Multiframe Motion Segmentation via Penalized MAP Estimation and Linear Programming

Han Hu
huh04@mails.tsinghua.edu.cn

Quanquan Gu
gqq03@mails.tsinghua.edu.cn

Lei Deng
dally211@163.com

Jie Zhou
jzhou@tsinghua.edu.cn

State Key Laboratory on Intelligent Technology and Systems
Tsinghua National Laboratory for Information Science and Technology (TNList)
Department of Automation, Tsinghua University, Beijing, China, 100084

Abstract

Motion segmentation is an important topic in computer vision. In this paper, we study the problem of multi-body motion segmentation under the affine camera model. We use a *mixture of subspace* model to describe the multi-body motions. Then the motion segmentation problem is formulated as an MAP estimation problem with model complexity penalty. With several candidate motion models, the problem can be naturally converted into a linear programming problem, which guarantees a global optimality. The main advantages of our algorithm include: It needs no priori on the number of motions and it has comparable high segmentation accuracy with the best of *motion-number-known* algorithms. Experiments on benchmark data sets illustrate these points.

1 Introduction

In various computer vision problems, we often confront dynamic scenes, in which both the camera and multiple objects move: e.g. traffic surveillance of a busy crossroad, television relaying of a sport game, DV recording using a hand-held camera, etc. Hence, there has been an increasing interest on the understanding and reconstructing of dynamic scenes in recent years. As an essential issue of this area, motion segmentation problem has been studied and many techniques have been proposed. Most of these approaches are based on the observation that the trajectory of each rigid motion lies in a linear subspace of four dimension or less, under the affine camera model. They can be roughly grouped into three types: factorization based, statistical and algebraic.

Most of the early researches on motion segmentation are based on subspace factorization or its variants [1, 3]. However, the performance of these algorithms drops quickly in the presence of degenerate and partially dependent motions. This short coming has led to the development of many statistical motion segmentation algorithms. Some of the algorithms

formulate the segmentation as a clustering problem and handle it with Expectation Maximization (EM) [7, 12], which carries out by recursively estimating parameters and segmentation membership. Other statistical methods use local information as a similarity metric, and make some achievements in segmentation accuracy [2, 12]. The interesting algebraic methods reveal another direction for this problem. Generalized Principal Component Analysis (GPCA) [15] is the best known method of this type. It receives much concern for its elegant style and the weak restriction on the relative orientations of the motion subspaces. However, the complexity and trajectories required for the method increase dramatically when the number and dimension of motions raise, which greatly limits the application of this method.

In this paper, we present a novel algorithm for multi-frame motion segmentation problem under the affine camera model. The algorithm is based on a *mixture-of-subspace* model, which describes the multi-body motions in a unified style. Using this model, the motion segmentation problem can be formulated as a maximum a posteriori estimation (MAPE) problem taking account of model complexity. To solve this problem, a list of candidate motion models is generated by a certain scheme. Then the problem is converted into a linear programming problem and therefore can be effectively solved. The main advantages of our algorithm include: (1) the algorithm needs no priori on the number of motions; (2) since MAP estimator incorporates prior information into the optimization problem, it is potentially more accurate than conventional methods; (3) the linear programming formulation guarantees a global optimality. Experiments on benchmark motion segmentation data sets illustrate that as long as the true motions are contained by the candidate motion models, the proposed method is very effective.

The rest of this paper is organized as follows: Section 2 presents the proposed method; Section 3 shows the experimental results on benchmark real data sets. Section 4 gives conclusions.

2 The Proposed Method

In this section, we first review the geometry of a single motion model. Then, a *mixture of subspace* model is presented for multi-body motions. After that, the motion segmentation problem is formulated as a maximum a posteriori estimation (MAPE) problem taking account of model complexity. Finally, a linear programming technique is adopted to solve the penalized MAPE problem with several derived candidate motion models.

