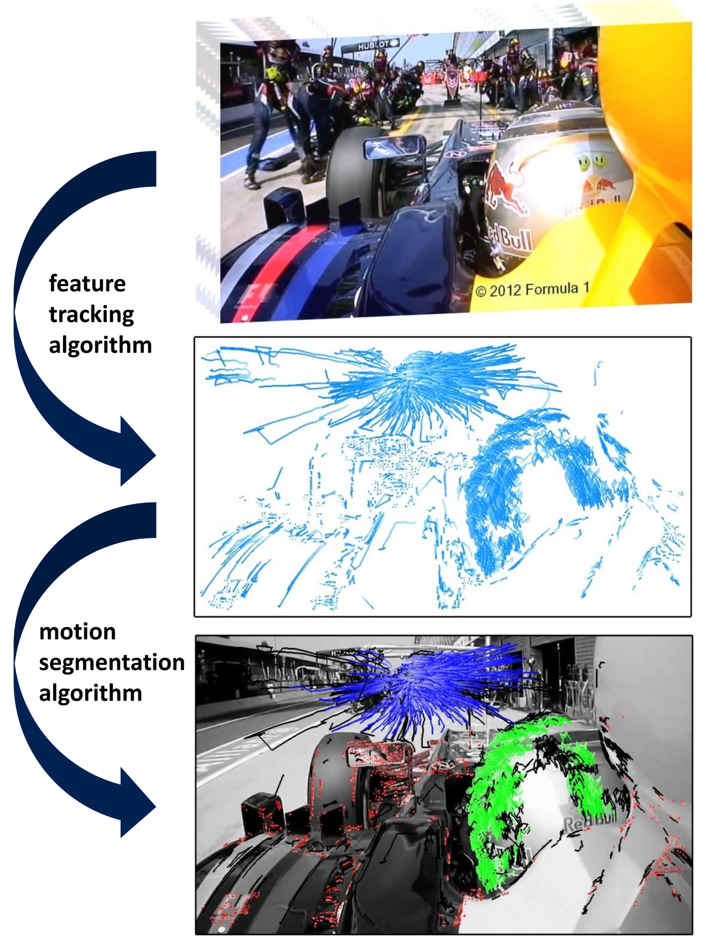
Fast Rigid Motion Segmentation via Incrementally-Complex Local Models

UNIVERSITY OF TORONTO

Fernando Flores-Mangas   Allan D. Jepson

{mangas, jepson}@ cs.toronto.edu

1. Motion Segmentation



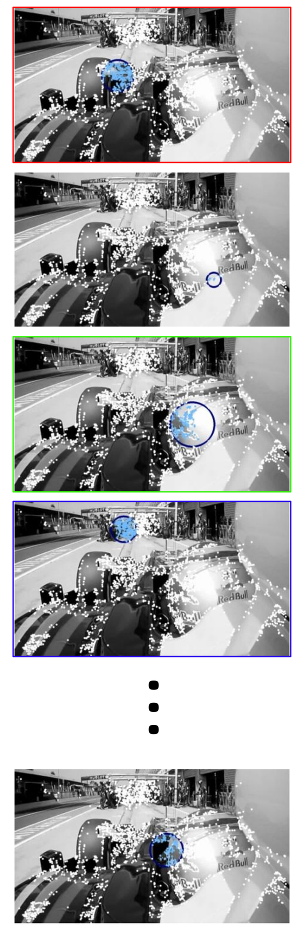
Goals: robust, accurate and efficient rigid motion segmentation of trajectory data.

Difficulties: structured noise, outliers, motion degeneracy, motion dependency.

Key ideas and contributions: local coherence for model estimation, penalized likelihood function for model selection, average of close-to-optimal segmentation results for improved accuracy.

2. Approach Overview

i) Instantiate spatially local motion models ($\mathbf{M}_i$).



→ $\mathbf{M}_1$

→ $\mathbf{M}_2$

→ $\mathbf{M}_3$

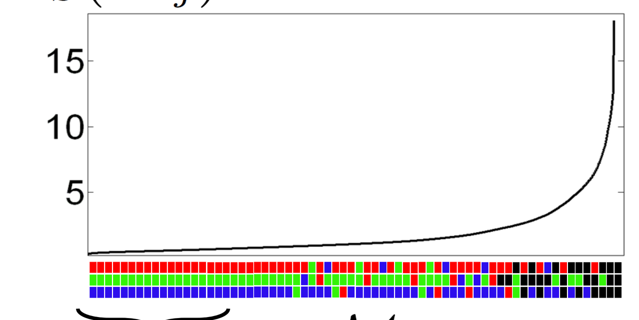
→ $\mathbf{M}_4$

⋮

→ $\mathbf{M}_M$

ii) Rank all possible model combinations using the penalized distance function $\mathcal{O}_S(\mathcal{M}_j)$

