3D Motion Segmentation from Straight-Line Optical Flow

Jing Zhang[1], Fanhuai Shi[2], Jianhua Wang[1], and Yuncai Liu[1]

[1] Inst. Image Processing and Pattern Recognition,
[2] School of Materials Science and Engineering,
Shanghai Jiao Tong University,
Shanghai 200240, P.R. China
{zhjseraph,fhshi,whomliu}@sjtu.edu.cn

Abstract. We present a closed form solution to the problem of segmenting multiple 3D motion models from straight-line optical flow. We introduce the multibody line optical flow constraint(MLOFC), a polynomial equation relating motion models and line parameters. We show that the motion models can be obtained analytically as the derivative of the MLOFC at the corresponding line measurement, without knowing the motion model associated with that line. Experiments on real and synthetic sequences are also presented.

1 Introduction

Motion segmentation is a fundamental problem in many applications in computer vision, such as traffic surveillance, recognition of human gaits, etc.And recently, there has been an increasing interest on the algebraic, geometrical and statistical models to resolve this problem.

Existing approaches on 3-D motion segmentation from two views include different variations of the Expectation Maximization (EM) algorithm [1,2,3]. But EM is very sensitive to initialization [4].

Algebraic approaches, which are based on polynomial and tensor factorization, have been proposed [5]. In [6],Vidal and Ma viewed the estimation of multiple motion models as the estimation of a single multibody motion model and proposed a unified algebraic approach to 2-D and 3-D motion segmentation from two-view correspondences or optical flow, which applies to most of the two-view motion models adopted in computer vision. In [7,8], they introduce the multibody brightness constancy constraint(MBCC) and presented a closed form solution to direct motion segmentation.

Other related works include [9], which considers the problem of modeling a scene containing multiple dynamic textures undergoing multiple rigid-body motions. And in [10], X.Fan and R.Vidal study the rank and geometry of the multibody fundamental matrix.

In addition, efforts for the robustness of multibody motion segmentation include those works that uses RANSAC with priors to do articulated motion segmentation [11]. L. Hajder et.al. [12] consider robust 3D segmentation of multiple moving objects under weak perspective.

N. Sebe, Y. Liu, and Y. Zhuang (Eds.): MCAM 2007, LNCS 4577, pp. 85–94, 2007.

However, all the previous methods dealing with 3-D motion segmentation use point correspondences. Generally, the choice of types of features depends on their availability in the images and the reliability of their measurement. When points are not available in large quantities, all the previous methods will have problems. Image line is also a common feature used in computer vision and was successfully used for 3-D motion estimation from 1980s to the latest. So, it is significant to consider the problem of 3-D motion segmentation from line correspondences. In [13], Shi et.al proposed a method to do motion segmentation using line correspondences. But they only dealt with translating objects. In [14],they proposed a general method for 3-D motion using line correspondences, but this method requires too many features when the number of motion is large.

1.1 Contributions of This Paper

In this paper, we proposed a novel method to segment multiple rigid-body 3D motions from the optical flow of line correspondences in two views. The remainder of this paper decomposes as follows. In Section 2, we introduce the concept of line optical flow. In Section 3, we show now to segment the 3D motion using the line optical flow and the method of GPCA [5]. Finally, Section 4 validates our algorithm by experiments with simulated data and real scenes.

2 Line Optical Flow [15]

A line l in the image plane is represented by a vector $m = (m_x, m_y, m_z)^T$ giving its equation:

$$m_x x + m_y y + m_z f = 0 \tag{1}$$

where f is the focal length. The interpretation of m is that it is the normal to the plane defined by the 2D line and the optical center of the camera(see Figure 1); note that this plane contain also the corresponding 3D line, L. We use two relationships between the representation of a 3D line and its image in the image plane. The first relation is that the direction v of the 3D line is perpendicular to the normal m to the plane it defines with the optical center.