2.1 A Brief Review of Single Rigid Motion Model

Let $\{(u_{fi}, v_{fi})^T \in \mathbb{R}^2\}_{f=1,\ldots,F}^{i=1,\ldots,P}$ be the 2-D projections in F frames of P 3-D points $\{(X_i, Y_i, Z_i)^T\}_{i=1,\ldots,P}$ from a rigid object. Under the affine camera model, the trajectories and their 3-D points satisfy the following equation,

$$W = \begin{bmatrix} u_{11} & u_{12} & \ldots & u_{1P} \\ v_{11} & v_{12} & \ldots & v_{1P} \\ \vdots & \vdots & \ddots & \vdots \\ u_{F1} & u_{F2} & \ldots & u_{FP} \\ v_{F1} & v_{F2} & \ldots & v_{FP} \end{bmatrix}_{2F \times P} = \begin{bmatrix} A_1 \\ \vdots \\ A_F \end{bmatrix}_{2F \times 4} \begin{bmatrix} X_1 & \ldots & X_P \\ Y_1 & \ldots & Y_P \\ Z_1 & \ldots & Z_P \\ 1 & \ldots & 1 \end{bmatrix}_{4 \times P} = MS, \quad (1)$$

where $A_i = K_i \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_i & \mathbf{t}_i \\ \mathbf{0}^T & 1 \end{bmatrix} \in \mathbb{R}^{2 \times 4}$ is an affine matrix at frame i, which

depends on the camera intrinsic parameters K_i and the object pose relative to the camera $(R_i, \mathbf{t}_i)$.

Eq.(1) indicates that,

$$2 = rank(A_i) \leq rank(W) \leq 4. \tag{2}$$

Eq.(2) assumes that trajectories from the same rigid motion lie in a linear subspace of $\mathbb{R}^{2F}$ with dimension between two and four. Denote the dimension of the motion by r, then we can use a set of orthonormal vectors $C = \{\mathbf{u_i}\}_{i=1}^r$ to completely represent a motion model. The orthonormal vectors can be obtained from the first r left singular vectors of U by SVD factorization [4],

$$W = U_{2F \times 2F} \Sigma_{2F \times P} V_{P \times P}^T. \tag{3}$$

2.2 Mixture of Subspace Model for Multibody Motions

For multi-body motion segmentation problem, the membership between trajectories and motion models is unknown, and the model parameters are also unknown. Our goal is to establish an optimization framework with a unified cost function, for all possibilities. To this end, one way is to construct the product of all possible motion memberships: $\prod_{j=1}^K d(\mathbf{w}, C_j) = 0$ (the motion number K is known as a priori), and then use a scheme to fit this constraint. Here, the distance from a trajectory $\mathbf{w}$ to a motion model C is defined as [4],

$$d(\mathbf{w}, C) = ||C_\perp \mathbf{w}||, \tag{4}$$

where $C_\perp = I - \sum_{i=1}^r \mathbf{u_i} \mathbf{u}_i^T$ is the orthogonal complement space of C in $\mathbb{R}^{2F}$ space. The GPCA algorithm [15] is of this kind. When there are sufficient samples with no noise, the algorithm can give a pretty good result. However, limited by the nonlinear form and a large quantity of parameters, the fitting procedure can hardly be effective and accurate.

An alternative way is to define indicator variables $L_{ij}, (i = 1, \dots, p; j = 1, \dots, K)$, of which $L_{ij} = 1$ means the ith trajectory belongs to the jth model, and $L_{ij} = 0$ otherwise. Then a *mixture of subspace* model can be formulated for each trajectory $\mathbf{w}_i$ as follows,

$$\begin{aligned} & \sum_{j=1}^K L_{ij} d(\mathbf{w}_i, C_j) = 0, \\ & s.t. \ \sum_{j=1}^K L_{ij} = 1, L_{ij} \in \{0, 1\}, \end{aligned} \tag{5}$$

2.3 Penalized Maximum a Posteriori Estimation (PMAPE)

If the number of motions K is known, then the maximum a posteriori estimation (MAPE) method can be used to solve C_j and L_{ij}.