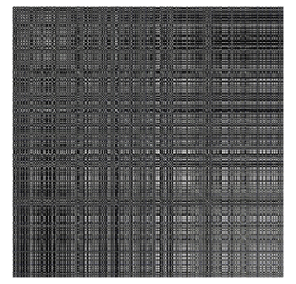
$\mathcal{O}_S(\mathcal{M}_j)$



$\mathcal{M}_j$

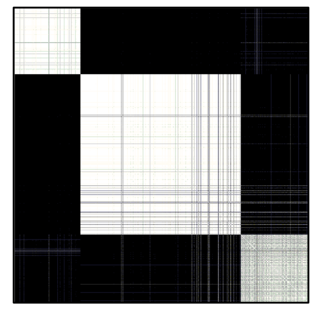
Labels of N best model combinations

iii) Incorporate the segmentation labels from the best N combinations of motions and cluster the resulting affinity matrix.



Affinities

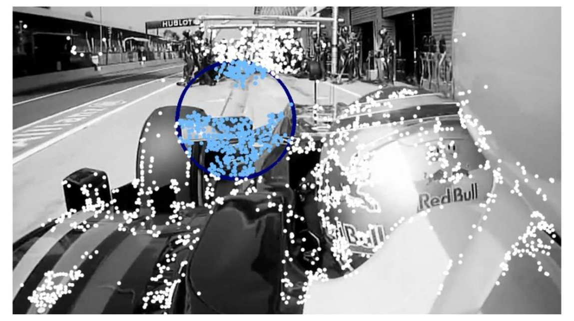
Spectral clustering



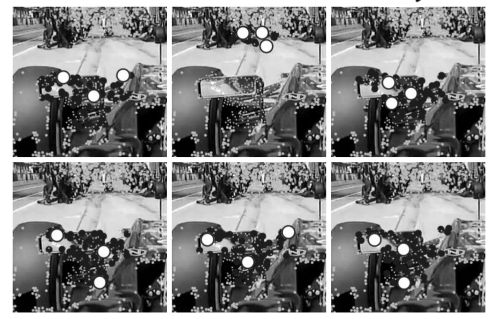
Final labels

3. Motion Model Instantiation

Given the trajectories $\hat{\mathbf{w}}$ within a support region



choose the i^th set of 3 control points $\mathbf{C}_i$



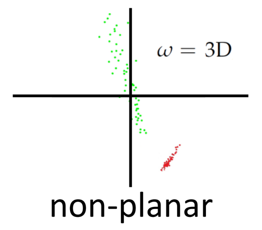
estimate a 2D affine model for each frame

$$\mathbf{A}^f = \begin{bmatrix} \mathbf{c}_1^f & \mathbf{c}_2^f & \mathbf{c}_3^f \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{c}_1^s & \mathbf{c}_2^s & \mathbf{c}_3^s \\ 1 & 1 & 1 \end{bmatrix}^{-1}$$

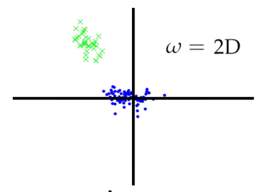
and evaluate the residuals

$$\begin{bmatrix} \mathbf{r}^f \\ \vec{0} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{w}}^f \\ \vec{1} \end{bmatrix} - \mathbf{A}^f \begin{bmatrix} \hat{\mathbf{w}}^s \\ \vec{1} \end{bmatrix}$$

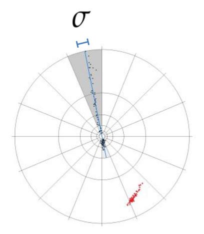
characterize



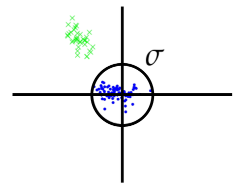
non-planar



planar

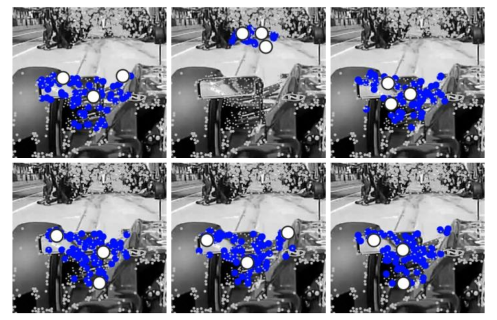


line fitting



gap finding

determine model inliers ($\hat{\mathfrak{b}}$)



