ACCV 2012

Robust Multi-Structure Fitting

Presented by:

Hanzi Wang

Xiamen University, China

Multi-structure fitting

- **Contamination in data is unavoidable:**
  - Sensor noise;
  - Faulty feature extraction;
  - Segmentation errors;
  - Multiple structures, etc.

What do we need to estimate for multi-structure data:

- The number of model instances
- The scale of inlier noise of each model instance.
- The parameters of each model instance.

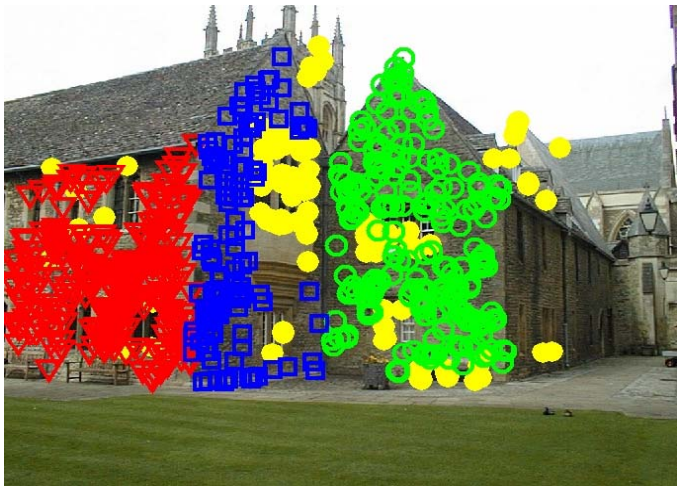


Figure 1. Multi-structure data

Review of Previous Works

- Traditional robust methods (such as LMedS, LTS, M-estimators) that cannot deal with more than 50% outliers.
- Some robust methods (such as RANSAC, MSAC, HT, ALKS, RESC, MSSE, MUSE, pbM, HBM, MINPRAN, ASSC, etc.) can deal with > 50% outliers.

To deal with multi-structure data, the above methods employ a sequential "**fitting-and-removing**" framework.



Review of Previous Works

- Problems in the sequential "**fitting-and-removing**" framework:
  - The **errors** (in model parameters or/and inlier scales) **can be propagated** to the following step;
  - It's **not computationally efficient**, it requires generating a large number of hypotheses in each step;
  - It requires a user to **specify the number of the structures**;
  - Some methods requires the user to **specify the inlier scale**, which may causes problems for multiple structures with different inlier scales.

One example

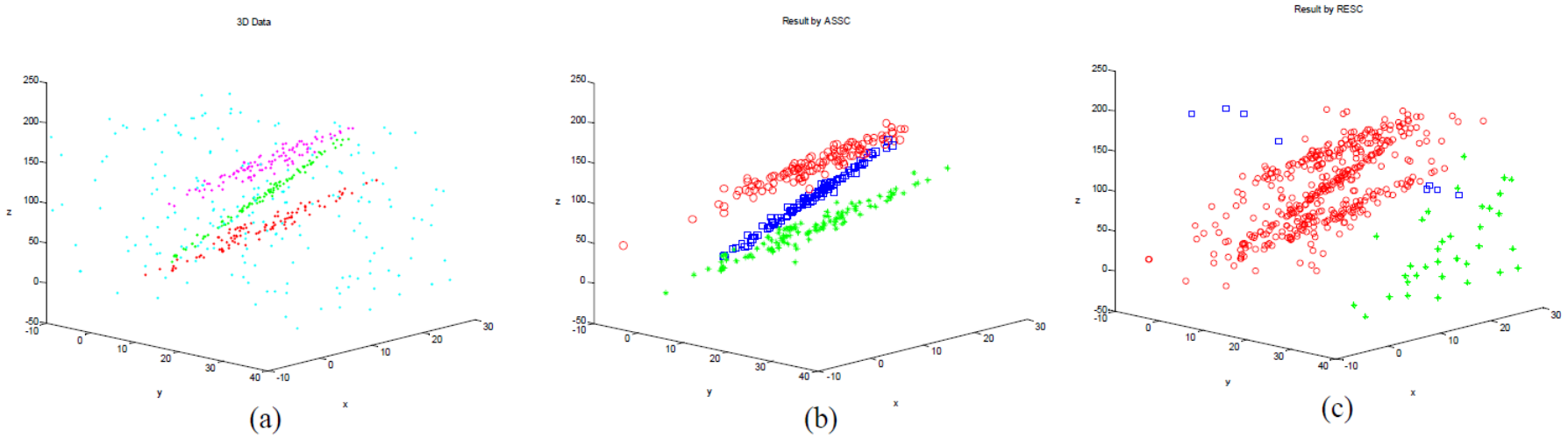


Figure 2. One example showing the segmentation results obtained by using the sequential "fitting-and-removing" framework(b). **Note (c), if the 1st structure is wrongly fitted, it will affect the fitting of the remaining structures.**

Review of Previous Works

Some methods use different strategies to fit multiple structures:

- multiRANSAC, Mean Shift (MS), HT/RHT, RHA, J-linkage, KF, AKSWH, etc.

However, most of these methods require some user-specified thresholds whose values are crucial in determining the number of model instances and affecting the performance of these methods.



Scale Estimation and Segmentation

Scale estimation plays an important role in model fitting and segmentation. It can greatly affect the performance of many robust estimators, because:

- It can be used to dichotomize inliers and outliers (such as RANSAC, multiRANSAC, etc.);
- It can be used to select the best hypothesis (such as ASSC, MDPE);
- It can be used to determine the bandwidth value or the bin size (e.g., PBM, HT/RHT).



Inliers/Outliers Dichotomy

- Given the scale of inlier noise *s*, the residual corresponding to the ith data sample is r_i, the inliers can be dichotomized from outliers by using the following equation:

$$|r_i/s| < \mathrm{E},$$

where E is a threshold (98 percent of inliers of a Gaussian distribution are included when E is set to 2.5).