$$m^T v = 0 \tag{2}$$

Then we take its time derivative and obtain

$$g^T v + m^T \partial v / \partial t = 0 \tag{3}$$

where $g = (\partial m_x/\partial t, \partial m_y/\partial t, \partial m_z/\partial t)^T$. We define g as the *line optical flow*. We assume that the 3D line under consideration is attached to a rigid body whose motion is described by its instantaneous angular velocity $W = (\omega_1, \omega_2, \omega_3)^T$ and linear velocity $T = (t_1, t_2, t_3)^T$, its kinematic screw is at the origin o.Then

$$\partial v / \partial t = W \times v \tag{4}$$

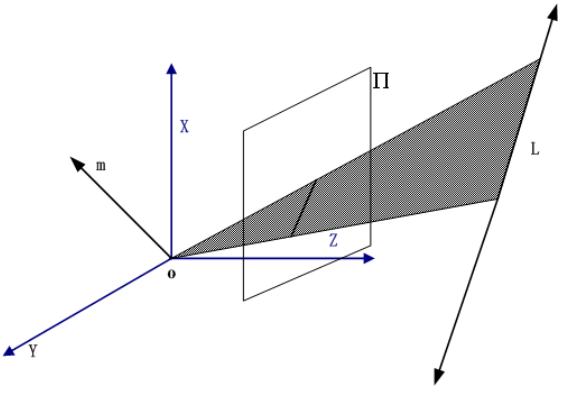


Fig. 1. The vector m

$$A = \begin{vmatrix} m_x & m_y & m_z \\ \omega_3 m_y - \omega_2 m_z + \Delta m_x & \omega_1 m_z - \omega_3 m_x + \Delta m_y & \omega_2 m_x - \omega_1 m_y + \Delta m_z \end{vmatrix} \tag{7}$$

replacing $\partial v/\partial t$ by this value, (3)give us the second relationship between the representation of a 3D line and its image

$$g^T v + m^T (W \times v) = 0 \tag{5}$$

From equation (5) and (2), one can get the linear constraints about v

$$\begin{cases} g^T v + m^T (W \times v) = 0 \\ m^T v = 0 \end{cases} \tag{6}$$

Let A is the coefficient matrix of the two equations(A is defined in equation(7)), then $R(A) = 1$ [16]. So,

$$\frac{m_x}{\omega_3 m_y - \omega_2 m_z + \Delta m_x} = \frac{m_y}{\omega_1 m_z - \omega_3 m_x + \Delta m_y} = \frac{m_z}{\omega_2 m_x - \omega_1 m_y + \Delta m_z} \tag{8}$$

Let $k = -\frac{m_x}{m_y}$ is the slope of image line and $b = -\frac{m_z f}{m_y}$is the intercept. Then one can get

$$\begin{cases} kb\omega_3 - b^2\omega_1/f - k\omega_2 \cdot f - \omega_1 \cdot f = -\Delta b \\ (1+k^2)\omega_3 + b\omega_2/f - kb\omega_1/f = -\Delta k \end{cases} \tag{9}$$

Let $\omega_1/f = \omega_{1f}, \omega_2/f = \omega_{2f}, \omega_1 \cdot f = \omega_{1F}, \omega_2 \cdot f = \omega_{2F}$,then

$$\begin{cases} kb\omega_3 - b^2\omega_{1f} - k\omega_{2F} - \omega_{1F} = -\Delta b \\ (1+k^2)\omega_3 + b\omega_{2f} - kb\omega_{1f} = -\Delta k \end{cases} \tag{10}$$

Let

$$y(k,b) = \begin{bmatrix} kb & -b^2 & 0 & -1 & -k & \Delta b \\ (1+k^2) & -kb & b & 0 & 0 & \Delta k \end{bmatrix}^T \tag{11}$$

and

$$u = (\omega_3, \omega_{1f}, \omega_{2f}, \omega_{1F}, \omega_{2F}, 1)^T \tag{12}$$

Equation (10) can be expressed as

$$y(k,b)^T u = \mathbf{0} \tag{13}$$

As each line-line correspondence gives 2 independent equations, one may estimate u linearly from 3 correspondences for one motion.

