Articulated Motion Segmentation Using RANSAC With Priors

Jingyu Yan and Marc Pollefeys
University of North Carolina at Chapel Hill

Abstract

Articulated motions are partially dependent. Most of the existing segmentation methods, e.g. Costeira and Kanade[2], can not be applied to articulated motions.

We propose a novel algorithm for articulated motion segmentation called RANSAC with priors. It does not require prior knowledge of the number of articulated parts. It is both robust and efficient. Its robustness comes from its RANSAC nature. Its efficiency is due to the priors, which are derived from the spectral affinities between every pair of trajectories.

We test our algorithm with synthetic and real data. In some highly challenging case, where other motion segmentation algorithms may fail, our algorithm still achieves robust results.

Though our algorithm is inspired by articulated motions, it also applies to independent motions which can be regarded as a special case and treated uniformly.

1. Introduction

Motion segmentation has been an essential issue in feature-based dynamic scene reconstruction. The problem can be described as the following: given trajectories, group those belonging to the same motion.

Lots of work has been done for independent motion segmentation[5][6][2][7][8][10] while little attention has been paid to articulated motion segmentation even though articulated motions involves one of the most interesting motions, human motions. With a proper segmentation, articulated motions can be recovered with the same ease as independent motions(Yan and Pollefeys[16], Tresadern and Reid[17]).

A naive thought may be to apply independent motion segmentation algorithms to the articulated case. However, the motions of two linked parts are partially dependent. The shape subspace of one part is not orthogonal to that of its linked part(s). Segmentation algorithms assuming independent motions generally can not be applied to articulated motions.

We propose a novel algorithm for articulated motion segmentation called RANSAC with priors. It does not require prior knowledge of the number of articulated parts. It is robust and efficient. Its robustness comes from its RANSAC nature. Its efficiency is due to the priors, which are derived

from the spectral affinities between every pair of trajectories.

We test our algorithm with synthetic and real data. In some highly challenging case, where other motion segmentation algorithms may fail, our algorithm still achieves robust results.

Though our algorithm is inspired by articulated motions, it also applies to independent motions which can be regarded as a special case and treated uniformly.

1.1. Previous Work

Lots of work has been done for motion segmentation based on the factorization method proposed by Tomasi and Kanade[12]. We will discuss the most prominent ones at this section and point out why they are not suitable for articulated motion segmentation.

Boult and Brown[5] recursively segment tracks into linearly independent motion subspaces. For articulated motions, the motion subspaces are dependent(Yan and Pollefeys[16]), which makes the criteria for segmentation invalid.

Costeira and Kanade[2] proposed a very different approach. It constructs a shape interaction matrix whose zero and nonzero entries provide strong hints for feature grouping. Later work of Weiss[10] compared several segmentation algorithms that use eigenvectors of affinity matrices for grouping and drew a unifying view of all these methods including the one in Costeira and Kanade[2] which turns out to have a root in spectral clustering.

However, articulated motions are not independent. The shape subspaces are not orthogonal to each other, which breaks the assumption of these approaches and makes the shape interaction matrix or the affinity matrix not sparse.

Another different motion segmentation approach is from Vidal[13][14][15] which propose an algebraic framework called GPCA for subspace clustering. It has been applied successfully to simple articulated motions(Yan and Pollefeys[16]).

But GPCA needs a sample size that grows exponentially with the number of subspaces. So as the number of articulated parts increases, the exponentially increasing number of trajectories required by GPCA proves to be the Archilles' heel.

Zelnik-Manor and Irani[9] briefly discusses the segmentation of partially dependent motions. They construct an

affinity matrix similar to those discussed in Weiss[10] and use an approach similar to Kanatani[8] to separate the data.

Essentially Zelnik-Manor and Irani[9] follows the segmentation approaches for independent motions and demonstrates that it may work for partially dependent motions as well. However, when the dependency between motions gets higher and the number of motions increases, it will face the same difficulties as those segmentation approaches for independent motions: the criteria for segmentation becomes ambiguous.

Our approach differs from the above work in that it adopts a RANSAC approach, which is known for its robustness, and uses a constructed affinity matrix to provide priors for random sampling. The major advantages are: it does not require prior knowledge of the number of motions; it is efficient and robust; and it provides a unified framework for motion segmentation for not only partially dependent motions but also independent motions. Like previous work, our approach assumes orthographic or weak camera projection.

The following sections are arranged in such a way: Section 2 describes articulated motion subspaces and the shape interactions of two linked parts; Section 3 describes how to derive the prior of how likely two trajectories belong to the same motion and discuss our segmentation approach, RANSAC with priors; Section 4 demonstrates RANSAC with priors using two experiments; Section 5 draws a conclusion and discusses future work.