Influence of the Scale Estimate

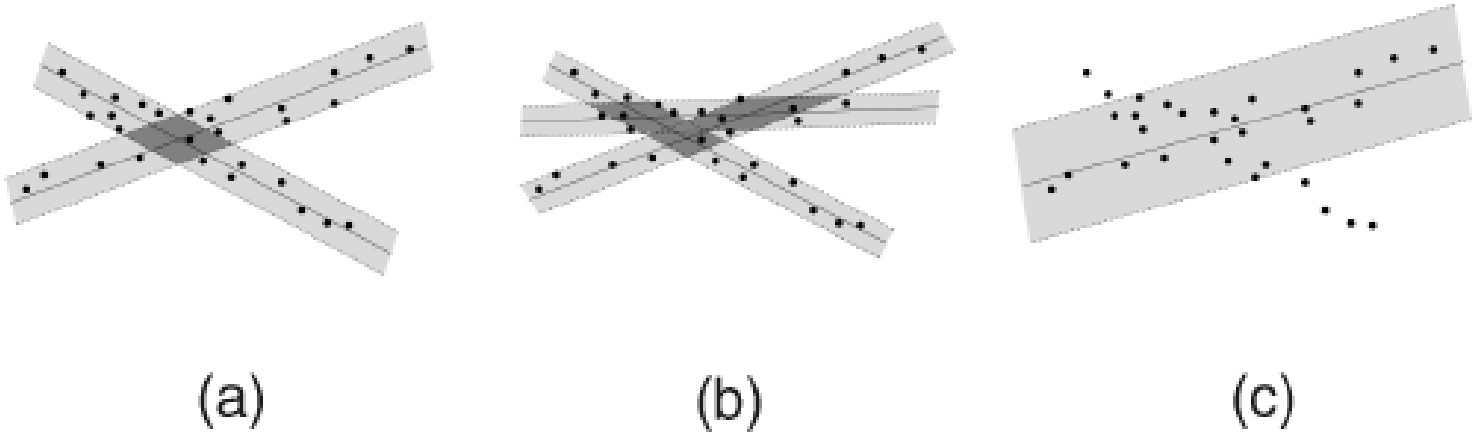


Figure 3. Influence of wrong thresholds on fitting multiple straight lines. (a) Correct threshold and fitted lines. (b) Thresholds that are too low encourage overfitting. (c) Thresholds that are too large encourage underfitting. (from Schindler and Suter, PAMI06)



Robust Scale Estimators

Median scale estimator

$$\hat{s}_{\mathrm{MED}} := 1.4826\left(1+\frac{5}{n-p}\right) med_i \left|r_i\right|$$

MAD scale estimator

$$\hat{s}_{\mathrm{MAD}} := 1.4826 med_i \left\{\left| r_i - med_i r_i \right|\right\}$$

Let $|\tilde{r}_i|$ be the sorted absolute residual. The KOSE scale estimator

$$\hat{s}_K := |\tilde{r}_{\mathrm{K}}| \Big/ \Theta^{-1}\left(\frac{1}{2}(1+\kappa)\right),$$

where $\kappa := \mathrm{K}/n$; $\Theta^{-1}(\bullet)$ is the argument of the normal cumulative density function.



Problems with MED/MAD/KOSE

When data contain multiple structures, MED, MAD and KOSE **can either break down or be badly biased.**

This is because :

- The breakdown is caused when the median (for MED/MAD), or the Kth ordered absolute residual (for KOSE), belongs to outliers;
- MED, MAD and KOSE are biased for multiple-structure data when *n* (the number of whole data) is used instead of the number of data belonging to the structure of interest.



The IKOSE scale estimator (Wang, et al., PAMI12):

$$\hat{s}_K^J := |\tilde{r}_K^J| \Big/ \Theta^{-1}\left(\frac{1}{2}\left(1+\kappa^J\right)\right)$$

$$\kappa^J := \mathrm{K}/n^J$$

It iteratively optimizes the estimate n^J, and **only use the data points belonging to the structure of interest**.

How to decide the K value?

- When the percentage of outliers in the data is not known, one should set K to be a small value to avoid breakdown;
- When the percentage of outliers is (approximately) known, one should set the K value as large as possible to achieve better statistical efficiency.



ALKS (Lee, et al., PAMI98)

- The Adaptive Least Kth order Squares (ALKS) algorithm optimizes the K value of KOSE by minimizing the variance of the normalized errors:

$$\zeta_{\mathrm{K}}^2 := \arg \min_K \frac{\hat{\sigma}_{\mathrm{K}}^2}{\hat{s}_{\mathrm{K}}^2}$$

$$= \arg \min_K \Theta^{-1}\left(\frac{1}{2}(1+\kappa)\right) \sum_{i=1}^{\mathrm{K}} r_i^2/(\mathrm{K}-p)|\tilde{r}_{\mathrm{K}}|,$$

where $\hat{\sigma}_K$ is the variance of the first K smallest absolute residuals. $\hat{S}_K$ is the estimated scale by KOSE.

MSSE (Bab-Hadiashar, et al, ROBOTICA99)

- The Modified Selective Statistical Estimator (MSSE) finds the K value which satisfies:

$$\frac{\hat{\sigma}^2_{K+1}}{\hat{\sigma}^2_K} > 1 - \frac{E^2 - 1}{K - p + 1}$$

