@swin.edu.au, rezah@unimelb.edu.au

Abstract. Various computer vision applications involve recovery and estimation of multiple motions from images of dynamic scenes. The exact nature of objects' motions and the camera parameters are often not known a priori and therefore, the most general motion model (the fundamental matrix) is applied. Although the estimation of a fundamental matrix and its use for motion segmentation are well understood, the conditions governing the feasibility of segmentation for different types of motions are yet to be discovered. In this paper, we study the feasibility of separating 2D translations of 3D objects in a dynamic scene. We show that successful segmentation of 2D translations depends on the magnitude of the translations, average distance between the camera and objects, focal length of the camera and level of noise. Extensive set of controlled experiments using both synthetic and real images were conducted to show the validity of the proposed constraints. In addition, we quantified the conditions for successful segmentation of 2D translations in terms of the magnitude of those translations, the average distance between the camera and objects in motions for a given camera. These results are of particular importance for practitioners designing solutions for computer vision problems.

Keywords: Motion segmentation, multibody structure-and-motion, fundamental matrix, robust estimation.

1 Introduction

Recovering structure-and-motion (SaM) from images of dynamic scenes is an indispensable part of many computer vision applications ranging from local navigation of a mobile robot to image rendering in multimedia applications. The main problem in SaM recovery is that the exact nature of objects' motions and the camera parameters are often not known a priori. Thus, a pure translation of 3D points needs to be modelled in the form of a fundamental matrix [12] (in the case where all moving points are in the same plane, the motion can be modelled as a homography). Motion estimation and segmentation based on the fundamental matrix are well understood and solved in the established work presented and

P. Foggia, C. Sansone, and M. Vento (Eds.): ICIAP 2009, LNCS 5716, pp. 82–91, 2009.

summarised in (Chpt.9-12,[6]). Soon after that work, researchers resumed to the more challenging multibody structure-and-motion (termed MSaM by Schindler and Suter in [10]) where multiple objects in motions need to be concurrently estimated and segmented. However, the conditions governing the feasibility of segmentation involving MSaM for different types of motions are yet to be established. These conditions are important as they provide information on the limits of current MSaM methods and would provide useful guidelines for practitioners designing solutions for computer vision problems.

Well known examples of previous works in motion segmentation using the fundamental matrix are by Torr [11], Vidal et.al [14] and Schindler and Suter [10]. Torr uses the fundamental matrix to estimate an object in motion and cluster it in a probabilistic framework using the Expectation Maximisation algorithm [11]. Vidal.et.al proposes to estimate the number of moving objects in motion and cluster those motions using the multibody fundamental matrix; the generalization of the epipolar constraint and the fundamental matrix of multiple motions [14]. Schindler and Suter have implemented the geometric model selection to replace degenerate motion in a dynamic scene using multibody fundamental matrix [10].

The focus of this paper is to study the feasibility to detect and segment an unknown 2D translation (of a rigid 3D object) in a dynamic scene (with images taken by an uncalibrated camera). The conditions for motion-background segmentation are established and provided by Basah et.al in [3]. In Section 2, we derive the conditions to detect and segment multiple 2D translations and provide quantitative measures for detectable translations using theoretical analysis. Section 3 details the experiments using synthetic and real images conducted to verify the theoretical analysis and the proposed conditions for successful segmentation of 2D translations. Section 4 concludes the paper.

1.1 Preliminary Information and Notations

This sections recalls some notations and equations governing motion segmentation using the fundamental matrix. Let $[X, Y, Z]^T$ and T refer to the location of a point and a translation in 3D world coordinate system and $[x, y]^T$ and t are their projected locations on a 2D image plane. The 3×3 rank 2 fundamental matrix F relates the epipolar geometry of homogeneous image points $m_{1i} = [x_{1i}, y_{1i}, 1]^T$ undergoing rotation and non-zero translation in world coordinate system to $m_{2i} = [x_{2i}, y_{2i}, 1]^T$ as [1,6,13,16]:

$$m_{2i}^T F m_{1i} = 0. \tag{1}$$

The image sequence generally contains noise and mismatches which result in uncertainty and errors in the segmentation result. Thus, robust estimators are usually applied. There are numerous robust estimators developed for motion estimation and segmentation. Recent examples of robust estimators are pbM-estimator [5], TSSE [15] and MSSE [2]. The segmentation step of MSSE [2] is

used because of its desired asymptotic and finite sample bias properties [7]. Although we use MSSE, the analysis is general and similar results will be obtained if other robust estimators are used.