3 Direct Motion Segmentation Using MLOFC

3.1 Multibody Line Optical Flow Constraint(MLOFC)

Now consider a scene with n rigid-body motions (n is known) with associated motion component $u_i \in \mathbb{R}^6$, where u_i is the motion component associated with the motion of the i^{th} object.

Therefore, we define the *multibody line optical flow constraint*(MLOFC)

$$MLOFC = \prod_{i=1}^{n} (y(k,b)^T u_i) = \mathbf{0} \tag{14}$$

MLOFC is a homogeneous polynomial of degree n in y, which can be written as a linear combination of the monomials $y_1^{n_1} y_2^{n_2} \cdots y_6^{n_6}$ with $n_1 + n_2 + \cdots + n_6 = n$. If we stack these $M_n^6 = \binom{n+6-1}{6-1}$ independent monomials into a vector $\nu_n(y) \in \mathbb{R}^{M_n^6}$,we get

$$MLOFC = \nu_n(y)^T \mathcal{U} = \sum \mathcal{U}_{n_1,n_2,\cdots,n_6} y_1^{n_1} y_2^{n_2} \cdots y_6^{n_6} \tag{15}$$

The vector $\mathcal{U} \in \mathbb{R}^{M_n^6}$is called *multibody line optical flow*, and $\nu_n : \mathbb{R}^6 \rightarrowtail \mathbb{R}^{M_n^6}$ is called the *Veronese* map of degree n.

In the following subsections, we will demonstrate that in the case of 3-D rigid motion models, the MLOFC can be expressed linearly in terms of a multibody motion model $\mathcal{W}$. By exploiting the algebraic properties of $\mathcal{W}$, we will derive an algebraic closed form solution to the following problem [8]:

Problem 1 (Direct multiple-motion segmentation from line optical flow). *Given the slopes and intercepts of line correspondences of a motion sequence generated from n 3-D rigid motion models, estimate the model parameters* $\{\mathcal{W}_i\}_{i=1}^n$*, without knowing which image measurements correspond to which model.*

3.2 Computing the Multibody Motion Model

Let $l(k,b) \leftrightarrow l'(k',b')$ be an arbitrary line-line correspondence associated with any of the n motions in two consecutive image frames. We may estimate $\mathcal{U}$ linearly from $M_n^6/2$ correspondences, using equations (15).

But note that the entries (1,3),(2,4) and (2,5)of each y_i are zero, $y_1^{n_1}y_2^{n_2}\cdots y_6^{n_6}=0$ in both equations of (15) when $(n_3\neq 0 \quad and \quad n_4\neq 0)$ or$(n_3\neq 0 \quad and \quad n_5\neq 0)$. After enforcing these equations we obtain

$$\sum \mathcal{Y}_n^T\widetilde{\mathcal{U}}=\mathbf{0} \tag{16}$$

where $\mathcal{Y}\in\mathbb{R}^{M_n^6-Z_n}, \widetilde{\mathcal{U}}\in\mathbb{R}^{M_n^6-Z_n}$ is equal to $\nu_n(y), \mathcal{U}$ without the both zero entries in two equations of (15).When $n=2$, then $Z_n=2$; when $n>2$, then $Z_n=3M_{n-2}^3+\sum_{i=1}^{n-1}M_{n-2-i}^3\cdot M_i^3+\sum_{i=1}^{n-1}M_{n-2-i}^3\cdot M_i^2+M_{n-1}^2$.

As one line-line correspondence gives 2 independent equations(16), we may estimate $\widetilde{\mathcal{U}}$ linearly from $(M_n^6-Z_n)/2$ correspondence.