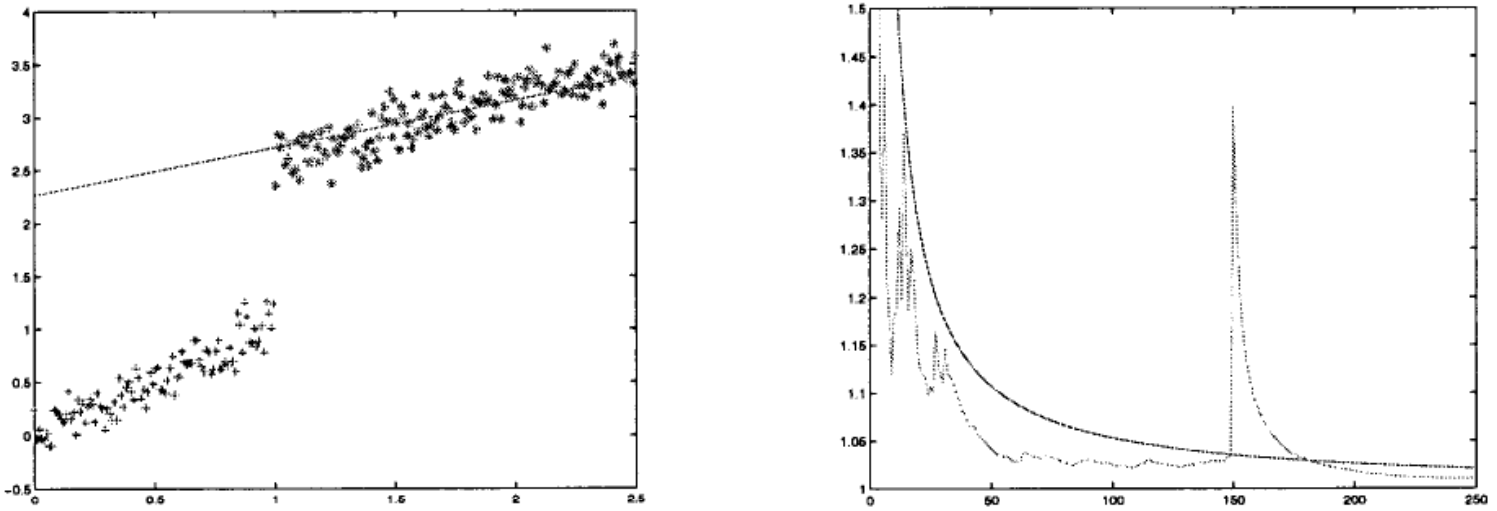


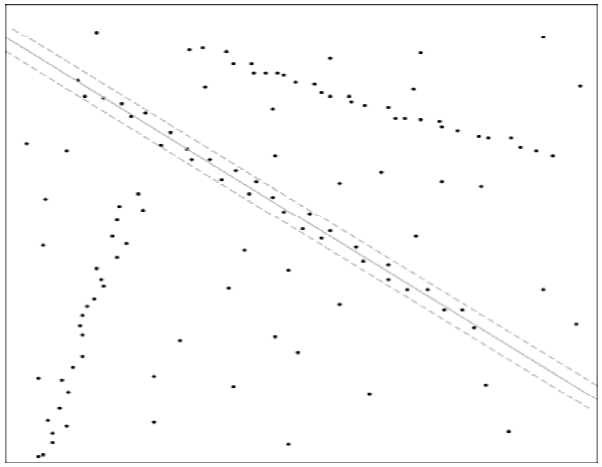
Fig. 4. Both sides of the inequality are plotted. The intersection of these two graphs provides the unbiased scale estimate. (from Bab-Hadiashar and Suter, Robotica99)

TSSE (Wang and Suter, PAMI04)

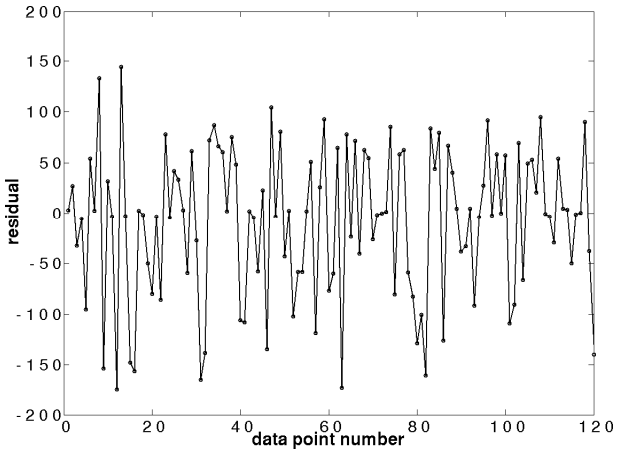
The Two-Step Scale Estimator (TSSE):

- Step 1: Using MS, with initial center zero, to find the local peak (P_k), and then using the MSV to find the valley next to the peak.

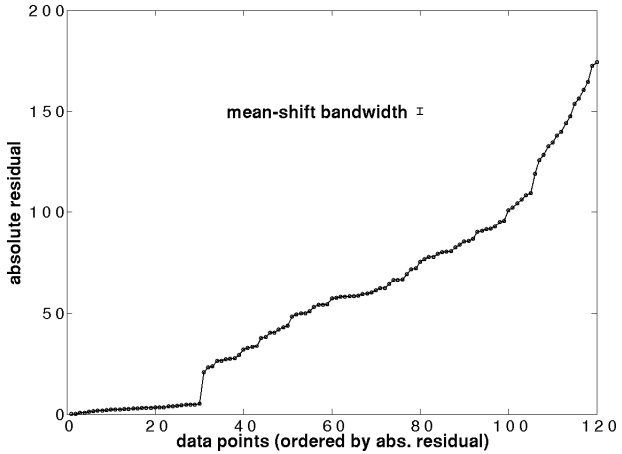
- Step 2: Estimating the scale of the inliers by the median scale estimator on the points within the obtained band centered at P_k.



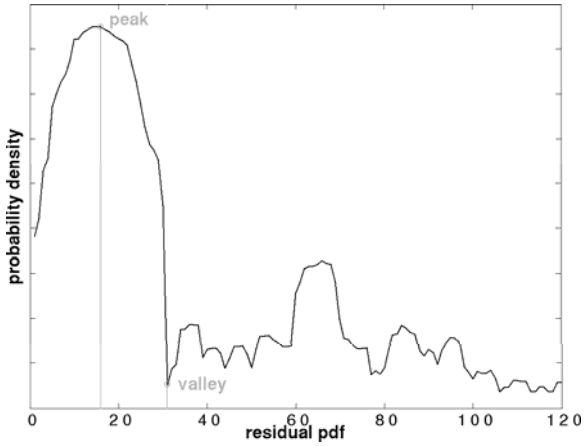
Original data distribution



Residuals



Sorted absolute residuals



Found peak and valley

Figure 5. Simultaneous scale estimation and outlier detection with the TSSE-estimator. (from Schindler and Suter, PAMI06)