Assume the noises are of Gaussian distribution, and the standard deviation of noise on the jth model is σ_j. Given motion models $\{C_j\}_{j=1,\dots,K}$ and memberships L_{ij}, the conditional probability of a trajectory $\mathbf{w}_i$ is,

$$\begin{aligned} & p(\mathbf{w}_i | C_j, L_{ij}, j = 1, \dots, K) \\ & = \sum_{j=1}^K L_{ij} \cdot \frac{1}{\sqrt{2\pi}\sigma_j} \exp(-\frac{d^2(\mathbf{w}_i, C_j)}{\sigma_j^2}). \end{aligned} \tag{6}$$

Given the prior probability of C_j and L_{ij}, under the independent assumption, the posteriori probability is,

$$
\begin{aligned}
&p(C_j, L_{ij}, i=1,...,P; j=1,...,K | \mathbf{w}_i, i=1,...,P) \\
&= \frac{p(\mathbf{w}_i, i=1,...,P | C_j, L_{ij}, i=1,...,P; j=1,...,K)p(C_j, L_{ij}, i=1,...,P; j=1,...,K)}{p(\mathbf{w}_i, i=1,...,P)} \\
&= \frac{\prod_{i=1}^{P}(p(\mathbf{w}_i | C_j, L_{ij}, j=1,...,K)p(C_j, L_{ij}, j=1,...,K))}{p(\mathbf{w}_i, i=1,...,P)}
\end{aligned}
\tag{7}
$$

The prior $p(\mathbf{w}_i, i=1,...,P)$ is constant irrespective of the choice of C_j and L_{ij}. Thus after some algebraic simplifications, we reach the log-MAPE cost function:

$$
\begin{aligned}
\max_{C,L,K} \ln L &= \max_{C,L,K} \ln p(C_j, L_{ij}, i=1,...,P; j=1,...,K | \mathbf{w}_i, i=1,...,P) \\
&= \max_{C,L,K} \sum_{i=1}^{P}(\ln(p(\mathbf{w}_i | C_j, L_{ij}, j=1,...,K)) + \ln(p(C_j, L_{ij}))) \\
&= \max_{C,L,K} \sum_{i=1}^{P}(\ln(\sum_{j=1}^{K} L_{ij} \cdot \frac{1}{\sqrt{2\pi}\sigma_j} \exp(-\frac{d^2(\mathbf{w}_i, C_j)}{\sigma_j^2})) + \ln p(C_j, L_{ij})) \\
&= \min_{C,L,K} \sum_{i=1}^{P} \sum_{j=1}^{K} L_{ij}(\frac{d^2(\mathbf{w}_i, C_j)}{2\sigma_j^2} + \ln(\sigma_j) - \ln p(C_j, L_{ij}))
\end{aligned}
\tag{8}
$$

Using MAPE assumes that the number of motions K is known as a priori and motions with different dimensions are equal. However, in practice, K is usually unknown, and the motions with different dimensions cannot be treated as the same. One can easily reach the latter conclusion by thinking the fact that, if several trajectories belong to a 3-dimensional model, a 4-dimensional model will fit the trajectories well as long as the former model is contained by the latter one.

The log-likelihood cost function certainly rises when K grows or the dimension of motions increases. Hence, a tradeoff between fitting error and model complexity should be adopted. Several criteria can touch this purpose, e.g. Akainke information criterion (AIC) and Bayesian information criterion (BIC) [6]. Then we can obtain the cost function as,

$$
\mathscr{J} = -\ln L + \alpha \sum_{j=1}^{K} Pr_j
\tag{9}
$$

where Pr_j indicates the complexity of the jth model, and α is a penalizing factor. In this work, since the two parts of the cost function are at similar scale level, α may be chosen from 0.1 to 10. Therefore, we reach the following penalized MAPE problem:

$$
\min_{C,L,K,r} \sum_{i=1}^{P} \sum_{j=1}^{K} L_{ij}(\frac{d^2(\mathbf{w}_i, C_j)}{2\sigma_j^2} + \ln(\sigma_j) - \ln p(C_j, L_{ij})) + \alpha \sum_{j=1}^{K} Pr_j
\tag{10}
$$
$$
s.t. K \geq 1, K \in \mathbb{Z}, \sum_{j=1}^{K} L_{ij} = 1, L_{ij} \in \{0,1\}, r_j = \{0,2,3,4\}.
$$