resulting inliers

using the inlier trajectories

$$\mathbf{W}_{\hat{\mathfrak{b}}} = \{\mathbf{w}_i \in \mathbf{W} \mid b_i = 1\}$$

find the orthonormal basis $\mathbf{P}$

$$\mathbf{PQ} = \mathbf{P}\Sigma\mathbf{Q}' = \mathtt{svd}(\mathbf{W}_{\hat{\mathfrak{b}}})$$

the model parameters are

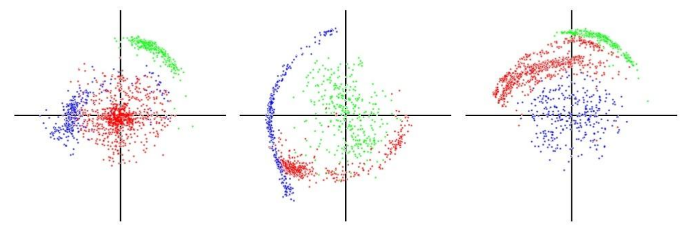
$$\mathbf{M} = \{\mathbf{P}, \hat{\mathfrak{b}}, \sigma, \omega\}$$

4. Model Subset Evaluation

accuracy

$$E(\mathbf{w}, \mathbf{M}) = |\mathbf{P}\mathbf{P}^\top \mathbf{w} - \mathbf{w}|^2$$

residuals for each model in $\mathcal{M}_j$



$$\mathcal{O}_S(\mathcal{M}_j) = \sum_{n=1}^{N} \sum_{p=1}^{P} \pi_{p,n} E(\mathbf{w}_p, \mathbf{M}_{j,n})$$

$$+ \lambda_\Phi \sum_{i=1}^{P} \Phi(\mathbf{w}_p, \mathbf{M}_{j,n}) + \lambda_\Psi \sum_{n=1}^{N} \Psi(\mathbf{M}_{j,n})$$

model overlap

$$\Phi(\mathbf{w}, \mathcal{M}_j) = -\log \frac{\max_{\mathbf{M}\in\mathcal{M}_j} \hat{E}(\mathbf{w}, \mathbf{M})}{\sum_{\mathbf{M}\in\mathcal{M}_j} \hat{E}(\mathbf{w}, \mathbf{M})}$$

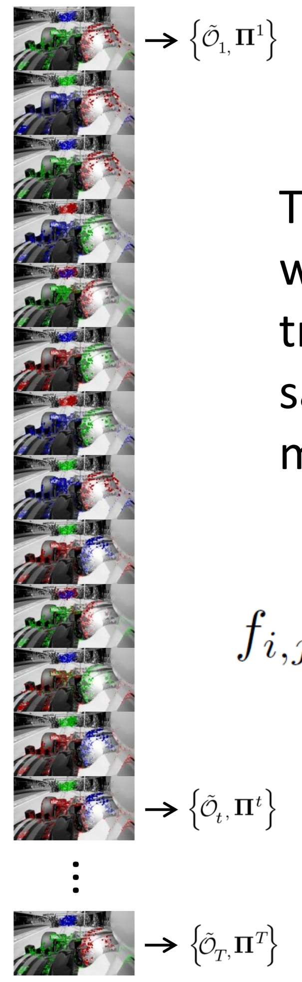
with $\hat{E}(\mathbf{w}, \mathbf{M}) = \exp\{-E(\mathbf{w}, \mathbf{M})\}$

model complexity

$$\Psi(\mathbf{M}) = \begin{cases} 6(F-1), & \text{if } \omega = 2D \\ 8(F-1) + (I-3), & \text{if } \omega = 3D \end{cases}$$

where I is the number of inliers produced by model $\mathbf{M}$

5. Segmentation affinities

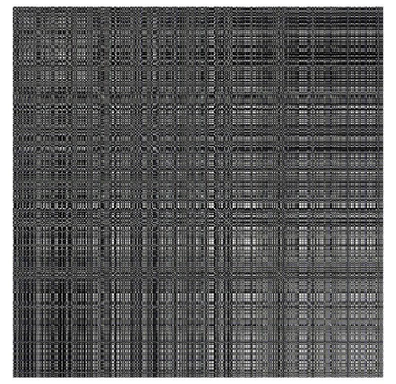


The affinity $f_{i,j}$ is the weighted frequency of trajectories i and j having the same label ($\boldsymbol{\pi}$) in the best T model combinations $\{\mathcal{M}_t\}_{t=1}^T$

