Multi-frame Motion Segmentation via Penalized MAP Estimation and Linear Programming

Han Hu, Quanquan Gu, Lei Deng and Jie Zhou

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

Outline

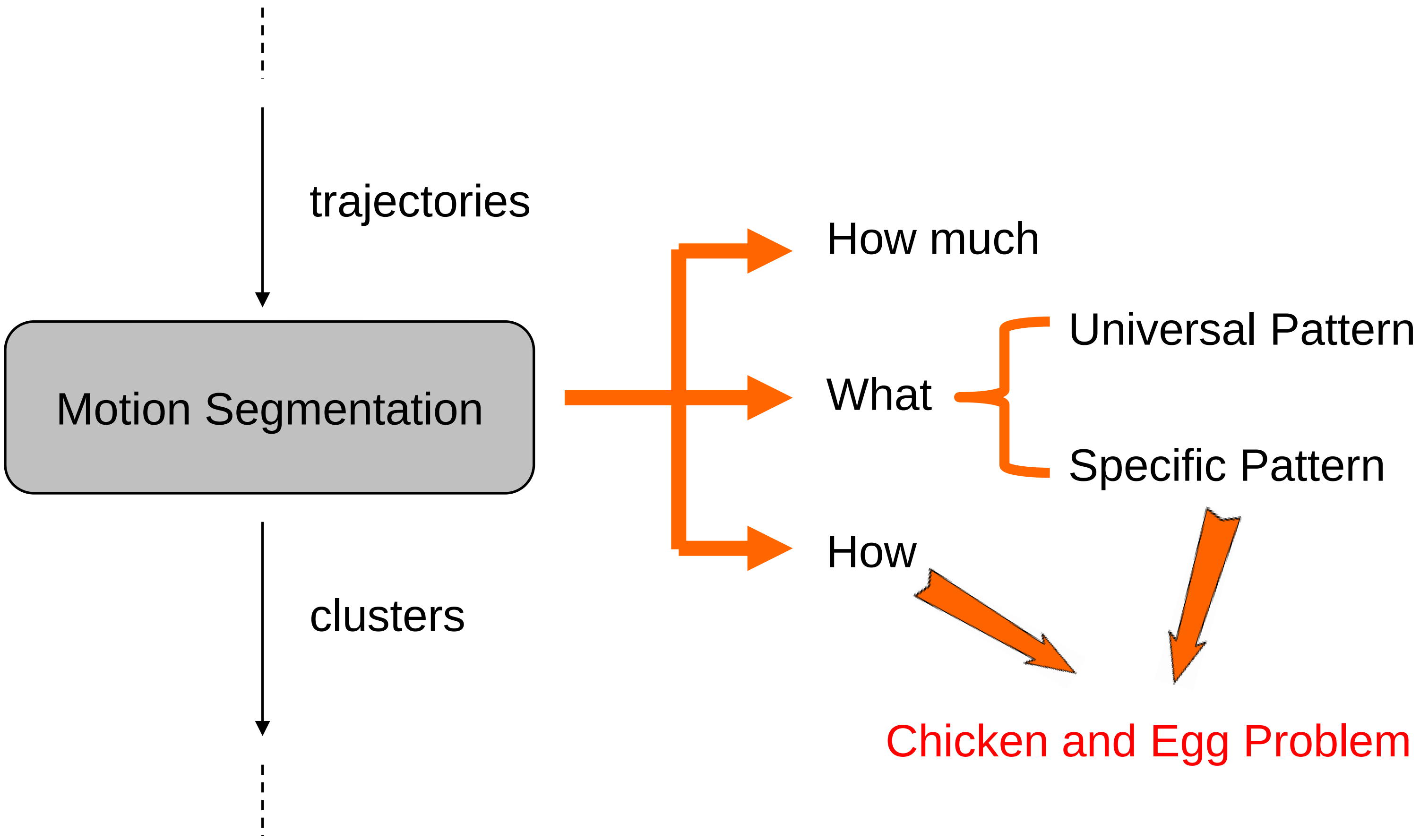
- Background
- Proposed Method
- Experiments
- Conclusions

Outline

- Background
- Proposed Method
- Experiments
- Conclusions

Applications

Problem State



Subspace Model

- A single rigid motion model

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

where, $A_i = K_i \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_i & \mathbf{t}_i \\ \mathbf{0}^T & 1 \end{bmatrix} \in \mathbb{R}^{2\times 4}$

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

Traditional Methods

- Factorization Based
  - Can only deal with full-dimensional and independent motions
  - Sensitive to noises
- Statistical
  - Local optimality
- Algebraic
  - GPCA (Vidal et al 2005)
  - Limited by the number of motions
  - Non-linear

Outline

- Background
- Proposed Method
- Experiments
- Conclusions

Motivation

- How much?
  - Most existing algorithms need the number of motions is known as a priori.
- What? And How?
  - Chicken and Egg problem
  - Iteratively optimize the objective function

→ Local optimality

To design a novel algorithm which can automatically determine the number of motions and reach global optimality.