2. Articulated Shape Subspaces and The Shape Interaction

We will discuss in detail the shape subspaces of articulated motions with comparison to those of independent motions. Then we will discuss how the shape subspaces interact for articulated motions.

2.1. Articulated Shape Subspaces vs. Independent Shape Subspaces

The articulated motion subspace is a set of intersecting rigid motion subspaces(Yan and Pollefeys[16]). They are not orthogonal to each other as independent motions are. We will show that by the following canonical factorization forms of both independent motions and articulated motions.

- Independent motions

$$W = (R_1|T_1|R_2|T_2|...|R_N|T_N) \begin{pmatrix} S_1 & & & \\ 1 & & & \\ & S_2 & & \\ & 1 & & \\ & & ... & \\ & & & S_m \\ & & & 1 \end{pmatrix} \quad (1)$$

Each motion has its own rotation and translation while the shape matrix consists of columns belonging to orthogonal shape subspaces.

- Articulated motions

There are two cases for articulated motions.

i. Two parts connected by a joint

$$W = [R_1 \ R_2 \ T] \begin{bmatrix} S_1 & 0 \\ 0 & S_2 \\ 1 & 1 \end{bmatrix} \quad (2)$$

Both motions share a translation T which is the motion of the joint, while two shape subspaces have a one-dimensional intersection.

ii. Two parts connected by an axis

$$W = [r_1 \ r_2 \ r_3 \ r_2' \ r_3' \ T] \begin{bmatrix} x_1 & x_2 \\ y_1 & 0 \\ z_1 & 0 \\ 0 & y_2 \\ 0 & z_2 \\ 1 & 1 \end{bmatrix} \quad (3)$$

where $R_1 = [r_1 \ r_2 \ r_3]$ and $R_2 = [r_1 \ r_2' \ r_3']$

Both motions share a translation T and a rotation axis. Two shape subspaces have a two-dimensional intersection.

2.2. Shape Interaction of Articulated Motions

Each trajectory has a corresponding column vector in the shape matrix which is the right most matrix in Equation (1), (2) and (3).

For independent motions(Equation (1)), column vectors of different shape subspaces have zero inner products while column vectors of the same subspace generally do not. The shape interaction matrix(Costeira and Kanade[2]) consists of these inner products of every pair of trajectories, so it can be used to group features of the same motion.

For articulated motions(Equation (2) and (3)), though the shape subspaces are not orthogonal, column vectors of the

same shape subspace generally have larger inner products then those from different shape subspaces in magnitude. We will show that in the following. The first shape subspace in Equation (2) can be represented by a base (e_1, e_2, e_3, e_7) where $e_i = [0, ..., 1, ...0]^T$ with i indicating the position of 1. Similarly, the second shape subspace can be represented by a base (e_4, e_5, e_6, e_7). It is easy to see that the inner product of column vectors from different shape subspaces has only one coefficient not canceled out while that of column vectors from the same shape subspace has four. This observation implies that the magnitude of the former is generally smaller than that of the later. A similar analysis applies to Equation (3).

So the inner products of column vectors may tell us how likely two trajectories are of the same motion. This key observation is what RANSAC with priors builds upon. And in the following section we will describe how to estimate the priors with regard to how likely every pair of trajectories belong to the same motion and then our segmentation approach, RANSAC with priors.

3. RANSAC with Priors

In this section, we will first describe how to build the priors which are to guide the RANSAC. Then we will discuss RANSAC with priors.

3.1. The Prior Matrix

Though the magnitude of the entries in the shape interaction matrix(Costeria and Kanade[2]) may be used directly for estimating how likely two trajectories are of the same motion. There is a better way.

The shape interaction matrix is actually an affinity matrix(Weiss[10]). We adopt spectral clustering approaches to analyze the affinity matrix without carrying out the clustering part. Instead, we build an affinity matrix from the normalized spectral representations of each trajectory and use it to estimate the priors of how likely every pair of trajectories are of the same motion.

The procedure is described as followed.

- Build an affinity matrix M from the trajectory matrix W: $M = W^T W$.

- Normalize M into $N = D^{-1/2} M D^{-1/2}$ where $D_{ii} = \sum_j M_{ij}$

- Form a matrix $X_{p \times k}$ whose columns are the k dominant egenvectors of N.

- Normalize each row vector of $X_{p \times k}$. This new matrix is $Y_{p \times k}$. Each row y_i of Y is the normalized spectral representation of trajectory i in R^k.