2.4 Linear Programming Formulation

Eq.(10) is a combinatorial optimization problem, which is NP hard. In this paper, a method based on linear programming relaxation is adopted to solve the problem.

Revisit the cost function of Eq.(10). Suppose that somehow we have already obtained a list of candidate motion models $\Phi\{C_1,...,C_N\}$, and the K true motions are contained by the

list. Define indicating variables x_j with $x_j = 1$ if the jth candidate motion is a true motion and $x_j = 0$ otherwise. Obviously we have $x_j = \max_{1 \leq i \leq P}\{L_{ij}\}$. Denote the rank of the jth candidate motion by r_j, and then the cost function of Eq.(10) can be rewritten as,

$$\min_{C,L,x}\sum_{i=1}^{P}\sum_{j=1}^{N}L_{ij}\left(\frac{d^2(\mathbf{w}_i,C_j)}{2\sigma_j^2}+\ln(\sigma_j)-\ln p(C_j,L_{ij})\right)$$
$$+\alpha P\left(2\sum_{\{j:r_j=2\}}x_j+3\sum_{\{j:r_j=3\}}x_j+4\sum_{\{j:r_j=4\}}x_j\right) \tag{11}$$

Since we assume that the N candidate motion models have been obtained, the terms of $d^2(\mathbf{w}_i,C_j)$, σ_j and prior probability $p(C_j,L_{ij})$ can all be pre-computed and they together can be considered as coefficients of the variables L_{ij}. Now the problem becomes linear to the unknown variables L and x.

However, due to the integer constraints on L and x, Eq.(10) is still NP-hard. In order to solve the problem, we exploit the idea of continuous relaxation, which is a popular technique in Operational Research [5]. Applying this idea to Eq.(10), we relax the variables L and x as:

$$0 \leq L_{ij} \leq 1, \ 0 \leq x_{ij} \leq 1, \forall i, j. \tag{12}$$

Then the ultimate optimization problem can be written as:

$$\min_{C,L,x}\sum_{i=1}^{P}\sum_{j=1}^{N}L_{ij}\hat{d}^2(\mathbf{w}_i,C_j)+\alpha P\left(2\sum_{\{j:r_j=2\}}x_j+3\sum_{\{j:r_j=3\}}x_j+4\sum_{\{j:r_j=4\}}x_j\right)$$
$$s.t. \quad \sum_{j=1}^{N}L_{ij}=1,\forall i;$$
$$L_{ij}\leq x_j,\forall i,j; \tag{13}$$
$$0\leq L_{ij}\leq 1, \ 0\leq x_{ij}\leq 1,\forall i,j,$$

where $\hat{d}^2(\mathbf{w}_i,C_j) = \frac{d^2(\mathbf{w}_i,C_j)}{2\sigma_j^2}+\ln(\sigma_j)-\ln p(C_j,L_{ij})$ is *normalized distance* between $\mathbf{w}_i$ and the C_j. And the constraint $L_{ij}\leq x_j$ is from $x_j = \max_{1\leq i\leq P}\{L_{ij}\}$ by certain algebraic transformation.

Rounding scheme. Now we have converted the combinatorial optimization problem into a linear programming problem. It is further hoped that the continuous variables $\widehat{L}$ and $\widehat{x}$ is 0 or 1. Thus a rounding procedure should be carried out. Many schemes can be chosen for this purpose [5]. In this paper, a simple strategy is adopted: let $L_{ij}=1$ if $\widehat{L}_{ij}=\max_{k=1,..,P}\{\widehat{L}_{kj}\}$; and $L_{ij}=0$ otherwise.

