submitted to CVPR 2005

On Articulated Motion and Its Recovery

Jingyu Yan and Marc Pollefeys

Abstract

This paper addresses the subspace properties and the recovery of articulated motion. We point out that the nature of the motion subspace of an articulated object is a combination of a number of intersecting rigid motion subspaces. The rank of that motion subspace is less than that of each articulated part combined, depending on the connection between every two linked parts, either a rotation axis or a rotation joint. The reduced dimension(s) results from an intersection between two motion subspaces of the linked parts which is exactly the motion subspace of the axis or joint that connects them. From these observations, we describes the rank constraints of articulated motion, gives an algorithm to recover the image motion of the axis or joint and propose a novel but simple approach to recover articulated shape and motion from a single-view image sequence, which is based on subspace clustering.

1. Introduction

Shape and motion recovery deals with the problem of estimating the geometry of scene objects and the motion of the camera. It has a wide variety of applications in robotics, navigation, 3D modeling and animation etc.

Due to scene complexity, a single method of shape and motion recovery has yet been found. However, a divide-and-conquer strategy is taking shape in the past several years. By dealing with different types of motion individually, more specifically, by discovering the intrinsic properties of motion subspaces of the scene, major breakthroughs keep arriving. [1] points out that the rank of the motion subspace of a rigid scene under orthographic projection is at most 4. Using a factorization method, rigid shape and motion can be robustly recovered. [2] extends the factorization method to a paraperspective projection camera model approximating a perspective projection camera, whose motion subspace of a rigid scene remains at most of rank 4. [3] shows that the motion subspace of independently moving object are orthogonal to each other. This enables the segmentation of the motion subspace of different objects by deriving a shape invariant matrix. Once segmentation is done, the shape and motion of each object can be recovered individually. [4] shows that the motion subspace of linearly moving objects is at most of rank 6 without regard to the number of objects. A linear algorithm is described to recover the motions of the objects from this motion subspace. [5] shows that the motion subspace of a nonrigid shape can be approximated by a linear combination of the motion subspaces of a certain number of key shapes. Using a factorization method combining nonlinear iterations, the shape and motion of a nonrigid scene may be recovered.

To our knowledge, the motion subspace of an articulated object has not be discussed in detail before. We point out that the nature of the motion subspace of an articulated object is a combination of a number of intersecting rigid motion subspaces. Furthermore, we address two intrinsic subspace properties of it. One is that the largest possible rank of the motion subspace of an articulated object is less than the sum of that of each articulated part. *The correlation of the motions of two connected parts results in one or two dimensions less in motion subspace depending on the connection between the parts, either a rotation axis or a rotation joint*. The second property is that the dimension reduction results from an intersection between the two motion subspaces of the connected parts. *This subspace intersection has a physical meaning: it is the motion subspace of the rotation axis or the rotation joint*. That leads to two possibilities: the type of connection can be automatically detected by the rank of the intersection; the motion of the axis or the joint can be recovered from this intersection. We provide algorithms to achieve these.

Another advantage of viewing the motion subspace of an articulated object as a combination of a number of intersecting motion subspaces is that it enables a new and simple strategy to recover articulated shape and motion, reducing the problem into clustering and segmenting motion subspaces. Any robust reconstruction method of a rigid object can then be applied to each segmented subspace. By putting all the articulated parts into the same coordinate system, e.g. the camera coordinate, the articulated shape and motion as a whole get recovered.

Though we achieve the subspace properties and our recovery approach of articulated motion independently following the thread of research of motion subspaces of different types, we find some previous works on articulated motion recovery aiming at similar goals. The attempts to recover a rotation axis or joint as well as the articulated motion itself are not new. The originality of this paper lies in our perspective of viewing articulated motion as a set of intersecting motion subspaces and our recovery approaches whose simplicity benefits from this perspective, which only require basic operations such as subspace intersection, clustering and segmentation. We summarize others' works in the following and make comparison with ours. [8] treats each part of an articulated object as independently moving objects using [3]'s techniques at the initial stage and then apply translation constraint on top of it. During the process, they need to find the root object and solve the rest hierarchically. Our method treats each articulated part equally as different intersecting subspaces and does not need any heuristic. The translation constraint is intrinsically expressed in the intersection of motion subspaces and does not need to be imposed explicitly. [9] first needs to recover the projective structure like the rotation of the camera before they can use a minimization approach to recover a rotation axis and it does not address the case of rotation joints. Our method directly achieves the axis or joint motion subspace from the measurement data without carrying out any recovery beforehand. [10] rigorously proves the algebraic geometry relationship between articulated parts supposing different types of connections are known. Our motion subspace method does not need any prior knowledge of that and automatically detects the type of connection from the motion subspaces.

There are also a large group of articulated motion tracking researches that is based on fitting a prior model with image data. These are even further away from our approach.

Section 2 discusses the rank constraints of the articulated motion; Section 3 gives the algorithms to compute the image motion of a rotation axis and joint by intersecting motion subspaces of articulated parts; Section 4 overviews our approach to recover articulated shape and motion; Section 5 discusses our experimental results; Section 6 draws our conclusion and describe future works.