Experiments

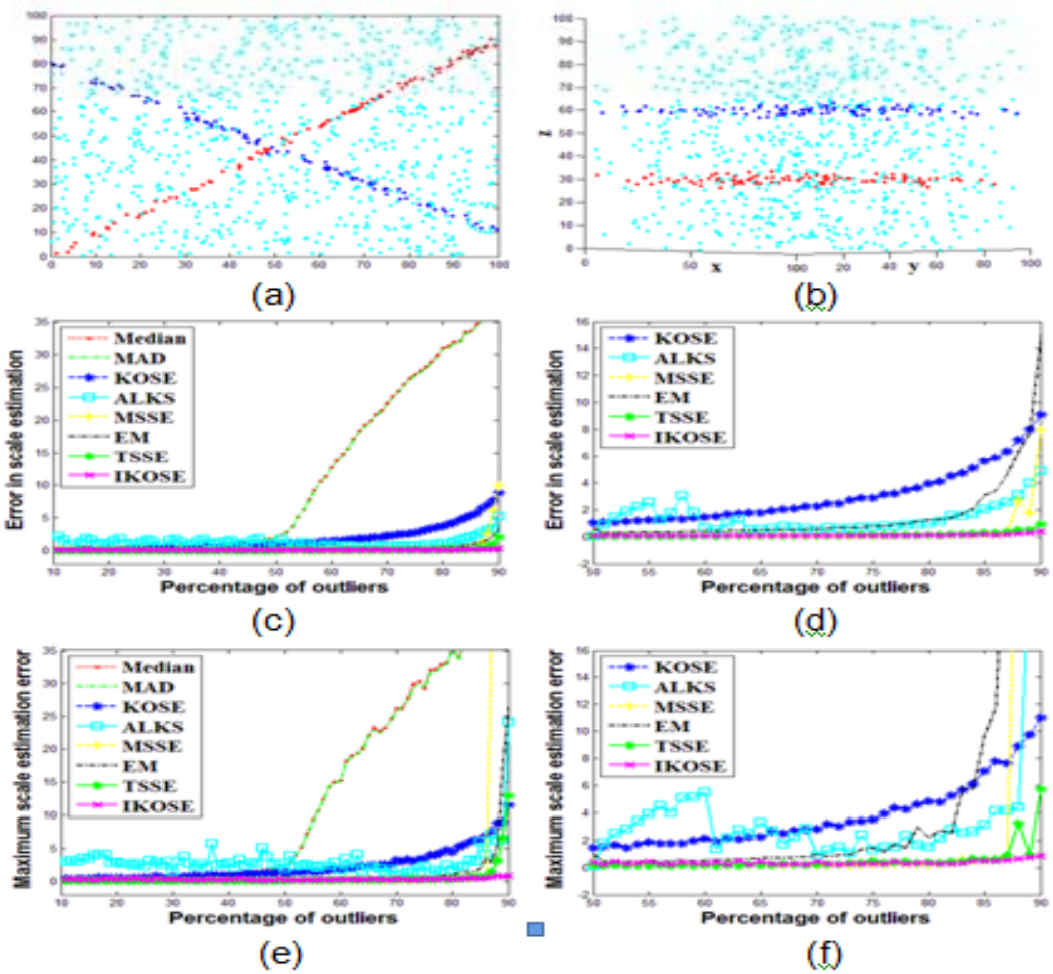


Figure 6. (a) and (b) Two snapshots respectively showing the “two-line” data and the “two-plane” data with 90% outliers. (c) and (d) are respectively the error plots of the scale estimation. (e) and (f) show the maximum errors in scale estimation.

Experiments

TABLE 1

QUANTITATIVE EVALUATION OF THE DIFFERENT SCALE ESTIMATORS ON THE TWO-LINE DATASET AND THE TWO-PLANE DATASET.

	Two-line dataset			Two-plane dataset		
	Mean	Std.Var.	Max. Err.	Mean	Std.Var.	Max. Err.
Median	11.1	12.9	39.8	35.7	2.32	40.8
MAD	11.0	12.8	39.9	27.3	1.33	36.7
KOSE	1.70	1.88	11.5	3.02	2.04	11.0
ALKS	1.27	0.71	24.1	1.40	1.02	23.2
MSSE	0.31	1.30	38.1	0.38	1.31	32.3
EM	0.33	0.52	26.6	1.57	2.70	87.4
TSSE	0.13	0.25	12.9	0.15	0.17	5.75
IKOSE	0.11	0.05	0.88	0.12	0.06	0.88

$$\left(\Lambda_s(\hat{s}, s_T) = \max\left(\frac{\hat{s}}{s_T}, \frac{s_T}{\hat{s}}\right) \right)$$

Multi-Structure Fitting (HT/RHT)

Hough Transform (HT) and Randomized Hough Transform (RHT, XU, et al, PRL90)

- ✓ Both use a random sampling scheme
- ✓ Both work in parameter space.
- ✓ Both assume that multiple model instances correspond to the multiple significant modes in parameter space.

The main differences are:

- ☐ HT discretizes the parameter space and computes the number of hypotheses in each bin, which is used to derive significant modes.
- ☐ RHT accumulates the number of the hypotheses whose distances are within a given tolerance, and detects significant modes based on the accumulated number.



Multi-Structure Fitting (HT/RHT)

The advantages of RHT over HT are:

- ✓ RHT can observe infinite parameter space while HT can only observe finite parameter space.
- ✓ RHT requires less storage space.
- ✓ It has higher computational speed.

However,

- ☐ RHT needs more user-specified thresholds than HT.
- ☐ It is not trivial to find a global optimal bin size for both HT and RHT.
- ☐ It is hard to both achieve accurate results and correctly localize multiple significant peaks in parameter space.