3.3 Estimate f

After the estimation of $\widetilde{\mathcal{U}}$, we compute focal length f

$$f^{2n}=\frac{\prod_{i=1}^n w_{1F}^i}{\prod_{i=1}^n w_{1f}^i}=\frac{\prod_{i=1}^n w_{2F}^i}{\prod_{i=1}^n w_{2f}^i} \tag{17}$$

$$f^2=\frac{\sum_{i=1}^n w_{1F}^i}{\sum_{i=1}^n w_{1f}^i}=\frac{\sum_{i=1}^n w_{2F}^i}{\sum_{i=1}^n w_{2f}^i} \tag{18}$$

where w^i is the parameter of the ith motion models. And

$$\prod_{i=1}^n w_{1F}^i=\mathcal{U}_{M_{n-1}^6+M_{n-1}^5+M_{n-1}^4+1},\quad \prod_{i=1}^n w_{1f}^i=\mathcal{U}_{M_{n-1}^6+1} \tag{19}$$

$$\prod_{i=1}^n w_{1F}^i=\mathcal{U}_{M_{n-1}^6+M_{n-1}^5+M_{n-1}^4+M_{n-1}^3+1},\quad \prod_{i=1}^n w_{2f}^i=\mathcal{U}_{M_{n-1}^6+M_{n-1}^5+1} \tag{20}$$

$$\sum_{i=1}^n w_{1F}^i=\mathcal{U}_{M_{n-1}^6+M_{n-1}^5+M_{n-1}^4+M_{n-1}^3},\quad \sum_{i=1}^n w_{1f}^i=\mathcal{U}_{M_{n-1}^6+M_{n-1}^5} \tag{21}$$

$$\sum_{i=1}^n w_{2F}^i=\mathcal{U}_{M_{n-1}^6+M_{n-1}^5+M_{n-1}^4+M_{n-1}^3+M_{n-1}^2},\quad \sum_{i=1}^n w_{2f}^i=\mathcal{U}_{M_{n-1}^6+M_{n-1}^5+M_{n-1}^4} \tag{22}$$

In our algorithm, we estimate f as the average value of (17)(18).

3.4 Refined MLOFC

After the estimation of f, we let

$$y'(k,b)=\begin{bmatrix} -(\frac{b^2}{f}+f) & -kf & kb & \triangle b \\ -\frac{kb}{f} & \frac{b}{f} & (1+k^2) & \triangle k \end{bmatrix}^T \tag{23}$$

$$u'=[w_1,w_2,w_3,1]^T \tag{24}$$

MLOFC(15) can be refined as

$$MLOFC' = \nu_n(y')^T \mathcal{U}' = \sum \mathcal{U}'_{n_1,n_2,n_3,n_4} {y'_1}^{n_1} {y'_2}^{n_2} {y'_3}^{n_3} {y'_4}^{n_4} \tag{25}$$

Then, we can solve for the multibody motion model $\mathcal{W} = \mathcal{U}'$ uniquely from (25).

3.5 Segmenting the Multibody Model

Given $\mathcal{U}'$, we can use the algorithm in [6] to compute the parameters u'_i of each individual motion model from the derivatives of $MLOFC'$, i.e.,

$$D_{MLOFC'} = \frac{\partial MLOFC'}{\partial y'} = \sum_{i=1}^{n} \prod_{l \neq i} ({y'}^T u'_l)) u'_i \tag{26}$$

If we evaluate $D_{MLOFC'}$ at a line $y = z_i$ that correspondence to the *i*th motion, i.e.if z_i is such that $z_i^T u'_i = 0$, then we have $D_{MLOFC'} \sim u'_i$. Therefore, given $\mathcal{U}'$ we can obtain the motion parameters as

$$u'_i = \frac{D_{MLOFC'}}{e_K^T D_{MLOFC'}}|_{y=z_i} \tag{27}$$

where $e_K = [0, \ldots, 0, 1]^T \in \mathbb{C}^K$ and $z_i \in \mathbb{C}^K$ is a nonzero vector that $z_i^T u'_i = 0$.

Once the right individual motion parameters have been computed, one may compute the Sampson error by assigning each feature to $\{u'_i\}_{i=1}^n$, and cluster the correspondence to the one that minimize the error.

We summarize the whole procedure in Algorithm 1.