2. Subspace Properties of Articulated Motion

In this section, we introduce the rank constraints of different motions from rigid to independently moving objects. Then, we derive the rank constraint of the articulated motion matrix and describe the relationship between the motion of a rotation axis or joint and the intersection of the motion subspaces of two connected articulated parts. Our discussion is based on a weak perspective cameral model.

2.1 Rank constraint of the motion of a rigid body

The motion matrix W is a set of n tracked feature points of a rigid body across a number of f frames:

$$W = \begin{pmatrix} u_{1,1} & u_{1,2} & ... & u_{1,n} \\ v_{1,1} & v_{1,2} & ... & v_{2,n} \\ ... & ... & ... & ... \\ ... & ... & ... & ... \\ u_{f,1} & u_{f,2} & ... & u_{f,n} \\ v_{f,1} & v_{f,2} & ... & v_{f,n} \end{pmatrix} \quad (1)$$

The homogeneous world coordinates of these feature points are represented by S which we call the shape matrix:

$$S = \begin{pmatrix} x_1 & x_2 & ... & x_n \\ y_1 & y_2 & ... & y_n \\ z_1 & z_2 & ... & z_n \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

M is the camera rotation matrix for all the frames and M_i is a 2×3 matrix for the ith frame.

$$M = \begin{pmatrix} M_1 \\ M_2 \\ ... \\ M_f \end{pmatrix}$$

T is the camera translation vector for all the frames.

$$T = \begin{pmatrix} t_1^x \\ t_1^y \\ ... \\ t_f^x \\ t_f^y \end{pmatrix}$$

W, M, T and S are related by:

$$W = (M|T)S \quad (2)$$

The motion matrix of a rigid body is at most of rank 4.

2.2 Rank constraint of the motion of independently moving objects

Similar to (1), now the n tracked feature points of W belong to independently moving objects. Suppose these tracked points are grouped by the object that they belong to. Suppose there are m objects and their shape matrices are S_1, ..., S_m. Their motion matrices and translation vectors are M_1,...,M_m and T_1,...,T_m respectively. The relation of these matrices is:

$$W = (M_1|T_1|M_2|T_2|...|M_m|T_m) \begin{pmatrix} S_1 & & & \\ & S_2 & & \\ & & ... & \\ & & & S_m \end{pmatrix} \quad (3)$$

W is at most of rank $4 \times m$.

2.3 Subspace properties of the motion of an articulated object

Given a set of n tracked feature points belonging to an articulated object across a number of f frames, the rank constraint of W is what we intend to find out.

2.3.1 Rank constraint of two connected articulated parts of the object

Without loss of generality, we arrange the world coordinate such that:

- The shape of one part, S_1, are fixed.
- The z-axis of the world coordinate system coincides with the axis in the rotation axis case; the origin coincides with the joint for the rotation joint case.

Thus, for the i^{th} frame, the shape of the other part can be expressed as

$$S_{2i} = \begin{pmatrix} R_i & 0 \\ 0 & 1 \end{pmatrix} \cdot S_2$$

where R_i can be any 3 by 3 rotation matrix for the rotation joint case and

$$R_i = \begin{bmatrix} \cos\theta_i & \sin\theta_i & \\ -\sin\theta_i & \cos\theta_i & \\ & & 1 \end{bmatrix} \quad (4)$$

for the rotation axis case.

Let W_i be the image coordinates of the i^{th} frame in the motion matrix W.

$$\begin{aligned} W_i &= (M_i|T_i|M_i|T_i) \begin{pmatrix} S_1 & \\ & S_{2i} \end{pmatrix} \\ &= (M_i|T_i|M_i|T_i) \begin{pmatrix} S_1 & \\ & \begin{pmatrix} R_i & 0 \\ 0 & 1 \end{pmatrix} \cdot S_2 \end{pmatrix} \\ &= (M_i|T_i|(M_i \cdot R_i)|T_i) \begin{pmatrix} S_1 & \\ & S_2 \end{pmatrix} \end{aligned}$$

There are two observations of the above equation.

- For the rotation axis case, besides two identical T_i for each frame, it is easy to see from (4) that the last columns of M_i and $M_i \cdot R_i$ are identical to. So M is of at most rank 6.

  $$M = \begin{pmatrix} M_1|T_1|(M_1 \cdot R_1)|T_1 \\ ... \\ M_f|T_f|(M_f \cdot R_f)|T_f \end{pmatrix}$$

  Thus, the motion matrix W is of at most rank 6.

  $$W = M \begin{pmatrix} S_1 & \\ & S_2 \end{pmatrix}$$

- For the rotation joint case, M is at most of rank 7 because R_i can be any 3 by 3 rotation matrix. Thus, W is of at most rank 7.

2.3.2 The intersection of the subspaces of two connected articulated parts

Each articulated part is basically a rigid body and has a motion subspace of at most rank 4. The motion matrix of two connected articulated parts combined loses 1 or 2 dimensions, which means the two subspaces intersect. For the case of a joint connection, the 1-dimensional subspace intersection is the motion subspace of the joint. For the case of a rotation axis, the 2-dimensional subspace intersection is the motion subspace of the rotation axis. We provide the proof in the following.