The error measure is defined as a function of the distances such that the error is minimum for the target object (inliers) and larger for the rest (outliers). Thus, segmentation is based on automatic separation of small distances from the large ones. Four commonly used error measures to estimate F are the algebraic distance [12], geometric distance [6], Sampson distance [12,13] and Luong distance [9]. We adopt the Sampson distance [12,13]:

$$d_i = \frac{m_{2i}^T F m_{1i}}{\sqrt{\left[\frac{\partial}{\partial x_{1i}}^2 + \frac{\partial}{\partial y_{1i}}^2 + \frac{\partial}{\partial x_{2i}}^2 + \frac{\partial}{\partial y_{2i}}^2\right] m_{2i}^T F m_{1i}}}, \tag{2}$$

where d_i denotes the distance of the i-th matching points in the image from its epipolar lines. Sampson distance provides first order approximation to the geometric distance and it is computationally cheaper than geometric distance[12,13] and gives slightly better estimation results than Luong distance [16].

2 Separability Conditions for 2D Translations

Consider $n = n_i + n_o$ feature points belonging to two 3D objects $[X_i, Y_i, Z_i]^T$ undergoing pure translations in $X - Y$ plane denoted by T_a and T_b where $T_a = [T_{xa}, T_{ya}, T_{za}]^T$ and $T_b = [T_{xb}, T_{yb}, T_{zb}]^T$ with $T_{za} = T_{zb} = 0$ ($i = 0, 1 \ldots n_i$ denote the points belonging to T_a and $i = n_i + 1, n_i + 2 \ldots n$ denote points belonging to T_b). The location $[X_i, Y_i, Z_i]^T$ before and after the translations are visible on the image plane and are denoted by $[x_{1i}, y_{1i}]^T$ and $[x_{2i}, y_{2i}]^T$. All points in the image plane are contaminated by measurement noise e assumed to be independent and identically distributed (i.i.d) with Gaussian distribution:

$$x_{1i} = \underline{x}_{1i} + e_{ix}^1, \quad y_{1i} = \underline{y}_{1i} + e_{iy}^1, \quad x_{2i} = \underline{x}_{2i} + e_{ix}^2, \quad \text{and} \quad y_{2i} = \underline{y}_{2i} + e_{iy}^2, \tag{3}$$

where e_{ix}^1, e_{iy}^1, e_{ix}^2 and $e_{iy}^2 \sim N(0, \sigma_n^2)$ and σ_n is the unknown scale of noise. The underlined variables denote the true or noise-free locations of points in image plane. The relationship between all true matching points in the image plane and world coordinate points are:

$$\underline{x}_{1i} = \frac{fX_i}{Z_i}, \quad \underline{y}_{1i} = \frac{fY_i}{Z_i}, \quad \underline{x}_{2i} = \underline{x}_{1i} + \frac{fT_x}{Z_i} \quad \text{and} \quad \underline{y}_{2i} = \underline{y}_{1i} + \frac{fT_y}{Z_i}, \tag{4}$$

where f is the focal length of the camera matrix (equal focal length in X and Y directions is used and the offset distances are assumed zero for simplicity), $T_x = T_{xa}$; $T_y = T_{ya}$ for T_a and $T_x = T_{xb}$; $T_y = T_{yb}$ for T_b. In this scenario, we aim to segment matching points belonging to T_a from the mixture of matching

points belonging to T_a and T_b in two images, thus the points undergoing T_a are the inliers and the points undergoing T_b are the outliers.

The fundamental matrix of point belonging to T_a is computed using:

$$F = A^{-T}[T]_x R A^{-1}, \tag{5}$$

where A is the camera matrix, R is the rotation matrix of the motion and $[T]_x$ is the skew symmetric matrix or a null vector of translation T [1,6,16]. For $T = [T_{xa}, T_{ya}, 0]^T$ and $R = I_3$ (representing zero rotation), equation (5) yields:

$$F = \frac{1}{f}\begin{pmatrix} 0 & 0 & T_{ya} \\ 0 & 0 & -T_{xa} \\ -T_{ya} & T_{xa} & 0 \end{pmatrix}. \tag{6}$$