Mixture of Subspace Model

- Single rigid motion model

  A set of orthonormal vectors $C = \{\mathbf{u_i}\}_{i=1}^{r}$

- Distance Metric

$$C_\perp = I - \sum_{i=1}^{r} \mathbf{u_i}\mathbf{u}_i^T \qquad d(\mathbf{w}, C) = ||C_\perp \mathbf{w}||,$$

- Mixture of Subspace

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

Penalized MAP Estimation

- MAP Estimation

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

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

- Penalizing

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

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

LP Relaxation

- Formulation
  - Suppose we have obtained a list of motion model candidates $\Phi\{C_1, \ldots, C_N\}$
  - Indicating variables $x_j = \max_{1 \le i \le P}\{L_{ij}\}$ $\Rightarrow$ $L_{ij} \le x_j$
  - Relaxation

$$L_{ij} \in \{0,1\}$$

$$\mathbf{x}_{ij} \in \{0,1\}$$

$\Rightarrow$ $0 \le L_{ij} \le 1,\ 0 \le x_{ij} \le 1, \forall i, j.$

$$\min_{L,x} \sum_{i=1}^{P} \sum_{j=1}^{N} L_{ij} \hat{d}^2(\mathbf{w}_i, C_j) + \alpha P\left(2\sum_{\{j:r_j=2\}} x_j + 3\sum_{\{j:r_j=3\}} x_j + 4\sum_{\{j:r_j=4\}} x_j\right)$$

$$\begin{aligned} s.t.\quad & \textstyle\sum_{j=1}^{N} L_{ij} = 1, \forall i; \\ & L_{ij} \le x_j, \forall i, j; \\ & 0 \le L_{ij} \le 1,\ 0 \le x_{ij} \le 1, \forall i, j, \end{aligned}$$

PMAPE-LP

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,

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

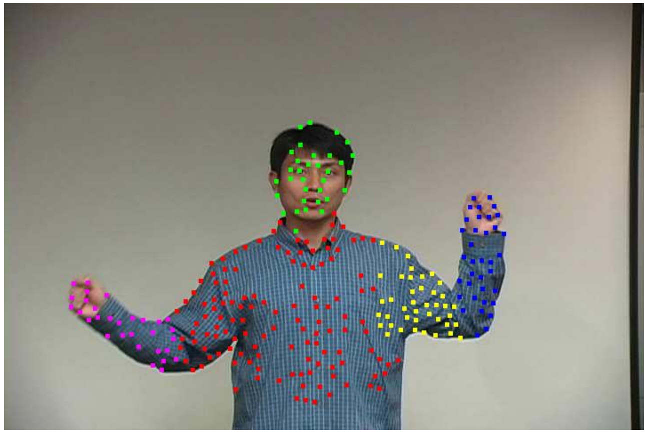
4. Construct linear programming problem, and solve it with any LP solver, e.g. Matlab's linprog in our experiment.
5. Round the LP solutions to binary variables, and output the groups.

Outline

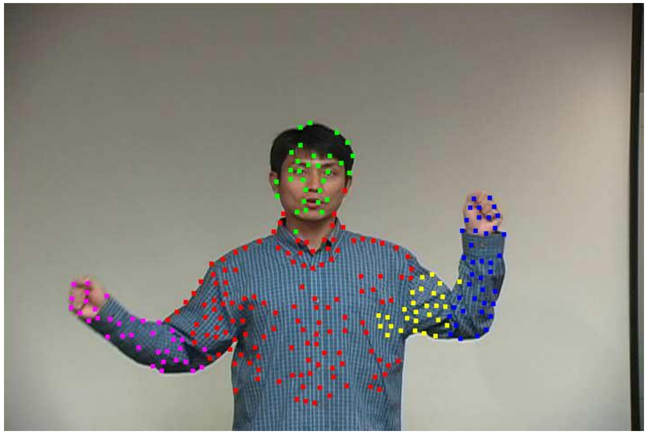
- Background
- Proposed Method
- Experiments
- Conclusions

Dancing Sequence

Ground truth



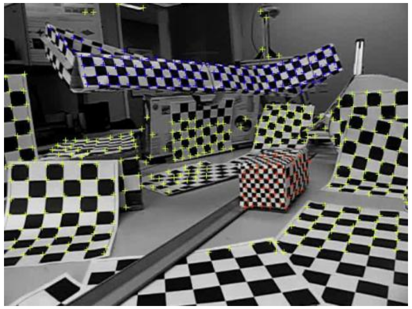
Segmentation Result



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.

Hopkins 155 Datasets



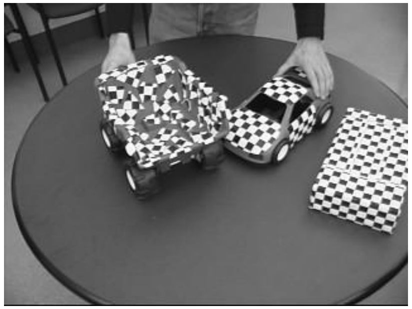


Table 1: Distribution of the number of points and frames.

	2 Groups			3 Groups		
	# Seq.	Points	Frames	# Seq.	Points	Frames
Check.	78	291	28	26	437	28
Traffic	31	241	30	7	332	31
Articul.	11	155	40	2	122	31
All	120	266	30	35	398	29
Point Distr.	35%-65%			20%-24%-56%		

Motion Number Constraint, $\sum_{j=1}^{N} x_j = K,$

Hopkins 155 Datasets

	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.

Outline

- Background
- Proposed Method
- Experiments
- Conclusions

Highlights

- Mixture of Subspace Model
- Penalized MAPE
  - MAP estimator is potentially more effective than the conventional maximum likelihood estimator.
  - The number of motions can be automatically estimated using model complexity penalizing.
- Non-constant Noise
  - More reasonable than the constant noise assertion.
- Linear Programming
  - Guarantee that the solutions are the best combination from the candidate motion models.
  - Easily incorporate other prior knowledge, e.g. the number of motions.

Thank you！