Suppose the motion matrices of the two parts are W_1 and W_2. Let W_{1i} and W_{2i} be the image coordinates of the i^{th} frame in W_1 and W_2 respectively.

$$\begin{aligned} W_{1i} &= (M_i|T_i)S_1 & (5) \\ W_{2i} &= (M_i|T_i) \begin{pmatrix} R_i & 0 \\ 0 & 1 \end{pmatrix} \cdot S_2 & (6) \end{aligned}$$

- For the case of a rotation joint, from either (5) or (6), it is easy to derive the image coordinates across all the frames of the joint, whose homogeneous world coordinates is $(0, 0, 0, 1)^T$:

  $$W_{JOINT} = \begin{pmatrix} T_1 \\ ... \\ T_f \end{pmatrix}$$

  First, we are going to prove that W_{JOINT} is in the linear subspaces of both W_1 and W_2. We assume that S_1 and S_2 are not degenerate. So there is a linear combination of columns of S_1 such that $S_1 \cdot C_1 = [0, 0, 0, 1]^T$; so is for S_2 such that $S_2 \cdot C_2 = [0, 0, 0, 1]^T$. So $W_1 \cdot C_1 = W_{JOINT}$ and $W_2 \cdot C_2 = W_{JOINT}$. W_{JOINT} is in the linear subspaces of both W_1 and W_2

  Secondly, because the intersection of W_1 and W_2 is at most of rank 1. W_{JOINT} is indeed within the subspace of the intersection of W_1 and W_2.

- For the case of a rotation axis, from either (5) or (6) plus (4), the image coordinates of two points on the axis, whose homogeneous world coordinates are $(0, 0, 0, 1)^T$ and $(0, 0, 1, 1)^T$ respectively, are:

  $$\begin{aligned} W_{(0,0,0,1)^T} &= \begin{pmatrix} T_1 \\ ... \\ T_f \end{pmatrix} \\ W_{(0,0,1,1)^T} &= \begin{pmatrix} M_1(3,:) \\ ... \\ M_f(3,:) \end{pmatrix} + \begin{pmatrix} T_1 \\ ... \\ T_f \end{pmatrix} \end{aligned}$$

$M_i(3,:)$ represent the last column of M_i.

First, we are going to prove that $W_{(0,0,0,1)^T}$ and $W_{(0,0,1,1)^T}$ are in the linear subspaces of both W_1 and W_2. Similar to the above, there are linear relations between S_1, S_2 and $[0,0,0,1]^T$ and $[0,0,1,1]^T$:

$$S_1 \cdot C_{11} = [0,0,0,1]^T$$
$$S_1 \cdot C_{12} = [0,0,1,1]^T$$
$$S_2 \cdot C_{21} = [0,0,0,1]^T$$
$$S_2 \cdot C_{22} = [0,0,1,1]^T$$

Thus,

$$W_1 \cdot C_{11} = W_{(0,0,0,1)^T}$$
$$W_1 \cdot C_{12} = W_{(0,0,1,1)^T}$$
$$W_2 \cdot C_21 = W_{(0,0,0,1)^T}$$
$$W_2 \cdot C_22 = W_{(0,0,1,1)^T}$$

So $W_{(0,0,0,1)^T}$ and $W_{(0,0,1,1)^T}$ are in the linear subspaces of both W_1 and W_2.

Secondly, because the intersection of W_1 and W_2 is at most of rank 2. $span(W_{(0,0,0,1)^T}, W_{(0,0,1,1)^T})$ is indeed the intersection of W_1 and W_2. The image coordinates across all the frames of a point on the axis is in this subspace.

3 Finding rotation axes and joints in articulated motion

Before we proceed, we need to introduce some notations. Bold capitals, e.g. $\mathbf{A}$, denote a matrix. $(\mathbf{A}, \mathbf{B})$ means the concatenation of two matrices $\mathbf{A}$ and $\mathbf{B}$. $\bigoplus \mathbf{A}$ means the sums along the rows. Bold lowercase letters, e.g. $\mathbf{t}$, denotes a column vector.

Let $\mathbf{W_1}$, $\mathbf{W_2}$ be the motion matrices of two connected parts of an articulated object. We assume that there is no degenerate motion, i.e.

$$rank(\mathbf{W_1}) = rank(\mathbf{W_2}) = 4 \tag{7}$$

As discussed in Section 2.3.1,$rank((\mathbf{W_1}, \mathbf{W_2})) = 6$ for the case of a rotation axis and $rank((\mathbf{W_1}, \mathbf{W_2})) = 7$ for the case of a rotation joint. Let us discuss the prior case first and the later case is similar.