Algorithm 1. (3D Motion segmentation using Straight-line optical flow)
Given $N > (M_n^6 - Z_n)/2$ line correspondence $\mathcal{L} = \{l_i(k,b) \leftrightarrow l'_i(k',b')\}_{i=1}^N$ in two consecutive image frames:

1. Construct Multibody Line Optical Flow Constraint $MLOFC = \prod_{i=1}^n (y(k,b)^T u_i) = \mathbf{0}$;
2. Compute $\widetilde{\mathcal{U}}$ using equation (16);
3. Estimate f using equations (17)(18);
4. Construct refined Multibody Line Optical Flow Constraint $MLOFC' = \nu_n(y')^T \mathcal{U}'$;
5. for $i = n : 1$

$$z_i = \arg\min_{z \in \mathcal{L}} \frac{\frac{|MLOFC'|}{\|\Pi D_{MLOFC'}\|} + \delta}{\frac{|z^T u'_{i+1}| \cdots |z^T u'_n|}{\|\Pi u'_{i+1}\| \cdots \|\Pi u'_n\|} + \delta};$$

where $\delta > 0$ is a small positive number,$\Pi = [I_{K-1} \quad 0] \in \mathbb{R}^{(K-1)\times K}$

$$u'_i = \frac{D_{MLOFC'}}{e_K^T D_{MLOFC'}},$$

end

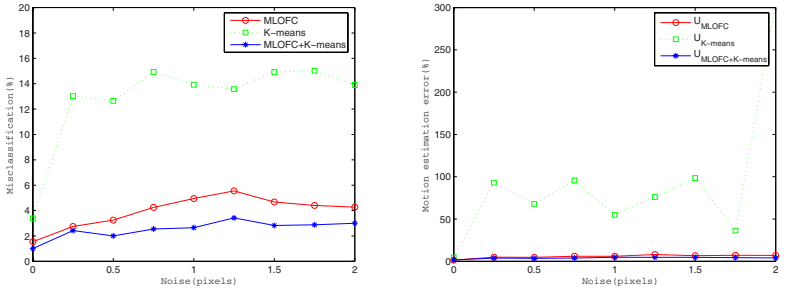


Fig. 2. Correctness of motion segmentation and motion estimation ($\{u_i'\}_{i=1}^2$) as a function of noise. The number of line correspondences of each object is set to be 20.

4 Experiments on Synthetic and Real Images

In this section, we evaluate our motion segmentation algorithms on both real and synthetic data. We compare our results with K-means methods and use our algorithms to initialize iterative technique. We consider the following algorithms:

1. MLOFC: Using Algorithms 1 to cluster the line correspondences.
2. K-means [17]: this algorithm alternates between computing(linearly) the parameters of different motion classes and clustering the line correspondences using the Sampons-distance error to the different motions.

4.1 Synthetic Experiments

First, we conduct some simulated experiments to check the correctness of the proposed algorithm with respect to the amount of noise in the image measurements for $n = 2$ and $n = 3$. We randomly generated n groups of $N = n \times m$ 3-D lines, each group has m lines. Then the n groups of 3-D lines are projected in two views. In the presence of noise, we added zero-mean Gaussian noise with standard deviation between 0 and 2 pixels in an image size of 1000×1000.

To show the general sensitivity to noise, the average percentage of misclassified correspondences, i.e., the percent of wrong classified line number in whole set and the average relative errors in the estimation of $\{u_i'\}_{i=1}^n$ are recorded. We first run K-means algorithm starting from random classification. The K-means algorithm converges to a local minimum due to bad initialization. Running the K-means algorithm starting from the clustering produced by the MLOFC algorithm result in a better result. The results when $n = 2$ are showed in Fig.2 and the results when $n = 3$ are showed in Fig.5.

