Simultaneous Recovery of Shape, Motion and Grouping by Applying Rank Constraints

Alireza Fathi

Georgia Institute of Technology

1 Introduction

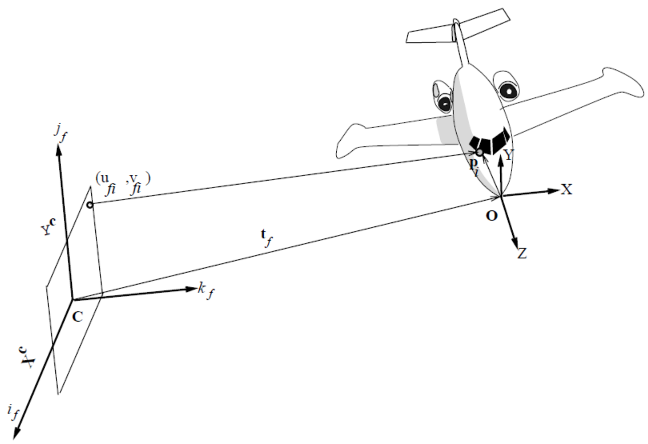

Fig. 1. The figure shows the camera and an object with its local coordinate system. Vectors i_f and j_f define the image plane and k_f is its normal axis. The image is taken from Costeira and Kanade [5].

Suppose there are two objects which are moving independently and rigidly in $3D$ space. Assume we have an orthographic camera[1]. The image plane is perpendicular to Z axis and passes from the origin. What we are given as input is the projection of $3D$ object points to the image plane at each time step, as a sequence of images. So basically we are given a video of two objects moving, and our goal is to recover the following information: group points as seperate objects, estimate the object shapes, and recover their motion in $\mathbb{R}^3$ space. We describe each in more detail below:

Grouping Points: Every pixel location in each image corresponds to projection of a $3D$ point of one of the two objects to the image plane. We want to group the set of points in all images that correspond to each object.

Object Shapes: Since objects are rigid, we define a local coordinate frame in $\mathbb{R}^3$ for each object. Our goal is to project back the pixel locations corresponding to each object, to their corresponding $3D$ location in its local coordinate frame.

Motion: We like to recover the rotation and translations applied to each object at each frame.

2 Previous Work

Here we survey some of the previous works. There is a tremendous number of papers that we will not be able to cover all in here.

[1] Orthographic camera model assumes there is an image plane in $3D$ space on which $3D$ points in the scene are projected vertically onto it, without any perspective. Perspective camera model is a more realistic model. However, given the projections of points in an orthographic model, and intrinsic parameters of the camera (calibration), it is possible to recover the actual pixel locations in image coordinates. In this report, we have assumed that the camera is orthographic for simplicity, however we will mention works that have extended the model for perspective camera in Section 2.

Shape and Grouping: There is a huge literature on recovering shape from point tracks in images. The one which is the most related to this paper is Tomasi and Kanade [16] which provide a method for recovering the $3D$ shape of a single moving object, given its point tracks in an image sequence. Tomas and Kanade use a factorization method by applying rank constraints on the point track matrix to recover the shape and motion. Their assumption is that there is only one rigid object moving with respect to the camera. Costeira and Kanade extend their work by handling the case of multiple independently moving objects [5]. Kanatani [10] proves their method works in theory. The method presented in Costeira and Kanade is sensitive to noise. There are works that extend their method to make it more robust to noise [20].

In these methods the assumption is that the camera is orthographic. There are works which extend the camera model to paraperspective and perspective models [18,19,14]. The other assumption in these works is that all the trajectories are given. However, this is not a right assumption ans is often violated in case of occlusions. Brand [1] introduces a method for incremental SVD which can handle missing data.

Deformable Objects: In the mentioned works the assumption is that the individual objects are rigid. However, Bregler and others extend those methods to deformable objects [2,17]. They represent the shape of each object as a linear combination of multiple shapes.