Multi-Structure Fitting (MS)

Mean-Shift (MS) based methods

- The MS-based methods (e.g., Tuzel, et al. CVPR05, Subbarao and Meer, IJCV09) also assume that **dominant modes represent multiple structures** and **the positions of the modes correspond to the parameters of the multiple structures .**

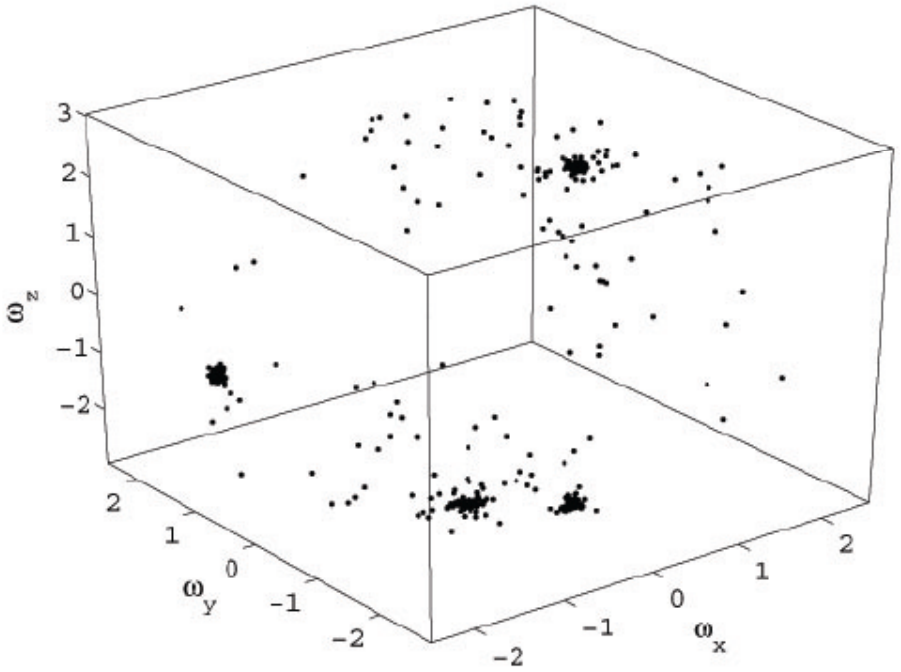


Figure 7. One example showing that each of the significant modes means one motion group. (from Tuzel, et al. CVPR05)

Some good results have been obtained by MS

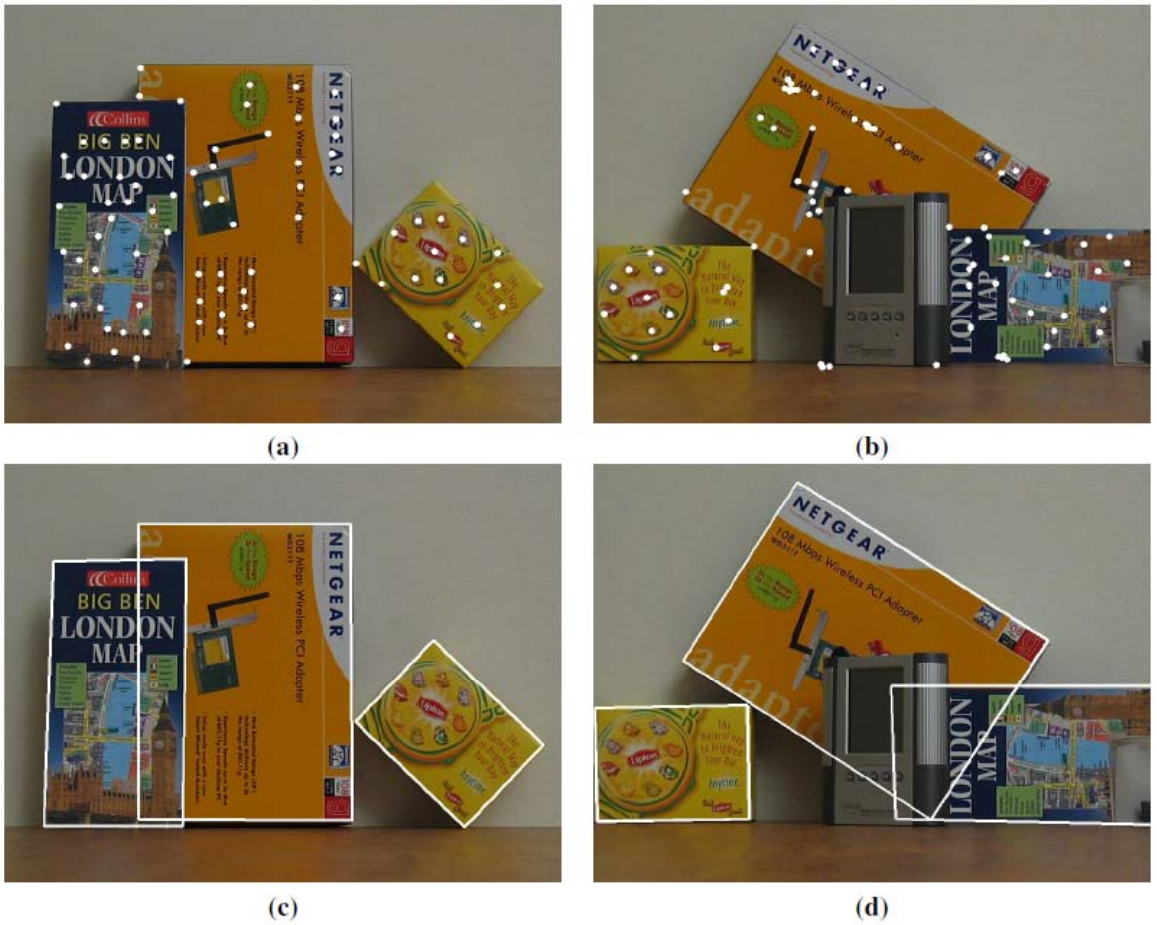


Figure 8. 2D image data. (a, b) Original and transformed scenes. (c) The boundaries of the bodies. (d) Transformed boundaries with the estimated motion parameters. (from Tuzel, et al. CVPR05)

	Mot. hyp.	kde
$\mathbf{M}_1$	581	0.469
$\mathbf{M}_2$	51	0.040
$\mathbf{M}_3$	54	0.036
M_4	8	0.016

Figure 9. Multibody Factorization. The left figure shows the 1st frame with all the points which are tracked. The right image shows the 5th (i.e., last) frame with only the inliers. **The table contains the properties of the first four modes.** (from Subbarao and Meer, IJCV09)

Multi-Structure Fitting (MS)

- However, **how to determine significant modes is not an easy task.**
- One way to select significant modes is to check **if the first N modes clearly dominate the (N+1)th mode**.
- But it is hard that how to judge when the Nth mode dominates the (N+1)th mode.
- We also found that the task becomes much more difficult for MS when the inlier noise scale is high or the percentage of outliers is high.