If the fundamental matrix for T_a is known, d_i can be computed using (2). Substitution of real plus noise forms in (3) and true F in (6), yields:

$$d_i = \frac{T_{ya}(\underline{x}_{2i} + e^2_{ix} - \underline{x}_{1i} - e^1_{ix}) + T_{xa}(\underline{y}_{1i} + e^1_{iy} - \underline{y}_{2i} - e^2_{iy})}{\sqrt{2(T^2_{ya} + T^2_{xa})}}. \tag{7}$$

For points belonging to T_a $(i = 0, 1 \ldots n_i)$, the expression without noise terms in (7) equal to zero according to equation (1) because the true F of T_a is used to compute the distances. Thus, equation (7) is simplified to:

$$d_i = \frac{T_{ya}(e^2_{ix} - e^1_{ix}) + T_{xa}(e^1_{iy} - e^2_{iy})}{\sqrt{2(T^2_{ya} + T^2_{xa})}} \sim e(N(0, \sigma_n^2)). \tag{8}$$

Distances d_i's of the points belonging to T_a in (8) turn out to be a linear combination of the i.i.d. noises therefore, they are also normally distributed with zero mean. The variance of d_i's $(i = 0, 1 \ldots n_i)$ also equals σ_n^2 as the numerator and denominator cancel each other. The points belonging to T_a are to be separated from the points belonging to T_b. Thus, d_i's of points belonging to T_b $(i = n_i + 1, n_i + 2 \ldots n)$ with respect to F of T_a is calculated using (7) where $\underline{x}_{2i}$ and $\underline{y}_{2i}$ are replaced with the terms given in (4) for T_b yields:

$$d_i = \frac{f(T_{ya}T_{xb} - T_{xa}T_{yb})}{Z_i\sqrt{2(T^2_{ya} + T^2_{xa})}} + e, \tag{9}$$

where $e \sim N(0, \sigma_n^2)$. Based on (9) and the normality assumption, 99.4% of the points belonging to T_a will be correctly segmented from points belonging to T_b if the following condition is satisfied[1]:

$$\frac{|W|}{Z_i} \geq 5, \quad \text{where} \quad W = \frac{f(T_{ya}T_{xb} - T_{xa}T_{yb})}{\sqrt{2(T^2_{ya} + T^2_{xa})}\sigma_n}. \tag{10}$$

[1] From probability theory it is well known that if the means of two normal populations with the same variance σ^2 are at least 5σ away, then only 0.6% of the points of each population will overlap.

Alternatively W could be express in term of the direction of T_a and T_b as:

$$W = \frac{f}{\sqrt{2}\sigma_n}\sqrt{T_{yb}^2 + T_{xb}^2}\sin(\phi_a - \phi_b), \tag{11}$$

where $\phi_a - \phi_b$ are the angle between the T_a and T_b.

In most computer vision problems, the distance between the camera and the object in motion is roughly known. Therefore, the condition for segmentation in (10) can be expressed in term of average distance between camera and translating objects $\bar{Z}$:

$$\frac{|W|}{Z_i} \approx \frac{|W|}{\bar{Z}} \geq 5. \tag{12}$$

We assume an accurate estimate for F of T_a is given by minimising the cost function of a robust estimator. Having F of T_a, the distances d_i's of all matching points are computed. Then d_i^2's for all points are used as residuals for segmentation to segment points belonging to T_a using MSSE. In MSSE, d_i^2's are sorted in an ascending order and the scale estimate given by the smallest k distances is calculated using [2]:

$$\sigma_k^2 = \frac{\sum_{i=1}^{k} d_i^2}{k-1}. \tag{13}$$

While incrementing k, d_{k+1} is detected as the distance of the first outlier if it is larger than 2.5 times the scale estimate given by the smallest k distances:

$$d_{k+1}^2 > 2.5^2\sigma_k^2. \tag{14}$$

With the above threshold, at least 99.4% of the inliers will be segmented if there are normally distributed [2].

From our analysis, the separability of 2D translations depends on the magnitude of $|W|/\bar{Z}$ in (12). This is the basis of the Monte Carlo experiments presented in Section 3.1. We also aim to determine the condition for segmentation in term of minimum $|W|/\bar{Z}$ required for successful segmentation. The correctness of these conditions will be verified by studying the variance of the result of the Monte Carlo experiments. The condition for segmentation of 3D translations are too complex to be derived theoretically. However, the derived condition for segmentation of 2D translations is served as the basis of approximation for 3D translations when T_z are very small or close to zero.