- Unlike spectral clustering which will cluster y_i into different groups at this step, we compute the affinity

between each pair of y_i and y_j and use it to build the prior matrix P.

$$P_{ij} = \frac{2}{\sqrt{\pi}} \int_0^{y_i y_j^T} e^{-t^2} dt \qquad (4)$$

P_{ij} represents the probability of trajectory i belonging to the same motion as trajectory j.

A few discussions:

- The choice of the number of eigenvectors k. Ideally, k should be the rank of N. In practice, due to noise, the rank of N can only be estimated. We may use a model selection algorithm inspired by a similar one in Vidal[15] to detect the rank.

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

with λ_i the ith singular value of the matrix and κ a parameter. If the sum of all λ_i^2 is below a certain threshold, the estimated rank is zero.

Notice that due to outliers the estimated rank may be larger than the rank of the motion subspace. However, the spectral affinity turns out to be not very sensitive to a larger k.

- Any reasonable distribution function may substitute for Equation (4). The point is to use the spectral affinity to build priors with regard to how likely two trajectories belong to the same motion.

3.2. RANSAC with Priors

P_{ij} represents the probability of trajectory(or data) i belonging to the same motion(or model) as trajectory(or data) j.

We outline our segmentation approach, RANSAC with priors, as followed.

- form a sample set of k data based on the priors P_{ij}:

1. Randomly choose the first data s_1 based on a probability distribution formed by the sums of each row of the prior matrix. The larger the row sum is, the more likely the corresponding data is in the same motion as more data and with higher probabilities.

2. Randomly choose the 2nd to the kth data, $s_2,...,s_k$, based on a probability distribution formed by the priors related to data s_1.

- instantiate a model from this sample set.

- Determine the set of data S_i that are within a threshold t of the model.

3

- Repeat this N times. The largest consensus set is selected and the model is re-estimated using all the points in that consensus set. If the largest consensus set has a size less than some threshold T, terminate.

- remove the data set S_i from the original data and repeat the above to find a new consensus set and its model until either the data is exhausted or no more models can be found from the remaining data.

A few discussions:

- The model that we use is the factorization model(Tomasi and Kanade[12]) which states that the trajectories of a full rigid motion generally span a rank-4 subspace. So k is 4 in our experiments.

- Model selection can be naturally combined with RANSAC with priors to deal with degenerate shape and motion. This will be discussed in Section 5.

4. Experiments

We test RANSAC with priors in two experiments.

The first experiment consists of a truck sequence with a moving shovel. Connected by an axis, the motion dependency is as high as it can get for articulated motions. To demonstrate the robustness of our approach, besides those erroneous trajectories due to tracking, outliers are added by adding random noise(the maximum error is 5 pixels) to some existing trajectories. The prior matrix is shown in Figure 1. The actual rank of the articulated motion subspace is 6 while the detected rank is 13 because of outliers and noise. For illustration purpose, the trajectories have been grouped into the truck body, the shovel and random outliers. Notice the priors for random outliers have very small values which makes them unlikely to be selected into a sample set. And the erroneous trajectories are rejected when RANSAC with priors tries to find a largest possible consensus set. 50 sampling times are tried each time to find a largest possible consensus set from the current data. RANSAC with priors finds 2 motions and terminates when the largest possible consensus set that it can find has a size 6 which is less than the threshold $T = 8$. Those 6 trajectories are some of the erroneous trajectories on the shovel and on the body. The remaining data consist of erroneous trajectories and the outliers that we add.

The second experiment is from a sequence of synthetic data of 4 linked parts. Each parts has 10 features to represent its 3D shape. Small random noise are added to the trajectories. And 4 outliers are added by adding larger random noise(maximum error is 5 pixels) to some trajectories.

This experiment is challenging. First, each part has a small number of trajectories which provides too few data
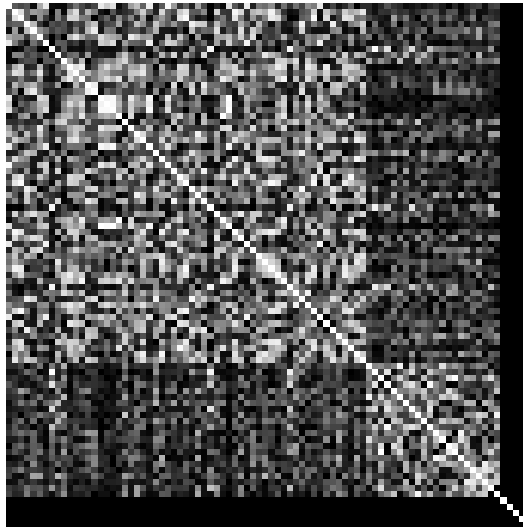

Figure 1: *The prior matrix of the truck sequence with added outliers.*

for GPCA(Vidal[13]) to work; secondly, RANSAC WITHOUT priors will require a large number of times of sampling before it may obtain a valid sample set,i.e. a sample set consisting of trajectories from the same part. For example, only $\frac{1}{4} \times \frac{1}{4} \times \frac{1}{4} \times \frac{1}{4} = \frac{1}{256}$ may be valid sample sets in this experiment. In practice, without knowing the number of motions beforehand, it is impossible to set a fixed threshold for the number of sampling times.