3.1 Finding the rotation axis

3.1.1 The intersection of $\mathbf{W_1}$ and $\mathbf{W_2}$

Using SVD, we can decompose $\mathbf{W_1}$ into $\mathbf{U_1} \cdot \mathbf{D_1} \cdot \mathbf{V_1}$ and $\mathbf{W_2}$ into $\mathbf{U_2} \cdot \mathbf{D_2} \cdot \mathbf{V_2}$. $\mathbf{U_i}(:, \mathbf{1} : \mathbf{4})$ $(i = 1, 2)$ represents the first 4 columns of $\mathbf{U_i}$. Let $\mathbf{c_i}$ $(i = 1, 2)$ be the two null vectors of $(\mathbf{U_1}(:, 1 : 4), \mathbf{U_2}(:, \mathbf{1} : \mathbf{4}))$, so

$$\begin{aligned} (\mathbf{U_1}(:, 1:4), \mathbf{U_2}(:, 1:4)) \cdot \mathbf{c_i} &= 0 \Rightarrow \\ \mathbf{U_1}(:, 1:4) \cdot \mathbf{c_i}(1:4) &= -\mathbf{U_2}(:, 1:4) \cdot \mathbf{c_i}(5:8) \Rightarrow \\ \mathbf{U_1}(:, 1:4) \cdot \mathbf{c_i}(1:4) &= \mathbf{U_2}(:, 1:4) \cdot -\mathbf{c_i}(5:8) \end{aligned}$$

Let $\mathbf{t_i} = \mathbf{U_1}(:, 1:4) \cdot \mathbf{c_i}(1:4) = \mathbf{U_2}(:, 1:4) \cdot -\mathbf{c_i}(5:8)$ $(i = 1, 2)$. It is easy to see that $\mathbf{t_1}$ and $\mathbf{t_2}$ are in both $\mathbf{span}(\mathbf{U_i}(:, \mathbf{1} : \mathbf{4})$ $(i = 1, 2)$, i.e. in the subspace of $\mathbf{W_i}$ $(i = 1, 2)$. Moreover, the intersection of the subspaces of $\mathbf{W_1}$ and $\mathbf{W_2}$ is of rank 2. So it is $\mathbf{span}(\mathbf{t_1}, \mathbf{t_2})$.

3.1.2 The constraint of the motion subspace of a point on the axis

We need to find a motion vector $\mathbf{t}$ in $\mathbf{span}(\mathbf{t_1}, \mathbf{t_2})$ such that it represents the image positions in all frames of a point on the rotation axis. There are infinite number of $\mathbf{t}$ corresponding to the infinite number of points on the axis.

The constraint for $\mathbf{t}$ in $\mathbf{span}(\mathbf{t_1}, \mathbf{t_2})$ is that the motion matrix $(\mathbf{W_1}, \mathbf{t})$ must be of the same rank as $\mathbf{W_1}$, which is of rank 4, because a point on the rotation axis is also a rigid part of that articulated part.

3.1.3 Computing the motion subspace of a point on the axis

In the following, we will describe our procedure to compute all the $\mathbf{t} = \alpha \cdot \mathbf{t1} + \beta \cdot \mathbf{t2}$ using this constraint. We form a new motion matrix from $\mathbf{W_1}$ and $\mathbf{t}$.

$$\mathbf{W} = (\mathbf{W_1}, \mathbf{t}) \tag{8}$$

The average of the sums of the rows of $\mathbf{W}$ is:

$$\overline{\mathbf{W}} = \frac{\bigoplus \mathbf{W}}{n+1} = \frac{\bigoplus \mathbf{W_1} + \mathbf{t}}{n+1} \tag{9}$$

After subtracting $\overline{\mathbf{W}}$ from every column of $\mathbf{W}$, $\mathbf{W}^*$ should be of rank 3.

$$\begin{aligned} \mathbf{W}^* &= \mathbf{W} - \overline{\mathbf{W}} \cdot \mathbf{1}_{1\times(n+1)} \\ &= (\mathbf{W_1}, \mathbf{t}) - \frac{\bigoplus \mathbf{W_1} + \mathbf{t}}{n+1} \cdot \mathbf{1}_{1\times(n+1)} \\ &= (\mathbf{W_1} - \frac{\bigoplus \mathbf{W_1}}{n+1} \cdot \mathbf{1}_{1\times n}, -\frac{\bigoplus \mathbf{W_1}}{n+1}) \\ &+ \mathbf{t}(-\frac{1}{n+1} \cdot \mathbf{1}_{1\times n}, \frac{n}{n+1}) \end{aligned}$$

We write $\mathbf{W}^*$ into two parts $\mathbf{W}_1^*$ and $\mathbf{W}_2^*$. $\mathbf{W}_1^*$ does not does not contain unknown $\mathbf{t}$ while $\mathbf{W}_2^*$ does.

$$\mathbf{W}_1^* = (\mathbf{W_1} - \frac{\bigoplus \mathbf{W_1}}{n+1} \cdot \mathbf{1}_{1\times n}, -\frac{\bigoplus \mathbf{W_1}}{n+1})$$

$$\begin{aligned}\mathbf{W}_2^* &= \mathbf{t}(-\frac{1}{n+1}\cdot\mathbf{1}_{1\times n}, \frac{n}{n+1}) \\ &= (\mathbf{t_1}, \mathbf{t_2})\begin{pmatrix}\alpha \\ \beta\end{pmatrix}\mathbf{K} \\ \mathbf{K} &= (-\frac{1}{n+1}\cdot\mathbf{1}_{1\times n}, \frac{n}{n+1})\end{aligned}$$