3 Experiments

3.1 Monte Carlo Experiments with Synthetic Images

The Monte Carlo experiments with synthetic images were divided into two parts. The first part was to verify the separability condition in (12) for separating T_a from T_b. The second part of the experiments was designed to determine the

conditions for successful segmentation in term of minimum $|W|/\bar{Z}$ required when the inlier ratio ϵ, noise level σ_n and the size of the objects in motions were varied.

In each iteration in those experiments, 2000 pairs of points in the world coordinate according to T_a (the inliers) were mixed with the pairs of matching points according to T_b (the number of matching points depends on ϵ). The X and Y coordinates of the matching points were randomly generated while Z coordinates were uniformly distributed according to $U(\bar{Z} - \sigma_Z, \bar{Z} + \sigma_Z)$. T_a and T_b (both with $T_z = 0$) were randomly selected based on $|W|/\bar{Z}$. Then, all matching points were projected to two images using a synthetic camera with equal focal length of 703 and offset distance of 256. Random noise with the distribution of $N(0, \sigma_n^2)$ were added to all points. Sampson distances were calculated using equation (2) based on the true F of T_a. After that, segmentation was performed using MSSE with d_i^2's as the segmentation residuals. The ratio of the number of segmented inliers over the true inliers ζ was calculated and recorded. Each experimental trial consists of 1000 experimental iterations and the mean and sigma of 1000 ζ's were recorded. The experiment trials were then repeated for various $|W|/\bar{Z}$, ϵ, σ_n and $\sigma_Z/\bar{Z}$ (representing different sizes of objects in motion).

The first part of the experiment was conducted with two cases of T_a and T_b randomly selected such that $|W|/\bar{Z} = 2$ and $|W|/\bar{Z} = 6$. The inlier ratios ϵ were 60%, $\sigma_n = 1$ and $\sigma_Z/\bar{Z} = 10\%$. The histogram of d_i^2 for all image points (the residuals for segmentation) are plotted in Fig.1(a) and 1(b). The ratio of segmented over true inliers ζ is 1.644 for $|W|/\bar{Z} = 2$, which means that 3288 points were segmented as belonging to T_a ($\zeta = 1$ and 2000 points for correct segmentation). In addition, the distribution of d_i^2 of T_a and T_b were well mixed and couldn't be distinguished from each other (as shown in Fig.1(a)). These observations show that points of T_a were not separable from the points of T_b when $|W|/\bar{Z} = 2$. As $|W|/\bar{Z}$ increased to 6, ζ reduced to 0.988 (1996 points segmented as belonging to T_a). Furthermore, the residuals (d_i^2's) of points belonging to T_a were easily distinguished from points belonging to T_b as shown in Fig.1(b). In this case, points belonging to T_a have successfully been identified and segmented. The result of possible segmentation when $|W|/\bar{Z} > 5$ in this case is consistent with our earlier separability analysis in (10) and (12).

In the second part of the experiments, the effect of varying parameters including $|W|/\bar{Z}$ (from 0 to 10), ϵ (from 30% to 80%), σ_n (from 0.25 to 1) and $\sigma_Z/\bar{Z}$ (from 5% to 20%) were examined. The mean and sigma of 1000 ζ's in each trial were recorded. Fig.1(c) and 1(d) show ζ versus $|W|/\bar{Z}$ when $\epsilon = 80\%$ and 50% while $\sigma_n = 1$ and $\sigma_Z/\bar{Z} = 10\%$. It was observed that, for small $|W|/\bar{Z}$ some points from T_b were mixed with points from T_a and segmented ($\zeta > 1$). In such cases, an inaccurate inlier-outlier dichotomy would result in an incorrect motion estimation and segmentation. As $|W|/\bar{Z}$ increased from 0 to 10, the mean of ζ reduced to 0.99 indicating accurate segmentation of T_a. Moreover, the very small sigma of ζ as $|W|/\bar{Z}$ increased indicated the consistency of the segmentation accuracy. The conditions for segmentation were then determined by interpolating $|W|/\bar{Z}$ when $\zeta = 0.994$ from Fig.1(c) and Fig.1(d). We found out that, the required condition for segmentation were $|W|/\bar{Z} \geq 4.25$ for $\epsilon = 80\%$