Multi-Structure Fitting (J-linkage)

- The J-linkage method (Toldo and Fusiello, ECCV08) generates M model hypothesis by random sampling.
- The Preference Set (PS) of each model hypothesis is computed, as in RANSAC.
- Then it builds a N ×M matrix where entry (i, j) is 1 if the distance of a point i to a model j is less than a threshold; otherwise the entry (i, j) is 0.
- Each column of the matrix is the PS of a model hypothesis. Each row indicates which model a point prefers.

Multi-Structure Fitting (J-linkage)

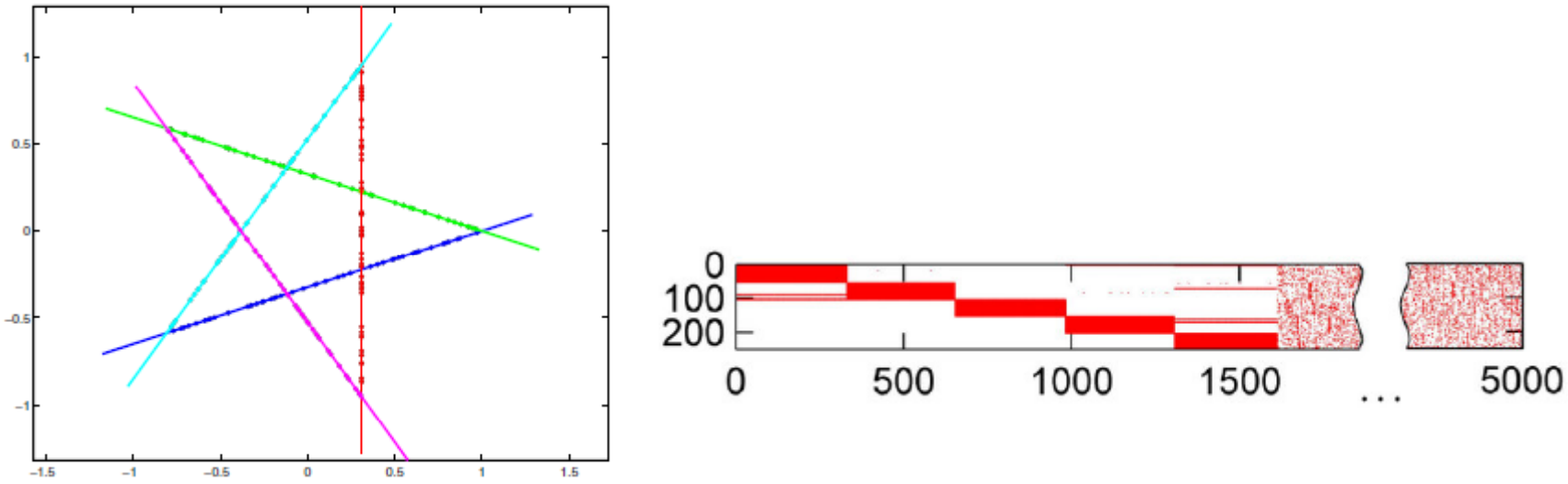


Figure 10. Left: the data consist of 250 points on 5 lines. Right: The NxM matrix. The rows are points (ordered by cluster), the columns are models (ordered by cluster size). (from Toldo and Fusiello, ECCV08)

Multi-Structure Fitting (J-linkage)

- Given two preference sets A and B, which correspond to two model hypotheses , the Jaccard distance between the two sets is:

$$d_{\mathrm{J}}(A,B) = \frac{|A \cup B| - |A \cap B|}{|A \cup B|}.$$

- The Jaccard distance measures the degree of overlap of the two sets and ranges from 0 (identical sets) to 1 (disjoint sets).

Multi-Structure Fitting (J-linkage)

- The cut-off value is set to 1. It means that the algorithm will only link together elements whose preference sets overlap.
- Each cluster of points defines (at least) one model.
- The final model for each cluster of points is estimated by least squares fitting.

Multi-Structure Fitting (J-linkage)