Degeneracies: The motion and shape matrices for a single independently moving object might have rank less than 4, in which case $rank(W){=}min(rank(M), rank(S))$. For example if object points lie on a plane, the rank of S will be 3. If they like on a line, the rank of S will be 2. However, here we assume both M and S have rank 4. Gear and others have studies the degenrate cases [7,21].

Local Motion Estimation: The approaches to shape and motion recovery require a set of point tracks in sequence of images as their input. Here we mention a basic method developed by Lucas and Kanade [13] for computing these tracks for a sequence of images. They assume that the displacement of the image content between two adjacent frames is small and constant within a neighborhood of an image point $p = (u, v)$. Namely they use Taylor expansion to show that local image flow $(V = (V_x, V_y))$ must satisfy

$$\begin{aligned}
I_x(q_1)V_x(p) + I_y(q_1)V_y(p) &= -I_t(q_1) \\
I_x(q_2)V_x(p) + I_y(q_2)V_y(p) &= -I_t(q_2) \\
&. \\
&. \\
I_x(q_w)V_x(p) + I_y(q_w)V_y(p) &= -I_t(q_w)
\end{aligned}$$

where $q_1, ..., q_w$ are points in a small window around (u, v). This set of equations is usually over-determined and $V_x(u)$ and $V_y(u)$ can be computed using leas square principle. However, if matrix A consisting of $I_x(q_i)$ and $I_y(q_i)$ is singular (has rank 1 or 0), there will be multiple answers for motion vector V. This happens when the small window around point (u, v) is a uniform patch (rank 0) or when it is on an edge (rank 1).

As a result, there is a whole set of papers on what windows are good for tracking and which ones are not [15,8]. There are more recent flow computation and tracking algorithms which we do not explore here [3,4,11,12]. Irani uses rank constraints to enhance the tracking results over multiple frames [9]. More complicated methods are developed to extend Irani's work to non-rigid objects [6].

Our contribution in this paper, is the proofs for theorems 2 and 3. Theorem 3 is specially important and is not addressed in previous work, while it is critical to be able to build a spanning tree of points belonging to each object.

3 Problem

Assume there is a $2D$ plane C in $\mathbb{R}^3$ whose normal is parallel to Z axis. Suppose we have two set of points $\mathcal{P}_1 = \{P_1, ..., P_{n_1}\}$ and $\mathcal{P}_2 = \{Q_1, ..., Q_{n_2}\}$ in $\mathbb{R}^3$. Assume both $\mathcal{P}_1$ and $\mathcal{P}_2$ have at least 5 points. Assume there is not a plane spanning any 4 points, i.e. any 4 points span the $3D$ space. For each set we define a local coordinate system with its origin at the mean of the points in that set. Each set of points is rigid meaning that the local coordinate of its points $\hat{P}_i$ are always constant. However, each set of points can move in $\mathbb{R}^3$ space. We transform $\mathcal{P}_1$ and $\mathcal{P}_2$ independently F times in $\mathbb{R}^3$ resulting in F instances of the points with different locations in space. Each transformation $T_{f,i}$ of point set $\mathcal{P}_i$, consists of a ratation $R_{f,i}$ and a translation $t_{f,i}$. We assume the set of

F 4×4 transformations of each point set have rank 4 and are linearly independent. Assume we are given the projection of all $N = n_1 + n_2$ points to C (x, y coordinates) at every instance $f \in 1, .., F$ as input. The projections are sorted, i.e. the i'th projection coordinate always correspond to a particular point. We show that it is possible to recover the corresponding location of each point in $\mathbb{R}^3$, it is possible to recover the transformation matrices $T_1, ..., T_F$ in the x and y axis (the first two rows) and it is possible to group the recovered points into two sets $\mathcal{P}_1$ and $\mathcal{P}_2$.