2.5 Candidate Motions

Generating Candidate Motions. Our scheme is as follows: Firstly, we fit 3 single local models with rank-2, rank-3, and rank-4 respectively for each trajectory by itself and its k-nearest neighbors, of which the distance is measured by cos-angle of two trajectories.

After that, we have obtained quantities of models, thus a model clustering scheme is adopted to merge duplicated models together. To achieve this goal, a similarity matrix of all generated models is computed, and then spectral clustering [9] is applied using the similarity matrix. Since each model is a subspace, we define the similarity of two models as,

$$s(C_\alpha,C_\beta)=\begin{cases} \exp(-\sum_{i=1}^{M}\sin^2(\theta_i)), & r_\alpha = r_\beta, \\ 0, & r_\alpha \neq r_\beta, \end{cases} \tag{14}$$

where $\theta_1 \leq ... \leq \theta_M$ are the principal angles [4] between two subspaces C_α, C_β , and M is the minimal dimension of two subspaces. Here we set the similarity of subspaces with different dimensions to be 0 is for the purpose of preserving low-rank motion models.

Lastly, for each cluster, we calculate the score of each model and choose the model with the highest score. The score is computed in this way:

1) calculate the inlier set by estimating the probability density function of the distances and discarding trajectories outside the first mode of that density. Due to lack of space, we refer the reader to [16] for details.

2) Calculate the standard deviation of noise by the inliers as,

$$\sigma_j = \frac{1}{N_{inliers(j)}} \sum_{i \in inliers(j)} d^2(\mathbf{w_i}, C_j). \tag{15}$$

3) Sort the models of the same cluster with ascending number of inliers and descending standard deviation respectively.

4) Record the sort orders D_1 and D_2 , then the score is,

$$score = D_1 + D_2. \tag{16}$$

The number of clusters is preset, e.g. 24 or 18 in our experiments. Thus we have obtained a list of candidate motion models. However, this scheme is not essential. Other random sample methods such as [8, 10] can also be used here to generate motion candidates.

Noise and Prior Probability Estimation. In Eq.(16), we have obtained the standard deviation of noise. The prior probability is simply set to be proportion to the number of inliers, namely,

$$p(C_j, L_{ij}) \propto N_{inliers(j)}. \tag{17}$$

2.6 PMAPE-LP Algorithm and Discussions

The proposed PMAPE-LP algorithm is listed below.

Algorithm 1. Penalized MAP Estimation and Linear Programming (PMAPE-LP)

---

1. Generate a list of N candidate motion models by a certain scheme. In this step, one must ensure that the true motion models are indeed contained in the list or very similar to one of the candidates.

2. Estimate the inlier set of each model, and calculate the standard deviation of noise and the prior probability as:

$$\sigma_j = \frac{1}{N_{inliers(j)}} \sum_{i \in inliers(j)} d^2(\mathbf{w_i}, C_j), \ p(C_j, L_{ij}) \propto N_{inliers(j)}.$$

3. Compute the normalized distance,

$$\hat{d}^2(\mathbf{w}_i, C_j) = \frac{d^2(\mathbf{w}_i, C_j)}{2\sigma_j^2} + \ln(\sigma_j) - \ln p(C_j, L_{ij}),$$

where $d(\mathbf{w}, C)$ is defined by Eq.(4).

4. Construct linear programming problem as Eq.(13), and solve it with any LP solver, e.g. Matlab's linprog in our experiment.

5. Round the LP solutions to binary variables, and output the groups.

Discussions.