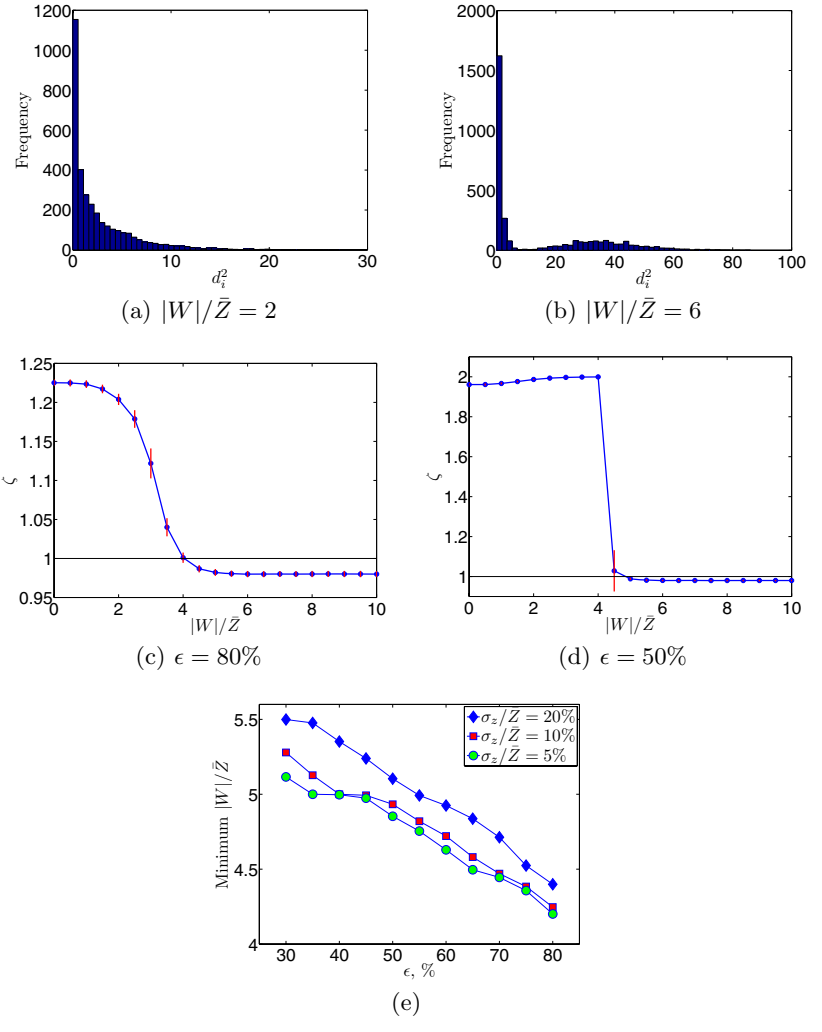
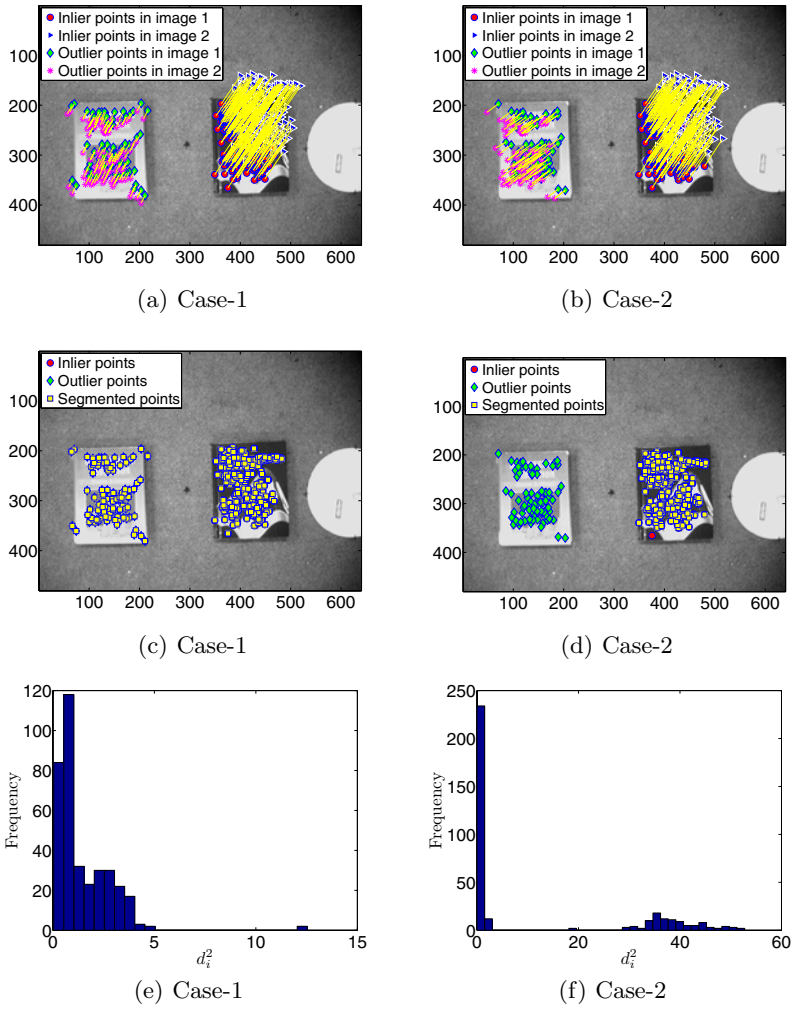