4 Approach

Consider an unknown point $P_{f,i}$ at instance f, with a local coordinate $\hat{P}_i$. Its projection on the image plane C at instant (frame) f is given by the following transformation,

$$P^C_{f,i} = R_f \hat{P}_i + t_f$$

where R_f and t_f are the rotation and translation matrices that transform the coordinate system of the object into camera's coordinate system at instant f. R_f and t_f are also unknowns. After simplifying the formulation using homogeneous coordinates and using the orthographic camera model we weill have

$$\begin{aligned} P^C_{f,i} = \begin{bmatrix} u_{f,i} \\ v_{f,i} \end{bmatrix} &= \begin{bmatrix} R_f(2 \times 3) \; t_f(2 \times 1) \end{bmatrix} \begin{bmatrix} \hat{P}_i \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} I_{x,f} & t_{x,f} \\ J_{x,f} & t_{y,f} \end{bmatrix} \begin{bmatrix} \hat{P}_i \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} i_{x,f} & i_{y,f} & i_{z,f} & t_{x,f} \\ j_{x,f} & j_{y,f} & j_{z,f} & t_{y,f} \end{bmatrix} \begin{bmatrix} \hat{p_{x,i}} \\ \hat{p_{y,i}} \\ \hat{p_{z,i}} \\ 1 \end{bmatrix} \end{aligned}$$

We were able to ommit the third dimension since all the points are being projected parallel to the Z axis and the plane C passes from origin.

We are given $N = n_1 + n_2$ projection tracks, each containing F instances as input. A projection track is a sequence of F corresponding projection points $P^C_{1,i}$, $P^C_{2,i}$,..., $P^C_{F,i}$ at instances 1 to F.

We collect the N tracks into a $2F \times N$ matrix W. We do not know the track groupings. However for know, we assume we know the groupings, and n_1 tracks are grouped together to build a $2f \times n_1$ matrix W_1 and n_2 tracks are grouped together to build a $2f \times n_2$ matrix W_2.

For W_1 we can write

$$W_1 = \begin{bmatrix} u_{1,1} & u_{1,2} & \dots & u_{1,n_1} \\ & \vdots & \vdots & \\ u_{F,1} & u_{F,2} & \dots & u_{F,n_1} \\ v_{1,1} & v_{1,2} & \dots & v_{1,n_1} \\ & \vdots & \vdots & \\ v_{F,1} & v_{F,2} & \dots & v_{F,n_1} \end{bmatrix} = \begin{bmatrix} i_{x,1} & i_{y,1} & i_{z,1} & t_{x,1} \\ & & \vdots & \\ i_{x,F} & i_{y,F} & i_{z,F} & t_{x,F} \\ j_{x,1} & j_{y,1} & j_{z,1} & t_{y,1} \\ & & \vdots & \\ j_{x,F} & j_{y,F} & j_{z,F} & t_{z,F} \end{bmatrix} \begin{bmatrix} \hat{p}_{x,1} & & \hat{p}_{x,n_1} \\ \hat{p}_{y,1} & \dots & \hat{p}_{y,n_1} \\ \hat{p}_{z,1} & & \hat{p}_{z,n_1} \\ 1 & & 1 \end{bmatrix} \tag{1}$$

and we can briefly write $W_1 = M_1 S_1$. M_1 is a $2F \times 4$ matrix and S_1 is a $4 \times n_1$ matrix. The observation is that the matrices M_1 and S_1 have rank 4 and as a result, W_1 should have rank 4 as well. Similarly $W_2 = M_2 S_2$ will have rank 4 as well.

4.1 Recovering M and S for one set of points

If we are given W_1 we can recover M_1 and S_1 as follows (These results can be applied to W_1, M_1 and S_1 as well). We first decompose W_1 using SVD:

$$W_1 = U_1 \Sigma_1 V_1^T$$

where $U_1 \in \mathbb{R}^{2F\times 4}$, $\Sigma_1 \in \mathbb{R}^{4\times 4}$ and $V_1 \in \mathbb{R}^{n_1\times 4}$, such that $U^T U = \mathcal{I}$ and $V^T V = \mathcal{I}$. We can write