1. In most previous works, the noises are set to be constant all over the sequence [8, 11, 12, 16]. However, it is more reasonable to think of the noises to be changeable among the candidate motion models, for the following reasons: (1) those who assume the noises are constant only consider the feature detector localization error. In fact, for real applications, some other effects should also be considered, e.g. perspective effects which are not accounted for by the affine model and act on the motion models in different degrees. (2) The estimated candidate models certainly cannot completely accord with the true motions, and the estimation errors of candidate motions are not the same with each other. (3) Even if the position errors of all trajectories are constant, the noises may still vary over the motions. Consider the motions with different ranks. It is obviously that the models with lower rank have higher noise level for their low degree of freedom.

2. The introducing of prior probabilities possibly improves the segmenting accuracy.

3. As long as the true motions are included in the candidate list, it is guaranteed that the algorithm will converge to the global optimum. In addition, linear programming is generally considered high efficient.

4. If two candidate motions are similar to each other, then the linear programming will surely pick the better one.

3 Experimental Results

We compare our algorithm (referred to as PMAPE-LP) with the state-of-the-art affine methods on large scale real data sets. These methods include:

1. Local Subspace Affinity (LSA) [14]: this method fits a subspace around each trajectory locally, and segments them by applying spectral clustering to an affinity matrix, which is built from the principal angles among subspace pairs.

2. GPCA [15]: this algebraic approach models K-body motion as a polynomial of degree K and fits it using a large quantity of trajectories. Then the differential of each trajectory is calculated and applied as a similarity metric for further clustering.

3. Multi-Stage Learning (MSL) [7]: this statistical approach uses the Expectation Maximization algorithm to get the segmentation results.

Before applying LSA, we project the trajectories onto a subspace of dimension 5 or 4K, where K is the number of motions, as suggested in [14]. We refer to the two versions as LSA-5 and LSA-4K.

In the first experiment, we mainly verify the effectiveness of our algorithm for estimating the number and dimensionality of a real sequence with complex motions. In the second experiment, we mainly evaluate the segmentation accuracy of our algorithm on a large scale benchmark data set (Hopkins 155).
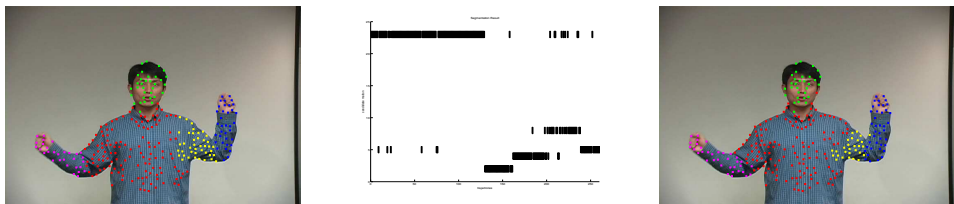
Figure 1: Multi-body motion segmentation experiment on dancing sequence: 5 motions, 268 trajectories. L: input sequence image and the ground truth of segmentation. M: Obtained segmentation membership. R: The motion segmentation result of PMAPE-LP.

Method	PMAPE-LP	LSA-5	LSA-4K	GPCA
Misclassification Rate	**9.33**	42.16	10.45	Not enough Samples [1]

Table 1: The misclassification rates (%) on dancing sequence.

3.1   Motion Number Estimation on Dancing Sequence

Dancing sequence [17] comes from a scene with a person dancing and has 5 motions, including one non-rigid motion and four other articulated motions (see Figure 1L). The motions are of rank-6, rank-3, rank-3, rank-3, and rank-3, respectively. It is a much challenging case due to the number and complexity of motions.

Since non-rigid motions are involved in this sequence, we set the maximum rank of motion models to be 6. In order to reduce the computational time, we projected the trajectory data onto a 8-dimensional subspace. The initial number of models was set to be 24. The number of trajectories used for fitting a subspace in the initial step was set to be 8, and α was set as 0.4.

The obtained motions come from the $2, 4, 5, 8, 23^{th}$ models in the candidate list(see Figure 1M), which are completely correct as well as the estimated dimensionality. We also compared PMAPE-LP with the *motion-number-known* methods, e.g. LSA-5, LSA-4K and GPCA on segmentation accuracy. Table 1 shows the results of misclassification rates given by different methods. Figure 1R shows the results of motion segmentation on the first frame of the sequence. From the results we can see that the proposed method PMAPE-LP performs best among these methods.