$\mathbf{W}_1^*$ is of rank 4. Using SVD,

$$\mathbf{W}_1^* = \mathbf{U}_{\mathbf{W}_1^*}\cdot\mathbf{D}_{\mathbf{W}_1^*}\cdot{\mathbf{V}_{\mathbf{W}_1^*}}^T \tag{10}$$

Then we transform $\mathbf{W}_2^*$ to the same space of $\mathbf{D_A}$.

$$\mathbf{W}_2^* = \mathbf{U}_{\mathbf{W}_1^*}(:,1:4)(\mathbf{m_1}, \mathbf{m_2})\begin{pmatrix}\alpha \\ \beta\end{pmatrix}\mathbf{n}^T\mathbf{V}_{\mathbf{W}_1^*}(:,1:4)^T$$

where

$$\begin{aligned}\mathbf{m_i} &= \mathbf{U}_{\mathbf{W}_1^*}(:,1:4)^T\cdot\mathbf{t_i} \\ \mathbf{n}^T &= \mathbf{K}\cdot\mathbf{V}_{\mathbf{W}_1^*}\end{aligned}$$

In order for $\mathbf{W}^*$ to be of rank 3, a necessary and sufficient condition is that

$$\mathbf{det}(\mathbf{U}_{\mathbf{W}_1^*}(:,1:4)^T\cdot\mathbf{W}^*\cdot\mathbf{V}_{\mathbf{W}_1^*}(:,1:4)) = 0$$

i.e.

$$\mathbf{det}(\mathbf{D}_{\mathbf{W}_1^*}(1:4,1:4)+(\mathbf{m_1},\mathbf{m_2})\begin{pmatrix}\alpha \\ \beta\end{pmatrix}\mathbf{n}^T) = 0 \tag{11}$$

We will show (11) is a linear equation of α and β. To do that, we are going to use the following rule of determinant to decompose (11):

$$\mathbf{det}(\mathbf{a}+\mathbf{b},\mathbf{c},\mathbf{d}) = \mathbf{det}(\mathbf{a},\mathbf{c},\mathbf{d}) + \mathbf{det}(\mathbf{b},\mathbf{c},\mathbf{d})$$

We rewrite (11) as

$$\mathbf{det}\begin{pmatrix}\mathbf{D}_{\mathbf{W}_1^*}(1:4,1)+(\alpha\mathbf{m_1}+\beta\mathbf{m_2})\mathbf{n}(1), \\ \mathbf{D}_{\mathbf{W}_1^*}(1:4,2)+(\alpha\mathbf{m_1}+\beta\mathbf{m_2})\mathbf{n}(2), \\ \mathbf{D}_{\mathbf{W}_1^*}(1:4,3)+(\alpha\mathbf{m_1}+\beta\mathbf{m_2})\mathbf{n}(3), \\ \mathbf{D}_{\mathbf{W}_1^*}(1:4,4)+(\alpha\mathbf{m_1}+\beta\mathbf{m_2})\mathbf{n}(4)\end{pmatrix} = 0$$

Denote $\mathbf{D}_{\mathbf{W}_1^*}(1:4,i)$ as $\mathbf{d_i}$ and $(\alpha\mathbf{m_1}+\beta\mathbf{m_2})\mathbf{n}(i)$ as $\mathbf{f_i}$. So we have

$$\mathbf{det}\begin{pmatrix}\mathbf{d_1}+\mathbf{f_1} \\ \mathbf{d_2}+\mathbf{f_2}, \\ \mathbf{d_3}+\mathbf{f_3}, \\ \mathbf{d_4}+\mathbf{f_4}\end{pmatrix} = 0$$

Keep decomposing the above. All the determinant terms that has more than one $\mathbf{f_i}$ result in 0 because all $\mathbf{f}_i$s are the same column $(\alpha\mathbf{m_1}+\beta\mathbf{c_2})$ multiplied by a different scalar $\mathbf{e}(i)$. The nonzero terms left are

$$\mathbf{det}\begin{pmatrix}\mathbf{d_1}, \\ \mathbf{d_2}, \\ \mathbf{d_3}, \\ \mathbf{d_4}\end{pmatrix} + \mathbf{det}\begin{pmatrix}\mathbf{f_1}, \\ \mathbf{d_2}, \\ \mathbf{d_3}, \\ \mathbf{d_4}\end{pmatrix} + \mathbf{det}\begin{pmatrix}\mathbf{d_1}, \\ \mathbf{f_2}, \\ \mathbf{d_3}, \\ \mathbf{d_4}\end{pmatrix} +$$

$$\mathbf{det}\begin{pmatrix}\mathbf{d_1}, \\ \mathbf{d_2}, \\ \mathbf{f_3}, \\ \mathbf{d_4}\end{pmatrix} + \mathbf{det}\begin{pmatrix}\mathbf{d_1}, \\ \mathbf{d_2}, \\ \mathbf{d_3}, \\ \mathbf{f_4}\end{pmatrix} = 0$$

This leads to the linear equation of α and β.

Finally, we have the motion vector of any point on the rotation axis:

$$\mathbf{t} = (\mathbf{t_1}, \mathbf{t_2})\begin{pmatrix}\alpha \\ \beta\end{pmatrix} \tag{12}$$

3.2 Finding the joint

Let us discuss the later case of finding the joint now. It is similar to finding the rotation axis.