$$\hat{M}_1 = U \Sigma^{\frac{1}{2}}$$

$$\hat{S}_1 = \Sigma^{\frac{1}{2}} V^T$$

We can then write $W_1 = \hat{M}_1 \hat{S}_1$, however this factorization is not unique. Since Given any invertible 4×4 matrix A_1, $W_1 = (\hat{M}_1 A_1)(A_1^{-1}\hat{S}_1)$ is also a solution. However, there are rotation and translation constraints that will make matrix A over determined. The first 3 columns of M_1 are related to rotational motion and its last column is related to translational motion. As M_1 is defined in Equation 1, we have

$$\begin{bmatrix} i_{x,f} \; i_{y,f} \; i_{z,f} \end{bmatrix} \begin{bmatrix} i_{x,f} \; i_{y,f} \; i_{z,f} \end{bmatrix}^T = 1$$

$$\begin{bmatrix} j_{x,f} \; j_{y,f} \; j_{z,f} \end{bmatrix} \begin{bmatrix} j_{x,f} \; j_{y,f} \; j_{z,f} \end{bmatrix}^T = 1$$

$$\begin{bmatrix} i_{x,f} \; i_{y,f} \; i_{z,f} \end{bmatrix} \begin{bmatrix} j_{x,f} \; j_{y,f} \; j_{z,f} \end{bmatrix}^T = 0$$

which means if we write $M_1 = \hat{M}_1 A = \hat{M}_1 \left[A_R | A_t\right]$, we will have

$$\hat{m}_f A_R A_R^T \hat{m}_f^T = 1, s.t. \forall f \in \{1, .., 2F\}$$

$$\hat{m}_f A_R A_R^T \hat{m}_{f+F}^T = 0, s.t. \forall f \in \{1, .., F\}$$

Since the center of points' local coordinate frame is at their mean, we have

$$\begin{aligned} \bar{W} &= \begin{bmatrix} \frac{1}{n_1}\Sigma u_{1,i} \\ . \\ . \\ . \\ \frac{1}{n_1}\Sigma v_{F,i} \end{bmatrix} \\ &= M_1.\bar{S}_1 = \begin{bmatrix} \hat{M}_1 A_R | \hat{M}_1 A_t \end{bmatrix} \begin{bmatrix} \bar{P} \\ 1 \end{bmatrix} = \hat{M}_1 A_t \end{aligned}$$

since $\bar{P} = 0$. We can solve for A_R and A_t using least square techniques.

Here we showed that it is possible to recover matrices M_1 and S_1 if we knew which projection tracks correspond to their point set. However, we do not know these groupings. In next part we show how we can group points into two sets.

4.2 Grouping Points

If we knew the groupings, we could permute matrix W to get W^* as below

$$W^* = [W_1|W_2] = [M_1|M_2] \begin{bmatrix} S_1 & 0 \\ 0 & S_2 \end{bmatrix}$$
$$= [U_1|U_2] \begin{bmatrix} \Sigma_1^{\frac{1}{2}} & 0 \\ 0 & \Sigma_2^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \begin{bmatrix} A_1^{-1} & 0 \\ 0 & A_2^{-1} \end{bmatrix} \begin{bmatrix} \Sigma_1^{\frac{1}{2}} & 0 \\ 0 & \Sigma_2^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} V_1^T & 0 \\ 0 & V_2^T \end{bmatrix}$$

We decompose matrix W using SVD as $W = U\Sigma V^T$, where U, Σ and V are $2F \times 8$, 8×8 and $N \times 8$ matrices respectively. We claim that the $N \times N$ matrix $Q = VV^T$ can be used to group the points into the two point sets. We claim that if $Q(\alpha, \beta)$ is not equal to 0, it means that points p_α and p_β belong to the same point set. Also we claim that for every α, there is at least one β for which $Q(\alpha, \beta) \neq 0$. As a result, we can recover the point sets given Q by building two spanning trees of the points given their connections in Q.

We prove this in the following theorems.