3.2   Test on Hopkins 155 database

The Hopkins 155 database is a large scale benchmark real data set for the evaluation of motion segmentation algorithms [13]. The entire database is available at *http://www.vision.jhu.edu*, which consists of 155 motion sequences, involving checkerboard, traffic, and other special topics (See Figure 2 for some samples). Among these sequences, some have two motions and others have three motions. The point trajectories are also provided in the database as well as the ground truth, which are obtained by a tracking tool implemented in OpenCV, a library freely available at *http://sourceforge.net/projects/opencvlibrary*.

Since most motion segmentation algorithms are designed under the assumption that the number of motions is known as a priori, in order to compare PMAPE-LP with them in the

---

[1]GPCA requires a large number of trajectories to work. Here, the number of motions is 5, thus at least $N = [(K+2)(K+1)/2]^2 - 1 = 440$ trajectories are required.

Figure 2: Sample images from some sequences in Hopkins 155 database. L: 2RT3RCT_B, from *checkerboard* sub-set. C: kanatani2, from *traffic* sub-set. R: three-cars, from *other* sub-set.

same condition, the PMAPE-LP method is slightly modified by adding an extra constraint to Eq.(13):

$$\sum_{j=1}^{N} x_j = K, \tag{18}$$

where K is the number of real motions.

In this experiment, we first transformed the trajectory data onto a 5 dimensional subspace (as suggested in [14]). Then we generated 18 motion models for each sequence. The penalizing factor α was set as 4 for all the sequences. Table 2 shows the average and median classification errors on the Hopkins 155 database given by different methods.

From the results we can see: (1) for both two-body and three-body motion segmentation, PMAPE-LP has the smallest overall misclassification rate compared to the state-of-the-art methods. (2) The segmentation results of PMAPE-LP remain consistent when going from one category to another, which reveals the robustness of PMAPE-LP for degenerate and partially dependent motions. (3) Though it happens rarely, PMAPE-LP sometimes gives large segmentation errors when the true motions are not fully contained in the candidate models. (4) In our experiment, the penalizing factor α was set to be constant all over the data sets. The misclassification rates can be further reduced by selecting different αs among the sequences.

4 Conclusions and Future Works

In this paper, we present an effective algorithm for multi-frame motion segmentation problem. It is obtained by an optimization problem based on a unified *mixture of subspace* model. This optimization problem is formulated as a maximum a posteriori estimation problem with model complexity penalty, which can be further converted into a linear programming problem with several candidate motion models. Therefore it can be effectively solved. Experimental results demonstrate that the proposed PMAPE-LP method can automatically estimate the number and dimensionality of motions, as well as achieves encouraging segmentation accuracy compared with the state-of-the-art methods.

Some highlights of the proposed method include:

1. Using *mixture of subspace* model to describe the multi-body motions makes the following unified formulation possible.

2. The proposed method uses penalized MAPE as the segmentation criterion. Since the prior probabilities of the candidate models are introduced, the MAP estimator is potentially more effective than the conventional maximum likelihood estimator. In addition,

	Methods		PMAPE-LP	LSA-5	LSA-4K	GPCA	MSL
2-body motions	Check.	Average	3.21	8.84	**2.57**	6.09	4.46
		Median	0.11	3.43	0.27	1.03	0.00
	Traffic	Average	**0.33**	2.15	5.43	1.41	2.23
		Median	0.00	1.00	1.48	0.00	0.00
	Other	Average	4.06	4.66	4.10	**2.88**	7.23
		Median	0.00	1.28	1.22	0.00	0.00
	All	Average	**2.20**	6.73	3.45	4.59	4.14
		Median	0.00	1.99	0.59	0.38	0.00
3-body motions	Check.	Average	8.34	30.37	**5.80**	31.95	10.38
		Median	5.35	31.98	1.77	32.93	4.61
	Traffic	Average	2.34	27.02	25.07	19.83	**1.80**
		Median	0.19	34.01	23.79	19.55	0.00
	Other	Average	8.51	23.11	7.25	16.85	**2.71**
		Median	8.51	23.11	7.25	16.85	2.71
	All	Average	**7.66**	29.28	9.73	28.66	8.23
		Median	5.60	31.63	2.33	28.26	1.76