However, RANSAC with priors generally gets one valid sample set out of every three in this experiment. And this rate does not depend on the total number of motions. It depends on the number of dependent motions. Parts that are further away generally have much less dependency which make the corresponding priors very small, thus their trajectories are unlikely to be chosen into a sample set. 50 sampling times are tried each time to find a largest possible consensus set from the current data in this experiment. The prior matrix of 4 linked parts is shown in Figure 3). The actually rank of the articulated motion subspace is 13 but the detected rank is 17 due to outliers. RANSAC with priors finds 4 motions within the trajectories and the segmentation is shown in Figure 4) with reference lines representing each part of the object for better illustration. The remaining data are 4 random outliers after RANSAC with priors can not find any consensus set of size more than $T = 8$. The result matches the ground truth.

At last, we put RANSAC with priors into the test of a more complex scenario. Using our approach, independent motions are only a special case and can be treated in the same fashion. The prior matrix for two independently moving articulated objects from a real sequence is shown in Figure 5). Each of these two articulated objects has two parts.
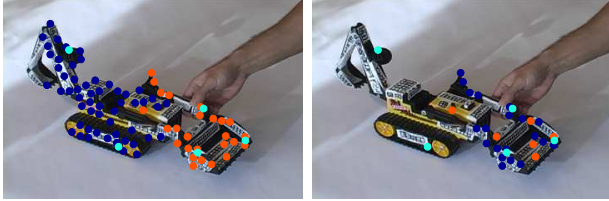
Figure 2: *(left) RANSAC with priors finds the first consensus set indicated by blue dots. The orange and light blue dots are the remaining data. The light blue dots are added outliers. The orange dots on the truck body are erroneous trajectories. (right) The blue dots indicate the second consensus set found by RANSAC with priors. The orange and light blue dots are the remaining data. The light blue dots are the added outliers. The orange dots are erroneous trajectories.*

Notice the priors between every pair of trajectories from different objects are very small. RANSAC with priors is able to segment 4 motions from these trajectories.

5. Conclusions and Future Work

We describe and demonstrate a motion segmentation algorithm called RANSAC with priors. It can segment articulated motions as well as independent motions. It does not require prior knowledge of how many motions there are. It is both efficient and robust. The priors are derived from the spectral affinity between every pair of trajectories.

Future work will involve combining model selection to deal with degenerate shape and motion. After having a sample set, we can estimate the model from several models. Furthermore, with priors, we may even consider forming a larger size of a sample set. This will not increase the computation too much as compared to common RANSAC WITHOUT priors because with the help of the priors the sample set has a far better chance consisting of data belonging to the same model.

We also plan to apply RANSAC with priors to highly challenging cases of complex articulated motions like human motions and complex scenes consisting of partially dependent and independent motions.

References

[1] M. A. Fischler, R. C. Bolles. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. Comm. of the ACM, Vol 24, pp 381-395, 1981.

[2] J.P. Costeira, T. Kanade, "A Multibody Factorization Method for Independently Moving Objects", *IJCV*, Vol. 29, Issue 3 pp. 159-179, 1998.

Figure 3: *The prior matrix of multiple linked parts with outliers*

[3] Ullman, S. 1983. Maximizing rigidity: The incremental recovery of 3D structure from rigid and rubbery motion. Technical Report A.I. Memo No. 721, MIT.

[4] Sinclair, D. 1993. Motion segmentation and local structure. In Proceedings of the 4th International Conference on Computer Vision.

[5] Boult, T. and Brown, L. 1991. Factorization-based segmentation of motions. In Proceedings of the IEEE Workshop on Visual Motion.

[6] Gear, C.W. 1994. Feature grouping in moving objects. In Proceedings of the Workshop on Motion of Non-Rigid and Articulated Objects, Austin, Texas

[7] N. Ichimura. Motion segmentation based on factorization method and discriminant criterion. In Proc. IEEE Int. Conf. Computer Vision, pages 600605, 1999.