Fig. 1. Histogram of d_i^2's for all image points for ($T_a = [-5.4, 22.6, 0]$cm, $T_b = [15.1, -46.2, 0]$cm and $\zeta = 1.644$) in (a) and ($T_a = [-9.3, -10.5, 0]$cm, $T_b = [-22.8, -7.6, 0]$cm and $\zeta = 0.988$) in (b). The mean and sigma of ζ versus $|W|/\bar{Z}$ ($\sigma_n = 1$ and $\sigma_Z/\bar{Z} = 10\%$) are shown in (c) and (d). (e) shows the minimum $|W|/\bar{Z}$ required for segmentation for various σ_n and $\sigma_z/\bar{Z}$.

and $|W|/\bar{Z} \geq 4.93$ for $\epsilon = 50\%$ for 99.4% of T_a to be correctly segmented when $\sigma_n = 1$ and $\sigma_Z/\bar{Z} = 10\%$. The overall conditions for segmentation of T_a from T_b over ϵ, σ_n and $\sigma_Z/\bar{Z}$ are shown in Fig.1(e). It was observed that, the minimum $|W|/\bar{Z}$ for segmentations increased as ϵ decreased and $\sigma_Z/\bar{Z}$ increased and remain unchanged when σ_n was varied. More precisely, the segmentation of two translations becomes more difficult when more outliers were involved and the



(a) Case-1 (b) Case-2

(c) Case-1 (d) Case-2

(e) Case-1 (f) Case-2

Fig. 2. Corresponding image points (the black book undergoing translation T_a (inliers) and the other book translated according to T_b (outliers)) superimposed in image-1 for Case-1 (T_a=[4.9, -7.0, 0]cm and T_b=[-1.0, 2.0, 0]cm) in (a) and Case-2 (T_a=[4.9, -7.0, 0]cm and T_b=[-3.0, 2.0, 0]cm) in (b), the segmented inliers in (c) and (d) and the histogram of d_i^2 for all image points in (e) and (f)

size of object belonging to the outlier translation was bigger. This is because as the contamination of outliers in image points increased and the size of object (outliers) was bigger, the density of outlier residuals became more spread and the likelihood of some of them to be included in inlier segment increased resulting in a higher magnitude of minimum $|W|/\bar{Z}$ required for correct segmentation.

3.2 Experiments with Real Images

The experiment with real images was set up using a commercial camera in a laboratory environment. The camera was calibrated using a publicly available camera calibration toolbox by Bouguet [4]. Two objects were moved according to 2D translations T_a and T_b and the images before and after the translations were taken. Then the corresponding image points were extracted from each pair of images using SIFT algorithm by Lowe [8]. Incorrect matches in each pair of images and background points were manually eliminated and the true image points belonging to the object moved according to T_a and T_b were manually extracted (in this case, points belonging to T_a was the inlier while points belonging to T_b was the outlier). Then, the true F for T_a was calculated using (5) with known T_a and the camera matrix. The segmentation step according to MSSE in (14) was performed to all image points to segment T_a (based on the Sampson distance d_i in (2) calculated using the true F for T_a). The segmented points was then compared with the true points belonging to T_a. The ratio ζ was computed and the histogram of d_i^2's of all image points were plotted.