3.2.1 The intersection of $\mathbf{W_1}$ and $\mathbf{W_2}$

We decompose $\mathbf{W_1}$ into $\mathbf{U_1}\cdot\mathbf{D_1}\cdot\mathbf{V_1}$ and $\mathbf{W_2}$ into $\mathbf{U_2}\cdot\mathbf{D_2}\cdot\mathbf{V_2}$. There is only one null vector of $(\mathbf{U_1}(:,1:4), \mathbf{U_2}(:,1:4))$. We denote it as $\mathbf{c_1}$. Let $\mathbf{t_1} = \mathbf{U_1}(:,1:4)\cdot\mathbf{c_1}(1:4) = \mathbf{U_2}(:,1:4)\cdot -\mathbf{c_1}(5:8)$ $(i=1,2)$. $\mathbf{t_1}$ is in both $\mathbf{span}(\mathbf{U_i}(:,\mathbf{1}:\mathbf{4})$ $(i=1,2)$, i.e. in the subspace of $\mathbf{W_i}$ $(i=1,2)$. Moreover, the intersection of the subspaces of $\mathbf{W_1}$ and $\mathbf{W_2}$ is of rank 1. So it is $\mathbf{span}(\mathbf{t_1})$.

3.2.2 The constraint of the motion subspace of the joint

To find the motion vector of the joint in the image coordinates, we need to find $\mathbf{t} = \alpha\cdot\mathbf{t_1}$. There is only one possible $\mathbf{t}$ corresponding to the uniqueness of the joint.
The constraint for $\mathbf{t}$ is that $rank(\mathbf{W_1}, \mathbf{t}) = 4$.

3.2.3 Computing the motion subspace of the joint

Our procedure to compute $\mathbf{t}$ is very similar to Section 3.1.3. The difference is

$$\mathbf{W}_2^* = \alpha\cdot\mathbf{t_1}\cdot\mathbf{K}$$

After transforming $\mathbf{W}_2^*$ to the same space of $\mathbf{D}_{\mathbf{W}_1^*}(1:4,1:4)$.

$$\mathbf{W}_2^* = \mathbf{U}_{\mathbf{W}_1^*}(:,1:4)\cdot\mathbf{m}\cdot\alpha\cdot\mathbf{n}^T\cdot\mathbf{V}_{\mathbf{W}_1^*}(:,1:4)^T \tag{13}$$

where

$$\begin{aligned}\mathbf{m} &= \mathbf{U}_{\mathbf{W}_1^*}(:,1:4)^T\cdot\mathbf{t_1} \\ \mathbf{n}^T &= \mathbf{K}\cdot\mathbf{V}_{\mathbf{W}_1^*}(:,1:4).\end{aligned}$$

In order for $\mathbf{W}^*$ to be of rank 3,

$$\mathbf{det}(\mathbf{D}_{\mathbf{W}_1^*}(1:4,1:4)+\mathbf{m}\cdot\alpha\cdot\mathbf{n}^T) = 0 \tag{14}$$

After decomposing (14) and discarding the zero determinant terms, we will have the linear equation of α like (12) where $\mathbf{d_i} = \mathbf{D}_{\mathbf{W}_1^*}(1:4,i)$ and $\mathbf{f_i} = \mathbf{m}\cdot\alpha\cdot\mathbf{n}(i)$.
The motion vector of the joint is

$$\mathbf{t} = \alpha\cdot\mathbf{t_1} \tag{15}$$

4. Articulated Motion Recovery

We summarize our approach to recover articulated motion from a single-view image sequence in this section.

- Tracking
  We use a KLT tracker to track features in image sequences and build the motion matrix. Then we impose the rank constraints of the articulated motion on the motion matrix.

- Subspace Clustering
  The motion subspace of each articulated part is intersecting with that of other parts that connect to it. These subspaces may not be orthogonal to each other. We use a generalized principle component analysis(GPCA)[6][7] to cluster and segment these subspaces.
  GPCA is an algorithm to find out a mixture of linear subspaces within a data set and then segment the data according to the underlying linear subspaces. Unlike PCA, it does not require the linear subspaces to be orthogonal to each other.

- Shape and Motion Recovery
  Each segmented subspace is the motion subspace of an articulated part. Any robust rigid motion recovery algorithm[1][2] can be used to recover the shape and motion of each part from these segmented subspaces.

4.1 Outlier Rejection

We reject outliers of feature trajectories in the motion matrix at two different stages. One is carried out on the raw motion matrix and the other, after the subspace clustering is done.

For the first outlier rejection, we impose the rank constraint of the articulated motion on the motion matrix and iteratively reject a column, i.e. one feature trajectory, that deviates from the constrained motion subspace most until the largest deviation is below a certain threshold. Without degenerate motion, the actual rank depends on the number of articulated parts, the rigid shapes(either 1D, 2D or 3D) of these parts, and their connections. It can be derived using the subspace properties of articulated motion discussed in Section 2.3. Another possibility is to automatically detect an effective rank, which will be discussed in Section 4.2.