During the procedure of segmentation using MLOFC, we didn't consider the influence of linear velocity $T = (t1, t2, t3)^T$. We treat it as noise. In the second simulated experiment, we evaluate the performance of the algorithms with respect to linear velocity. The linear velocity of the two motion objects are t times as large as the angular

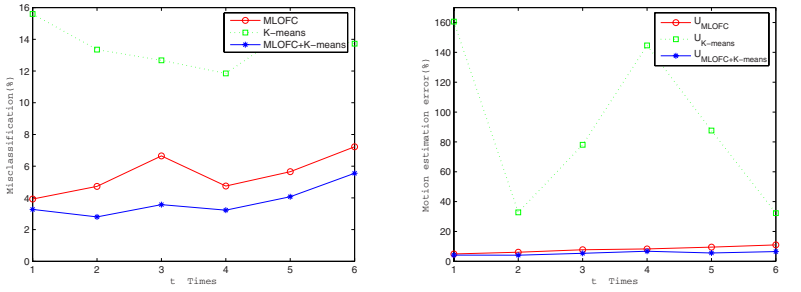


Fig. 3. Correctness of motion segmentation and motion estimation ($\{u_i'\}_{i=1}^2$) as a function of linear velocity. The image noise is set to be 1.0 pixel. The number of line correspondences of each object is set to be 20.

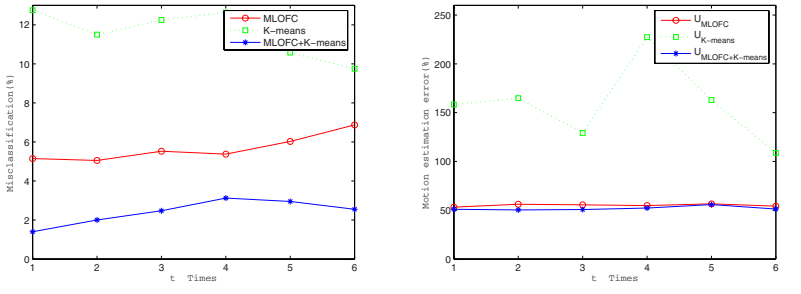


Fig. 4. Correctness of motion segmentation and motion estimation ($\{u_i'\}_{i=1}^2$) as a function of linear velocity, when the second object take pure translating motion. The image noise is set to be 1.0 pixel. The number of line correspondences of each object is set to be 20.

velocity. The result is showed in Fig.3. One can see that with the increasing of linear velocity, the correctness of motion segmentation is decrease. But the result is still good.

Then, we conduct another experiment to test the performance to pure translating motion. In this experiment, we set the angular velocity of second object as zero, i.e., this object take a pure translating motion. The linear velocity of the two objects are same. It is t times as large as the first object's angular velocity. The result is showed in Fig.4. One can see that the performance is good, although we did not compute the linear velocity.

4.2 Real Experiment

Figure 6(a) shows one frame of a 640×560 sequence taken by a static camera observing two moving boxes using different 3D rigid motions. Figure 6(b) shows the lines that detected by *Edge linking and line segment fitting* algorithm that developed by *Peter*

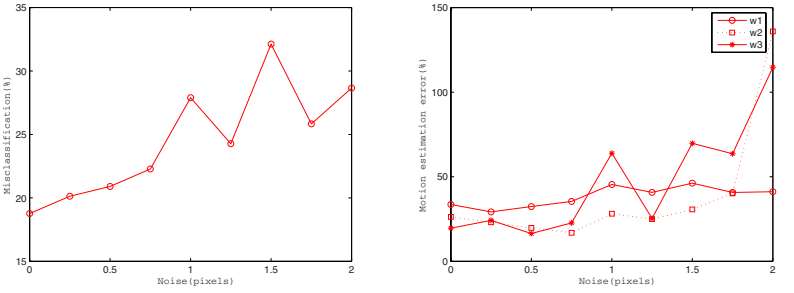
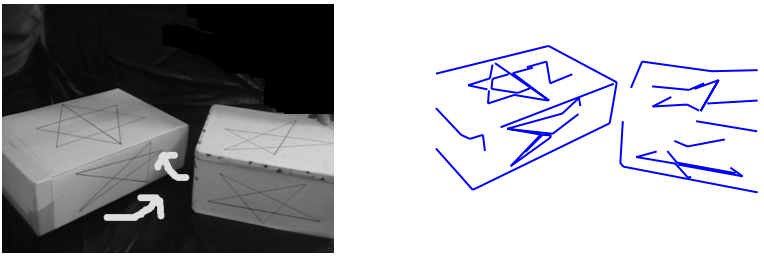


Fig. 5. Correctness of motion segmentation and motion parameters ($\{w_i\}_{i=1}^{3}$) as a function of noise. The number of line correspondences of each object is set to be 30.