4.3 Theorem 1

Let $\mathcal{T}$ be N points $(p_1, ..., p_N)$ that belong to an r dimensional subspace $\mathcal{L} \subset \mathbb{R}^n$. Define an $n \times N$ matrix $W = [p_1 ... p_N]$ that contains these points. Let $V_r = \{v_1, ..., v_r\}$ be W's r right singular vectors with non-zero singular values. Define the $N \times N$ matrix Q by

$$Q = V_r V_r^T = \sum_1^r v_i v_i^T$$

Divide the N points into two disjoint subsets $\mathcal{T}_i, i \in \{1, 2\}$, and let r_i be the dimension of subspace $\mathcal{L}_i$ defined by the ith subset $\mathcal{T}_i$. If the two subspaces are linearly independent and the αth and the βth points belong to different subspaces, the element $Q_{\alpha,\beta}$ is zero.

Proof: Assume the number of points in the first subset is N_1 and the number of points in the second subset is N_2. Suppose $p_1, ..., p_{N_1} \in \mathcal{L}_1$ and $p_{N_1+1}, ..., p_{N_2} \in \mathcal{L}_2$.

Since the subspace $\mathcal{L}_1$ has dimension r_1, the $n \times N_1$ matrix $W_1 = [p_1 ... p_{N_1}]$ has rank r_1. The null space $(\mathcal{N}_1)$ of W_1 will have dimension $v_1 = N_1 - r_1$. Similarly, the $n \times N_2$ matrix $W_2 = [p_{N_1+1} ... p_{N_2}]$ has rank r_2, and its null space $(\mathcal{N}_2)$ has dimension $v_2 = N_2 - r_2$. Let $\{n_1, ..., n_{v_1}\}$ be an arbitarary orthonormal basis of $\mathcal{N}_1$ and $\{n'_1, ..., n'_{v_2}\}$ be an arbitarary orthonormal basis of $\mathcal{N}_2$.

Let $\tilde{n}_i, i \in 1, .., v_1$ and $\tilde{n}'_i, i \in 1, ..., v_2$ be N dimensional vectors defined by padding n_i and n'_i with zero elements as follows:

$$\tilde{n}_i = \begin{pmatrix} n_i \\ 0 \end{pmatrix}$$

$$\tilde{n}'_i = \begin{pmatrix} 0 \\ n'_i \end{pmatrix}$$

The $v = v_1 + v_2 = N_1 + N_2 - r_1 - r_2 = N - r$ vectors $\{\tilde{n}_1, ..., \tilde{n}_{N_1 - r_1}, \tilde{n}'_1, ..., \tilde{n}'_{N_2 - r_2}\}$ produce an orthonormal system of $\mathbb{R}^N$ belonging to the null space $\mathcal{N}$ of the observation matrix W. As a result $\{\tilde{n}_1, ..., \tilde{n}_{N_1 - r_1}, \tilde{n}'_1, ..., \tilde{n}'_{N_2 - r_2}\}$ are an orthonormal basis of right singular vectors of W. If we let the $\{v_{r+1}, ..., v_N\}$ be orthonormal right singular vectors of W, there exist a $v \times v$ orthonormal matrix C such that

$$[v_{r+1}, ..., v_N] = [\tilde{n}_1, ..., \tilde{n}_{N_1 - r_1}, \tilde{n}'_1, ..., \tilde{n}'_{N_2 - r_2}]C$$