For the second outlier rejection, we impose the rank constraint of a rigid object on each segmented motion subspace. This rank constraint may be 2, 3 or 4 corresponding to a 1D, 2D or 3D shape of the articulated part. Or we may try to detect it automatically(Section 4.2) without any prior knowledge.Again, we iteratively reject one feature trajectory that deviates from the constrained motion subspace most until

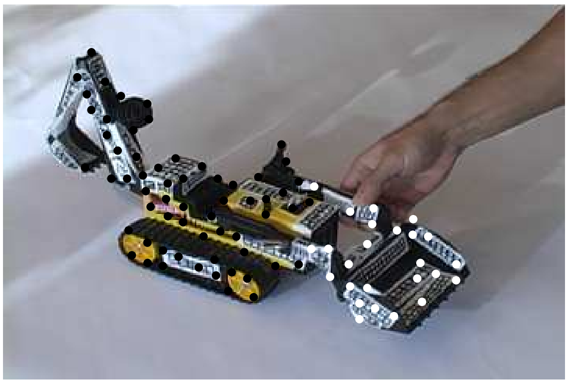

Figure 1: *The GPCA algorithm identifies two subspaces from the feature tracks and segments the tracks accordingly. The color of the feature shows the subspace it belongs to*

the largest deviation is below a certain threshold, then we stop the rejection process.

4.2 Effective Rank Detection

In practice, the motion matrix is corrupted by noise and outliers and thus its rank is usually full. Without prior knowledge of the scene and object, we may use a model selection algorithm to detect an effective rank.

$$r_n = \text{arg min}_r \frac{\lambda_{r+1}^2}{\sum_{k=1}^{r} \lambda_k^2} + \kappa\, r$$

with λ_i the i^{th} singular value of the matrix and κ a parameter.

5. Experiments

In this experiment, a toy truck with a moving shovel is videotaped. A KLT tracker successfully tracks 99 features over 70 frames while the truck moves and the shovel rotates along an axis on the truck. 87 features are left after the first stage of outlier rejection by imposing the rank constraint, rank 6 in this case, of the articulated motion. The GPCA algorithm identifies two subspaces from the feature tracks and segment the feature tracks accordingly(Figure 1). 74 features remain after the second outlier rejection by imposing the rank constraint of a rigid object to each articulated part. A rotation axis is identified by intersecting the motion subspaces of the articulated parts. Its image positions across all the frames are recovered using our algorithm(Figure 2). The shape and motion of each articulated part can be recovered from each subspace individually. But putting each part into the camera coordinates, the shape and motion of the whole articulated object gets recovered. Furthermore, with the axis recovered in the camera coordinates, we are ready

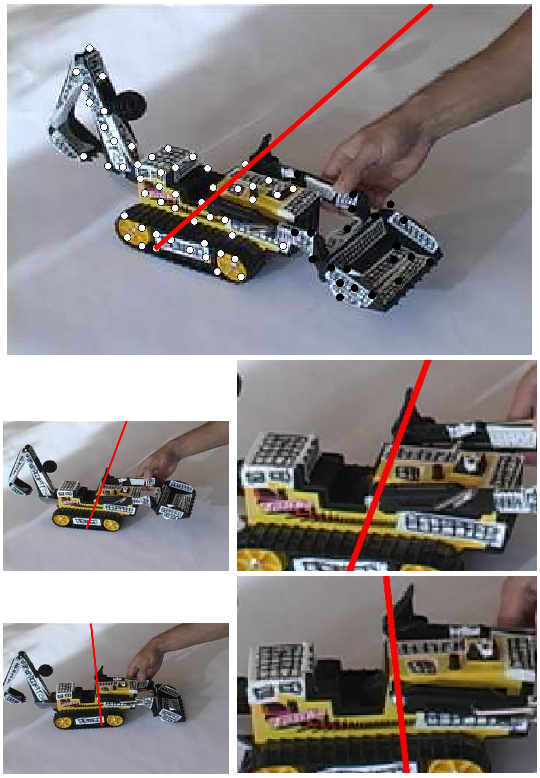

Figure 2: *(left) A rotation axis is identified by intersecting the motion subspaces of the articulated parts. (middle and right) The image coordinates of the axis across all frames gets recovered.*

to reanimate the articulated motion by rotating the articulated part around the axis and generate not only novel views but also novel motions. (Figure 3).

6. Conclusions and Future Works

With the rank constraints on articulated motion subspace, we may explore the possibility of dealing with missing data. There are also all kinds of degenerate shape and motion cases that we will look into in the future. Besides, by assuming a perspective cameral model, a wider class of scenes may be better recovered using the idea of this paper.

Acknowledgments

The partial support of the NSF ITR grant IIS-0313047 is gratefully acknowledged.