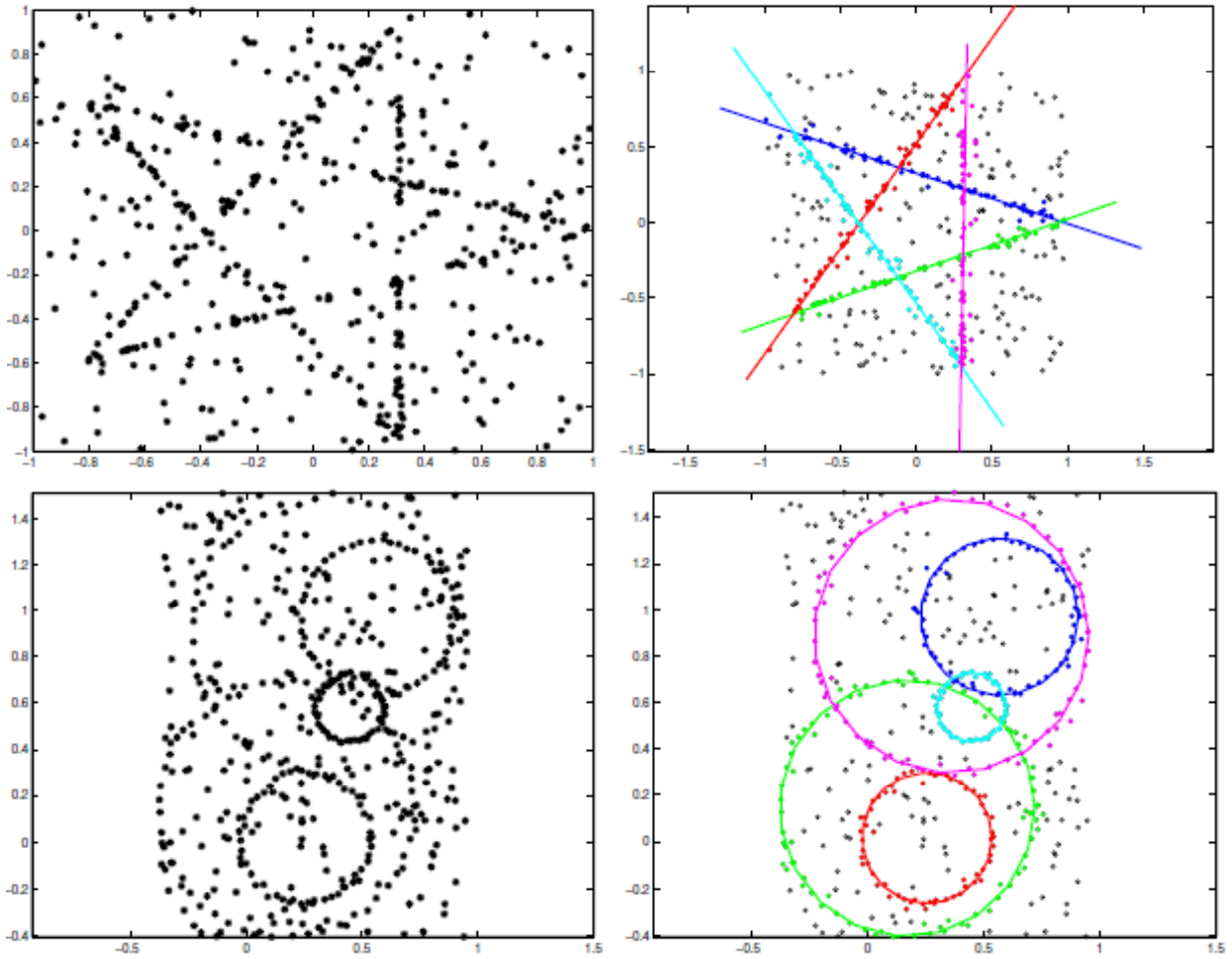


Figure 11. Some results obtained by J-linkage (from Toldo and Fusiello, ECCV08)



Multi-Structure Fitting (J-linkage)

Advantages:

- ✓ It can fit multiple structures simultaneously.
- ✓ It can estimate both the number of model instances and the parameters of the model instances.

Disadvantages:

- ☐ It dichotomizes inliers/outliers **by using a user-specified inlier scale**.
- ☐ Like RANSAC, **the performance of J-linkage greatly depends on the specified inlier scale**.
- ☐ The estimated number of model instances is heavily affected by **a threshold which is used in selecting significant bins of the hypothesis histogram**.

Multi-Structure Fitting (KF)

- For each data point x_i, KF (Chin, et al, ICCV09) computes its absolute residual set $\mathbf{r}_i = \{r_1^i, \ldots, r_M^i\}$ as measured to the M hypotheses.
- It sorts the absolute residual set to obtain the sorted residual set $\tilde{\mathbf{r}}_i = \{r_{\lambda_1^i}^i, \ldots, r_{\lambda_M^i}^i\}$, where the permutation $\{\lambda_1^i, \ldots, \lambda_M^i\}$ is obtained so that

$$r_{\lambda_1^i}^i \leq \cdots \leq r_{\lambda_M^i}^i$$

- Define the sorted hypothesis set of point x_i as

$$\tilde{\boldsymbol{\theta}}_i := \{\lambda_1^i, \ldots, \lambda_M^i\},$$

Multi-Structure Fitting (KF)

- The Ordered Residual Kernel (ORK) between two data points can be defined as:

$$k_{\tilde{r}}(x_{i_1}, x_{i_2}) := \frac{1}{Z} \sum_{t=1}^{M/h} z_t \cdot k_{\cap}^{t}(\tilde{\boldsymbol{\theta}}_{i_1}, \tilde{\boldsymbol{\theta}}_{i_2}),$$

where $z_t = 1/t$ are the harmonic series and is Z the (M/h)-th harmonic number.



Multi-Structure Fitting (KF)

- Step size *h* is used to obtain the Difference of Intersection Kernel (DOIK)

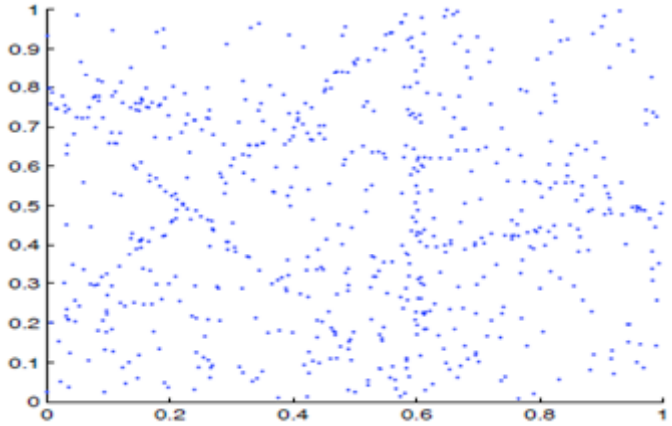
$$k_{\cap}^{t}(\tilde{\theta}_{i_1}, \tilde{\theta}_{i_2}) := \frac{1}{h}(|\tilde{\theta}_{i_1}^{1:\alpha_t} \cap \tilde{\theta}_{i_2}^{1:\alpha_t}| - |\tilde{\theta}_{i_1}^{1:\alpha_{t-1}} \cap \tilde{\theta}_{i_2}^{1:\alpha_{t-1}}|)$$

where $\alpha_t = th$ and $\alpha_{t-1} = (t\text{-}1)h$.



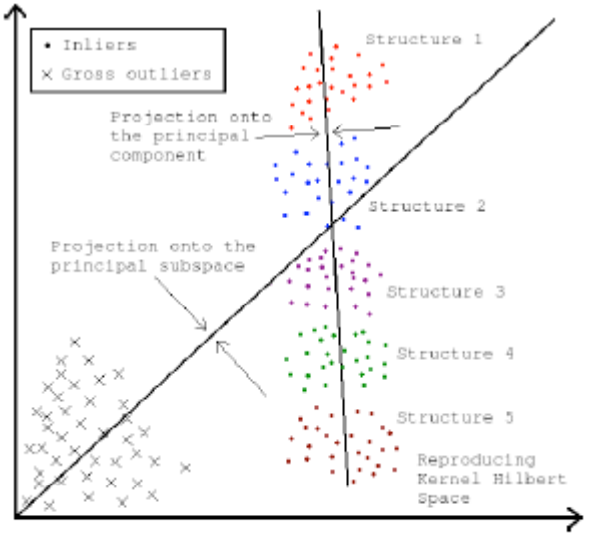
Multi-Structure Fitting (KF)

- A kernel matrix $\mathbf{K}$ is constructed by mapping the input data to a reproducing Kernel Hilbert Space (RKHS), which can be computed by using the kernel function $k_{\tilde{T}}$.
- Through the eigenvalue decomposition (EVD) and Singular Value Decomposition (SVD), the data can be projected onto the principal subspace.
- Directions of the principal subspace are dominated by inlier points.