(a) (b)

Fig. 6. Line segmentations

Kovesi. We manually select $m = 10$ line correspondences in two consecutive frames. And the average misclassification error using MLOFC is 17.197% for 10 pairs of consecutive frames. This result is not very satisfied and may due to the camera distortion and the insufficient of line correspondences. But may be improved by other optimization method, such as EM [17].

5 Conclusion

We present a closed form solution to the problem of segmenting multiple 3D motion models from straight-line optical flow. The algebraic method of motion classification involves computation of MLOFC and individual angular velocity. Our approach has the advantage that it provides a global, non-iterative solution, which was able to provide an initial classification for other optimization segmentation method. Experiments with simulated data and real scenes validate our algorithm. But we have to admit that this method can not deal with motions that all take the same angular velocity. So our future work is to develop a general algorithm to all kinds of 3D rigid motions.

References

1. Rittscher, J., Tu, P.H., Krahnstoever, N.: Simultaneous estimation of segmentation and shape. In: IEEE Conference on Computr Vision and Pattern Recognition (2005)
2. Torr, P.H.S.: Geometric motion segmentation and model selection Phil. Trans. Royal Society of London A (1998)
3. Feng, X., Perona, P.: Scene segmentation from 3d motion (1998)
4. Torr, P., Szeliski, R., Anandan, P.: An integrated bayesian approach to layer extraction from image sequences. IEEE Trans. on Pattern Analysis and Machine Intelligence 23, 297–303 (2001)
5. Vidal, R., Ma, Y., Sastry, S.: Generalized principal component analysis (gpca). In: IEEE Conference on Computr Vision and Pattern Recognition, pp. 621–628 (2003)
6. Vidal, R., Ma, Y.: A unified algebraic approach to 2-d and 3-d motion segmentation. In: European Conference on Computer Vision, pp. 1–15 (2004)
7. Singaraju, D., Vidal, R.: A bottom up algebraic approach to motion segmentation. In: Asian Conference on Computer Vision, pp. 286–296 (2006)
8. Vidal, R., Singaraju, D.: A closed-form solution to direct motion segmentation. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 510–515 (2005)
9. Vidal, R., Ravichandran, A.: Optical flow estimation and segmentation of multiple moving dynamic textures. In: IEEE Conference on Computer Vision and Pattern Recognition (2005)
10. Fan, X., Vidal, R.: The space of multibody fundamental matrices: Rank, geometry and projection. In: IEEE International Conference on Computer Vision (2005)
11. Yan, J., Pollefeys, M.: Articulated motion segmentation using ransac with priors. In: IEEE International Conference on Computer Vision (2005)
12. Hajder, L., Chetverikov, D.: Robust 3d segmentation of multiple moving objects underweak perspective. In: IEEE International Conference on Computer Vision (2005)
13. Shi, F., Wang, J., Zhang, J., Liu, Y.: Motion segmentation of multiple translating objects using line correspondences. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 315–320 (2005)
14. Zhang, J., Shi, F., Liu, Y.: Motion segmentation by multibody trifocal tensor using line correspondence. In: International Conference on Pattern Recognition (2006)
15. Faugeras, O., Navab, N., Deriche, R.: On the information contained in the motion field of lines and the cooperation between motion and stereo. International Journal of Imaging Systems and Technology (1991)
16. Jiang, S.F., Chen, Z., Yang, S.H.: Reconstructing 3d rotation motion parameters from 2d image straight-line optical flow. Journal of Nanchang Institute of Aeronautical Technology(Natural Science) (2005)
17. Hartley, R., Vidal, R.: The multibody trifocal tensor: Motion segmentation from 3 perspective views. In: IEEE Conference on Computer Vision and Pattern Recognition (2004)