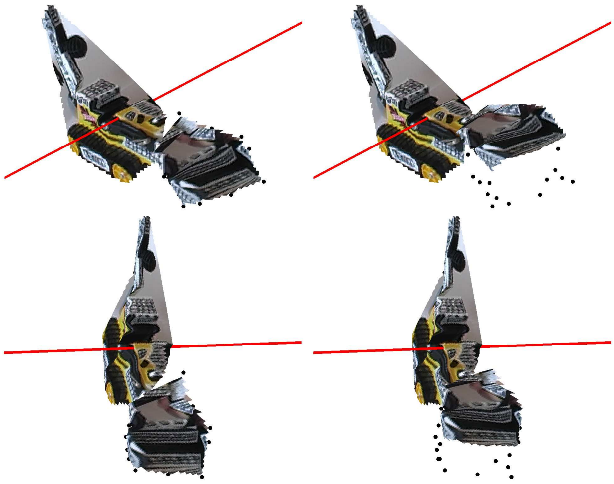

Figure 3: *The shape and motion of the truck get recovered and reanimated. Black dots shows the original position of the shovel. Not only novel views but also novel motions can be generated by rotating the shovel around the axis.*

References

[1] Carlo Tomasi, Takeo Kanade, "Shape and motion from image streams under orthography: a factorization method", *IJCV*, Vol. 9, Issue 2 pp. 137-154, 1992.

[2] C. Poelman and T. Kanade, "A Paraperspective Factorization Method for Shape and Motion Recovery", tech. report CMU-CS-92-208, Computer Science Department, Carnegie Mellon University, October, 1992.

[3] Joao Paulo Costeira, Takeo Kanade, "A Multibody Factorization Method for Independently Moving Objects", *IJCV*, Vol. 29, Issue 3 pp. 159-179, 1998.

[4] M. Han and T. Kanade, "Reconstruction of a Scene with Multiple Linearly Moving Objects", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '00), June 2000.

[5] Christoph Bregler, Aaron Hertzmann, Henning Biermann, "Recovering Non-Rigid 3D Shape from Image Streams", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '00), June 2000.

[6] Rene Vidal, Yi Ma and Shankar Sastry, "Generalized Principal Component Analysis (GPCA) ", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'03), June 2003.

[7] Rene Vidal, Yi Ma and Jacopo Piazzi, "A New GPCA Algorithm for Clustering Subspaces by Fitting, Differentiating and Dividing Polynomials", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'04), June 27 - July 02, 2004.

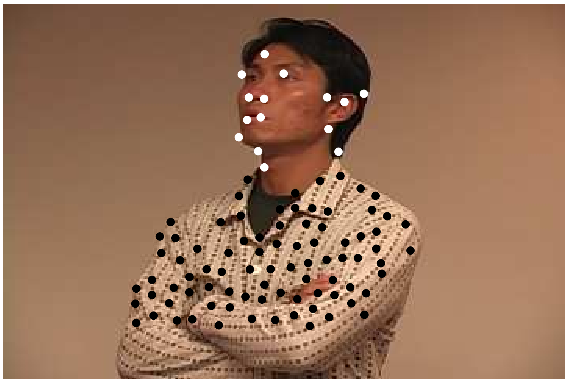

Figure 4: *The GPCA algorithm identifies two subspaces from the feature tracks and segments the tracks accordingly. The color of the feature shows the subspace it belongs to*

[8] Hanning Zhou, Thomas S. Huang, ”Recovering Articulated Motion with a Hierarchical Factorization Method”, Gesture Workshop 2003: 140-151

[9] Sinclair D., Paletta L. and Pinz A., ”Euclidean Structure Recovery through Articulated Motion”, Proc. 10th Scandinavian Conference on Image Analysis, Lappeenranta, Finland, 1997, pp 991-998.

[10] R. J. Holt, T. S. Huang, A. N. Netravali, and R. J. Qian, ” Determining articulated motion from perspective views: A decomposition approach”. Pattern Recognition, 30:14351449, 1997.

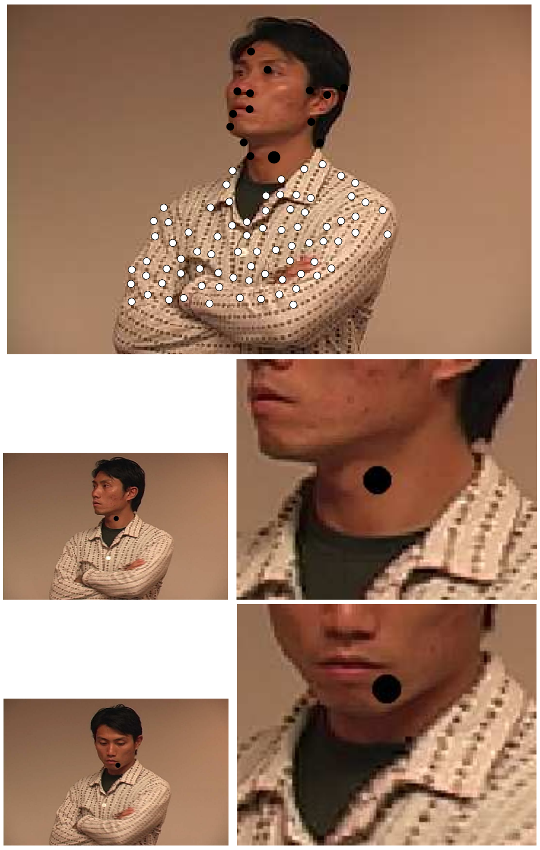

Figure 5: *(left) A rotation joint is identified by intersecting the motion subspaces of the articulated parts. (middle and right) The image coordinates of the joint across all frames gets recovered.*