We observe that

$$
\begin{aligned}
[v_{r+1},...,v_N]\,[v_{r+1},...,v_N]^T &= \left[\tilde{n}_1,...,\tilde{n}_{v_1},\tilde{n}'_1,...,\tilde{n}'_{v_2}\right] CC^T \left[\tilde{n}_1,...,\tilde{n}_{v_1},\tilde{n}'_1,...,\tilde{n}'_{v_2}\right]^T \\
&= \left[\tilde{n}_1,...,\tilde{n}_{v_1},\tilde{n}'_1,...,\tilde{n}'_{v_2}\right] \left[\tilde{n}_1,...,\tilde{n}_{v_1},\tilde{n}'_1,...,\tilde{n}'_{v_2}\right]^T \\
&= \begin{bmatrix} \tilde{n}_1 \ldots \tilde{n}_{v_1} & 0 \ldots 0 \\ 0 \ldots 0 & \tilde{n}'_1 \ldots \tilde{n}'_{v_2} \end{bmatrix} \begin{bmatrix} \tilde{n}_1 \ldots \tilde{n}_{v_1} & 0 \ldots 0 \\ 0 \ldots 0 & \tilde{n}'_1 \ldots \tilde{n}'_{v_2} \end{bmatrix}^T \\
&= \begin{bmatrix} \Gamma & 0 \\ 0 & \Upsilon \end{bmatrix}
\end{aligned}
$$

where Γ and Υ are $N_1 \times N_1$ and $N_2 \times N_2$ submatrices. This implies that αth and βth rows of $[v_{r+1},...,v_N]$ are orthogonal if p_α and p_β rows belong to different subspaces. Let $\{v_1,..,v_r\}$ be right singular vectors of W for non-zero singular values. Since that right singular matrix $V = [v_1,..,v_r,v_{r+1},..,v_N]$ is orthonormal, its rows are pairwise orthogonal. If α and β are two rows indices of V, and $\alpha \neq \beta$ we have

$$
\begin{gathered}
v_{\alpha 1}v_{\beta 1} + ... + v_{\alpha r}v_{\beta r} \\
+v_{\alpha(r+1)}v_{\beta(r+1)} ... + v_{\alpha N}v_{\beta N} = 0
\end{gathered}
$$

and since we have already shown that the summation in second row is 0, we will have

$$
v_{\alpha 1}v_{\beta 1} + ... + v_{\alpha r}v_{\beta r} = 0
$$

As a result the (α, β) rows of matrix $Q = V_r V_r^T$ is zero.

4.4 Theorem 2

Permuting columns of W does not change the set of values that appear in Q, however it changes their arrangement.

Proof: Assume we are given an initial W for which we compute its right singular matrix V^T and $Q = VV^T$. Let us create W' by swaping columns l and m for W. Right singular value of W', V' will be a result of swapping rows l and m in V. As a result, $Q'(\alpha,\beta)$ will be equal to $Q(\alpha,\beta)$ if non of α or β equal l or m. For the α or β related locations in Q', we have

$$
\begin{aligned}
Q'(\alpha, i) &= V'(\alpha,:)V'(i,:)^T \\
&= V(\beta,:)V(i,:)^T \\
&= Q(\beta, i)
\end{aligned}
$$

Similarly we will have $Q'(\beta, i) = Q(\alpha, i)$.

4.5 Theorem 3

In Theorem 1, let every r_1 points belonging to $\mathcal{T}_1$ span $\mathcal{L}_1$ and every r_2 points belonging to $\mathcal{T}_2$ span $\mathcal{L}_2$. We show that for every α, there is at least one β were $Q_{\alpha,\beta} \neq 0$.

Proof: Assume $W = [W_1|W_2]$. We will have

$$
\begin{aligned}
W &= U_r \Sigma_r V_r^T \\
[W_1|W_2] &= [U_{r_1}|U_{r_2}] \begin{bmatrix} \Sigma_{r_1} & 0 \\ 0 & \Sigma_{r_2} \end{bmatrix} \begin{bmatrix} V_{r_1}^T & 0 \\ 0 & V_{r_2}^T \end{bmatrix}
\end{aligned}
$$

and we will have

$$Q = V_r V_r^T = \begin{bmatrix} V_{r_1} & 0 \\ 0 & V_{r_2} \end{bmatrix} \begin{bmatrix} V_{r_1} & 0 \\ 0 & V_{r_2} \end{bmatrix}^T$$

$$V_r V_r^T = \begin{bmatrix} V_{r_1} V_{r_1}^T & 0 \\ 0 & V_{r_2} V_{r_2}^T \end{bmatrix}$$