$$f_{i,j} = \frac{1}{\sum_{t=1}^T \tilde{\mathcal{O}}_t} \sum_{t=1}^{T} \left(\boldsymbol{\pi}_{i,:}^t \boldsymbol{\pi}_{j,:}^{t\top}\right) \tilde{\mathcal{O}}_t$$

where the weights are $\tilde{\mathcal{O}}_t = \exp\left\{-\frac{\mathcal{O}_S(\mathbf{W}|\mathcal{M}_t^\star)}{\mathcal{O}_S(\mathbf{W}|\mathcal{M}^\star)}\right\}$

affinity matrix $\mathbf{F}$



spectral clustering

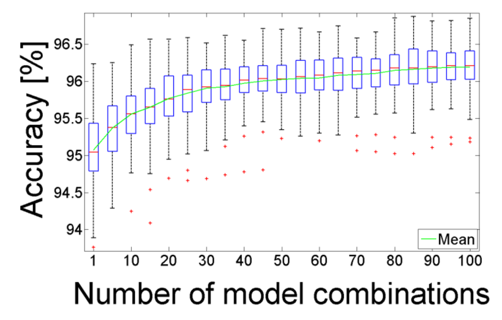
segmentation result

6. Evaluation and Results

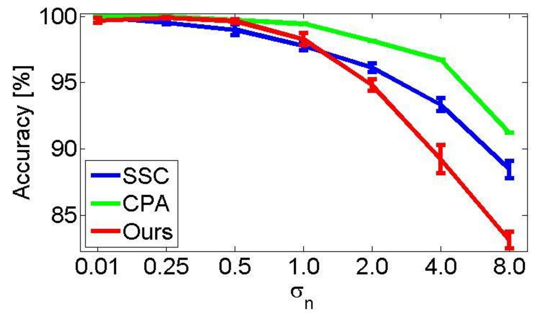
Comparison on Hopkins 155

Algorithm	Average Accuracy [%]	Computation time [s]
SSC [1]	98.76	14500
CPA [2]	98.75	147600
Ours	96.19	217
RANSAC	89.15	30

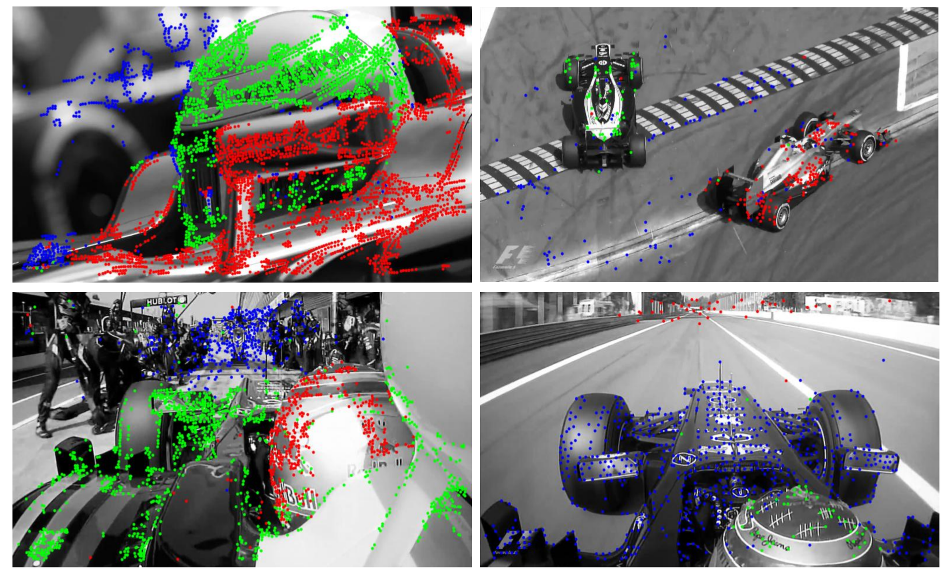
Accuracy vs. T



Synthetic noise on H155



Qualitative results with KLT tracks [3]



Benefits of the proposed method

- Computation time is 2 orders of magnitude faster than the state of the art.
- Comparable accuracy to state of the art in H155
- Useful for real-world trajectory data

[1] E. Elhamifar and R. Vidal. Sparse subspace clustering. In *Proc. CVPR*, 2009.

[2] L. Zappella, E. Provenzi, X. Lladó, and J. Salvi. Adaptive motion segmentation algorithm based on the principal angles configuration. *Proc. ACCV*, 2011.

[3] `http://www.ces.clemnson.edu/~stb/klt`