Table 2: The misclassification rates (%) on Hopkins 155 database.

the number of motions can be automatically estimated using model complexity penalizing.

3. The assumption that the noises are not the same among candidate motion models is more reasonable than the constant noise assertion.

4. Linear programming can guarantee that the solution are the best combination from the candidate motion models. Also it can easily incorporate other prior knowledge, e.g. the number of motions.

Future research mainly lies in how to overcome the contradiction between good initial candidates and low computational cost of linear programming problem. In order to raise the possibility of true motions contained by the candidate list, we need to increase the number of candidates, which on the other hand leads to the larger scale of linear programming problem. Although linear programming is generally considered to be much efficient, when the scale becomes larger, the computational cost may be intolerable and the problem may even be impossible to be solved by ordinary computers.

References

[1] J. Costeira and T. Kanade. A multibody factorization method for independently moving objects. *IJCV*, 29(3):108–121, 1998.

[2] Z. Fan, J. Zhou, and Y. Wu. Multibody grouping by inference of multiple subspaces from high-dimensional data using oriented frames. *IEEE T-PAMI*, 28(1):91–105, 2006.

[3] C. W. Gear. Multibody grouping from motion images. *IJCV*, 29(2):133–150, 1998.

[4] G. H. Golub and C. F. Van-Loan. *Matrix Computation*. 2nd edn, The John Hopkins University Press, Baltimore, 1989.

[5] D. Hochbaum. *Approximation Algorithms for NP-Hard Problems*. PWS Publishing Company, Boston, MA, 1997.

[6] K. Kanatani. Geometric information criterion for model selection. *IJCV*, 26(3):171–189, 1998.

[7] K. Kanatani and Y. Sugaya. Multi-state optimization for multi-body motion segmentation. In *Australia-Japan Advanced Workshop on Computer Vision*, 2003.

[8] H. D. Li. Two-view motion segmentation from linear programming relaxation. In *Proc. CVPR*, 2007.

[9] A. Ng, Y. Weiss, and M. Jordan. On spectral clustering: Anaysis and an algoritm. In *NIPS*, 2001.

[10] K. Schindler and D. Suter. Two-view multibody structure-and-motion with outliers through model selection. *IEEE T-PAMI*, 28(6):983–995, 2006.

[11] N. Thakoor and J. Ga. Branch-and-bound hypothesis selection for two-view multiple structure and motion segmentation. In *Proc. CVPR*, 2008.

[12] P. Torr. Geometric motion segmentation and model selection. *Philosophical Transaction of Royal Society*, 356(1740):1321–1340, 1998.

[13] R. Tron and R. Vidal. A benchmark for the comparison of 3-d motion segmentation algorithm. In *Proc. CVPR*, 2007.

[14] R. Vidal, Y. Ma, and S. Sastry. Generalized principal component analysis (gpca). *IEEE T-PAMI*, 27(12):1–15, 2005.

[15] R. Vidal, R. Tron, and R. Hartlet. Multiframe motion segmentation with missing data using powerfactorization and gpca. *IJCV*, 79(1):85–105, 2008.

[16] H. Wang and D. Suter. Mdpe: A very robust estimator for model fitting and range image segmentation. *IJCV*, 59(2):139–166, 2004.

[17] J. Yan and M. Pollefeys. A general framework for motion segmentation: Independent, articulated, rigid, non-rigid, degenerate and non-degenerate. In *Proc. ECCV*, 2006.