Now we want to show if there is a $\alpha \in \{1, ..., N_1\}$, there is at least one $\beta \in \{1, ..., N_1\}$ for which

$$v_{\alpha 1} v_{\beta 1} + ... + v_{\alpha r} v_{\beta r} \neq 0$$

Since V_{r_1} is rank r_1, $Q_{r_1} = V_{r_1} V_{r_1}^T$ is also rank r_1. If αth row of Q_{r_1} is all zero, but at (α, α), it means row α is independent from all other rows in V_{r_1}. This means that all the other rows span a subspace with dimension $r_1 - 1$. Which additionally means that all points $p_1, .., p_{\alpha-1}, p_{\alpha+1}, .., p_{r_1}$ span a subspace with dimension r_1. This is in contrast with the assumption in Theorem 1 that every r_1 points in $\mathcal{T}_1$ span the r_1 dimensional subspace $\mathcal{L}_1$.

References

1. M. Brand. Incremental singular value decomposition of uncertain data with missing values. In *ECCV*, 2002.
2. C. Bregler, A. Hertzmann, and H. Biermann. Recovering non-rigid 3d shape from image streams. In *CVPR*, 1999.
3. T. Brox, A. Bruhn, N. Papenberg, and J. Weickert. High accuracy optical flow estimation based on a theory for warping. In *ECCV*, 2004.
4. Thomas Brox, Christoph Bregler, and Jitendra Malik. Large displacement optical flow. In *CVPR*, 2009.
5. J.P. Costeira and T. Kanade. A Multibody Factorization Method for Independently Moving Objects. *IJCV*, 29(3):159–179, 1998.
6. R. Garg, L. Pizarro, D. Rueckert, and L. Agapito. Dense multi-frame optic flow for non-rigid objects using subspace constraints. In *ACCV*, 2010.
7. C. W. Gear. Multibody grouping from motion images. In *IJCV*, 1998.
8. C. Harris and M. Stephens. A combined corner and edge detector. *Proceedings of the 4th Alvey Vision Conference*, pages 147–151, August 1988.
9. M. Irani. Multi-frame optical flow estimation using subspace constraints. In *ICCV*, pages 626–633, 1999.
10. K. Kanatani. Motion segmentation by subspace separation and model selection. In *ICCV*, 2001.
11. D. G. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 60(2):91–110, 2004.
12. D.G. Lowe. Robust model-based motion tracking through the integration of search and estimation. *IJCV*, 8:113–122, 1992.
13. B. D. Lucas and Takeo Kanade. An iterative image registration technique with an application in stereo vision. In *Intl. Joint Conf. on AI (IJCAI)*, pages 674–679, 1981.
14. C.J. Poelman and T. Kanade. A paraperspective factorization for shape and motion recovery. In *ECCV*, pages B:97–108, 1994.
15. J. Shi and C. Tomasi. Good features to track. In *CVPR*, pages 593–600, 1994.
16. C. Tomasi and T. Kanade. Shape and motion from image streams under orthography: a factorization method. *IJCV*, 9(2):137–154, Nov. 1992.
17. L. Torresani, D. B. Yang, E. J. Alexander, and C. Bregler. Tracking and modeling non-rigid objects with rank constraints. In *CVPR*, 2001.
18. B. Triggs. Factorization methods for projective structure and motion. In *CVPR*, pages 845–851, 1996.
19. R. Vidal, Y. Ma, S. Soatto, and S. Sastry. Two-View Multibody Structure from Motion. *IJCV*, 68(1):7–25, 2006.
20. Y. Wu, Z. Zhang, T. S. Huang, and J. Y. Lin. Multibody grouping via orthogonal subspace decomposition. In *CVPR*, 2001.
21. L. Zelnik-Manor and M. Irani. Degeneracies, dependencies and their implication in multi-body and multi-sequence factorization. In *CVPR*, 2003.