[8] K. Kanatani. Motion segmentation by subspace separation and model selection:model selection and reliability evaluation. Intl. J. of Image and Graphics, 2(2):179197, 2002.

[9] L. Zelnik-Manor and M. Irani. Degeneracies, dependencies and their implications in multi-body and multi-sequence factorizations. In Proc. IEEE Computer Vision and Pattern Recognition, 2003.

[10] Y. Weiss. Segmentation using eigenvectors: A unifying view. In International Conference on Computer Vision, pages 975982, Corfu, Greece, September 1999.

[11] Ng, M. Jordan, and Y. Weiss. On spectral clustering: analysis and an algorithm. In Advances in Neural Information Processing Systems 14. MIT Press, 2002.
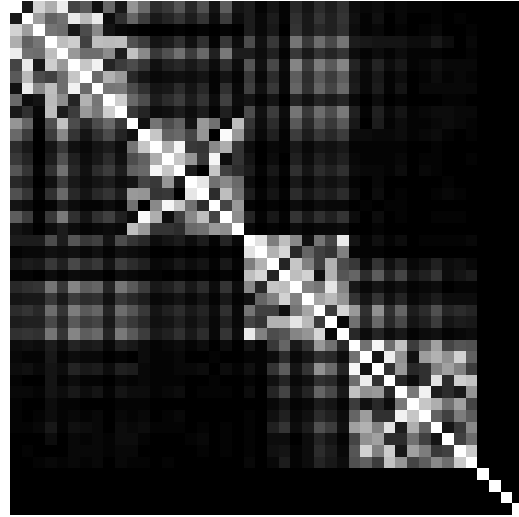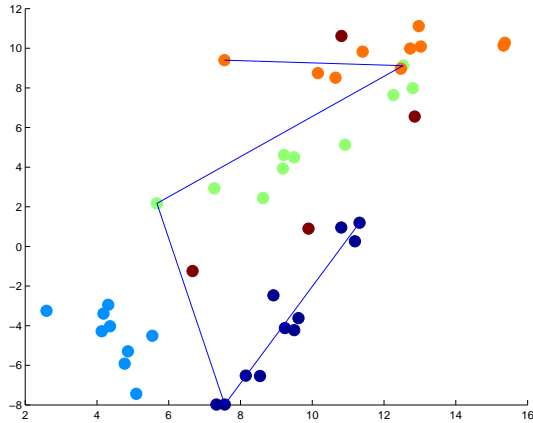
Figure 4: *The segmentation result of* 4 *linked parts using RANSAC with priors. Orange, green, light blue and dark blue indicate* 4 *parts of the articulated object.* 4 *Red dots are the remaining data rejected by the algorithm, which are the added outliers.*

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

[13] R. Vidal and R. Hartley. Motion Segmentation with Missing Data using PowerFactorization and GPCA. IEEE Conference on Computer Vision and Pattern Recognition, 2004

[14] R. Vidal, Y. Ma and S. Sastry, "Generalized Principal Component Analysis (GPCA) ", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'03), June 2003.

[15] R. Vidal, Y. Ma and J. 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.

[16] J. Yan, M. Pollefeys, A Factorization-based Approach to Articulated Motion Recovery, IEEE Conf. on Computer Vision and Pattern Recognition, 2005

[17] P. Tresadern and I. Reid, Articulated Structure From Motion by Factorization, Proc IEEE Conf on Computer Vision and Pattern Recognition, 2005

[18] Multiple View Geometry in Computer Vision, Richard Hartley and Andrew Zisserman, Cambridge University Press, 2002
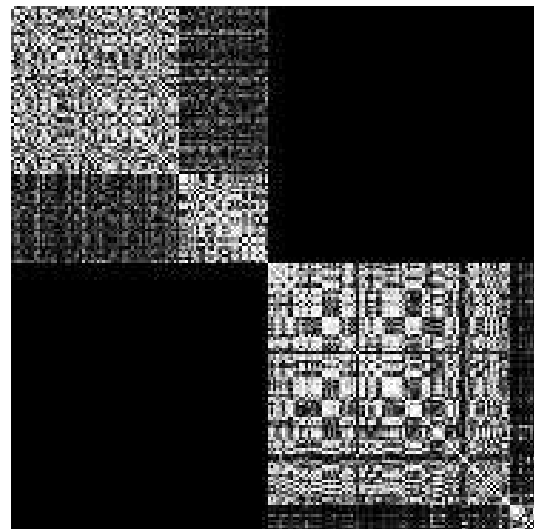
Figure 5: *The prior matrix of two independently moving articulated objects. One articulated object has two parts linked by an axis. The other object has two parts linked by a joint*