The sample results of the experiment are shown in Fig.2 where the black book undergoes translation T_a while the other book moved according to T_b. The noise level σ_n during the experiment was about 0.75 (estimated using (13) and the true d_i's of points belonging to T_a) and the distance between the objects and camera was about 1.5m. The value $|W|/\bar{Z}$ was calculated using (10) from the magnitude of T_a and T_b, estimated σ_n, $\bar{Z} = 1.5$ and $f = 950.5$ according to camera calibration. In Case-1, ($|W|/\bar{Z} = 1.94$ in Fig.2(a)) it is observed that from Fig.2(c) and $\zeta = 1.463$ that the points belonging to T_a were incorrectly segmented. This is because the d_i^2's (the residuals for segmentation) of T_a were not distinguishable from the d_i^2's of T_b as shown in Fig.2(e). As $|W|/\bar{Z}$ increased to 7.73 in Case-2 as shown in Fig.2(b), the points belonging to T_a were correctly segmented as shown in Fig.2(d) and the ratio $\zeta = 0.996$. The successful segmentation of T_a was possible because d_i^2's of T_a could be easily distinguished from d_i^2's of T_b in Fig.2(f). These experimental results and observations are in agreement with the results of Monte Carlo experiments in Fig.1(a) and (b) and verify the segmentation analysis of 2D translations in (10) and (12).

4 Conclusions

The conditions for successful segmentation of 2D translations using the fundamental matrix depends on the term $|W|/\bar{Z}$ (consists of the magnitude of the translations, average distance between the camera and objects, focal length of the camera and level of noise). This relationship was explored both by theoretical analysis and experimentation with synthetic and real images. Monte Carlo experiments using synthetic images provided the required conditions for successful segmentation of 2D translations in term of the minimum $|W|/\bar{Z}$ required when the inlier ratio, noise level and the size of the objects in motion were varied. The minimum $|W|/\bar{Z}$ required for segmentations increased as the inlier ratio decreased and the size of the objects (belonging to the outliers) increased and

remains unchanged when the noise level was varied. This is because, when the outlier contamination increased and became more spread, the likelihood for an outlier to be segmented as inlier increased resulting in a higher magnitude of minimum $|W|/\bar{Z}$ required for successful segmentation.

References

1. Armangue, X., Salvi, J.: Overall view regarding fundamental matrix estimation. Image and Vision Computing 21(2), 205–220 (2003)
2. Bab-Hadiashar, A., Suter, D.: Robust segmentation of visual data using ranked unbiased scale estimate. Robotica 17(6), 649–660 (1999)
3. Basah, S.N., Hoseinnezhad, R., Bab-Hadiashar, A.: Limits of Motion-Background Segmentation Using Fundamental Matrix Estimation. In: Proceedings of the Digital Image Computing: Techniques and Applications, pp. 250–256. IEEE Computer Society, California (2008)
4. Camera Calibration Toolbox for Matlab, `http://www.vision.caltech.edu/bouguetj/calib_doc/index.html`
5. Chen, H., Meer, P.: Robust regression with projection based M-estimators. In: Proceedings of the IEEE International Conference on Computer Vision (2003)
6. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision, 2nd edn. Cambridge University Press, Cambridge (2004)
7. Hoseinnezhad, R., Bab-Hadiashar, A.: Consistency of robust estimators in multi-structural visual data segmentation. Pattern Recognition 40(12), 3677–3690 (2007)
8. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. IJCV 60(2), 91–110 (2004)
9. Luong, Q.T., Deriche, R., Faugeras, O.D., Papadopoulo, T.: Determining The Fundamental Matrix: Analysis of Different Methods and Experimental Results. INRIA (1993)
10. Schindler, K., Suter, D.: Two-view multibody structure-and-motion with outliers through model selection. PAMI 28(6), 983–995 (2006)
11. Torr, P.H.S.: Motion Segmentation and Outlier Detection. Phd Thesis, Department of Engineering Science, University of Oxford 28(6), 983–995 (1995)
12. Torr, P.H.S., Zisserman, A., Maybank, S.: Robust detection of degenerate configurations while estimating the fundamental matrix. CVIU 71(3), 312–333 (1998)
13. Torr, P.H.S., Murray, D.W.: The Development and Comparison of Robust Methods for Estimating the Fundamental Matrix. IJCV 24(3), 271–300 (1997)
14. Vidal, R., Ma, Y., Soatto, S., Sastry, S.: Two-view multibody structure from motion. IJCV 68(1), 7–25 (2006)
15. Wang, H., Suter, D.: Robust adaptive-scale parametric model estimation for computer vision. PAMI 26(11), 1459–1474 (2004)
16. Zhang, Z.: Determining the Epipolar Geometry and its Uncertainty: A Review. IJCV 27(2), 161–195 (1998)