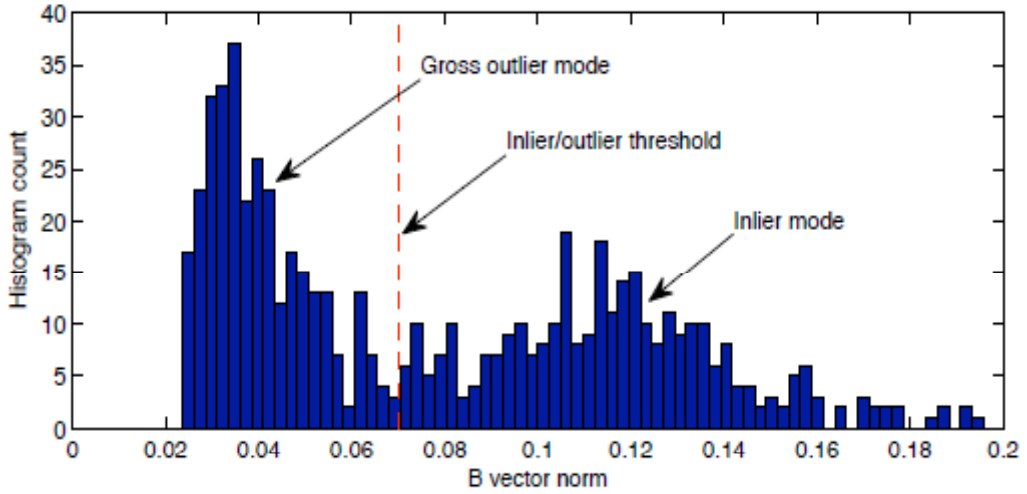


Input data: 92% outliers, inlier $\sigma = 0.01$.

(a) Original data



(b) Input data mapped to the RKHS $F_{k_{\tilde{r}}}$



(c) Histogram of vector norms in the kernel principal subspace

Figure 12. (b) Gross outlier removal with Kernel SVD and structure discovery with Kernel PCA. (c) The histogram is obtained from the actual input data.

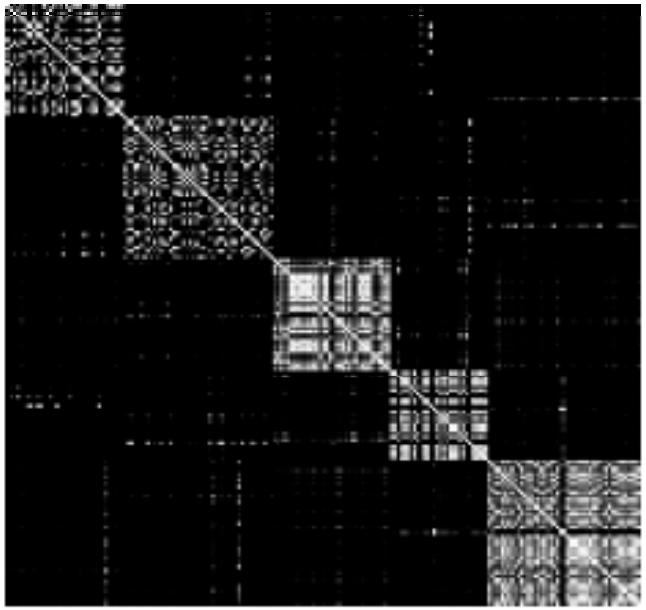
Multi-Structure Fitting (KF)

- The cutoff threshold can be set by employing the Gaussian Mixture Model (GMM) with two components to find the threshold.
- The outlier removal scheme is more tractable than the mode seeking-based method (for the unknown number of structures), as it is known beforehand that there are at most two modes in the norm distribution.

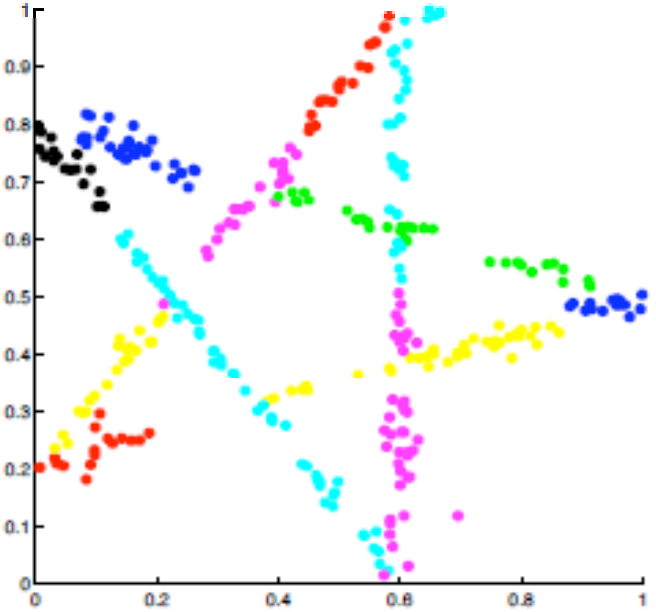
Multi-Structure Fitting (KF)

- To fit multiple model instances, it is based on the idea that points from the same structure concentrate at a location in RKHS.
- The Normalized Cut method is used to cluster the data.

Multi-Structure Fitting (KF)



(a) Weighted adjacency matrix with points re-arranged based on true cluster membership.



(b) Ncut reveals 12 clusters. The figure is best viewed in colour. Note that the colours repeat.

Figure 13. (a) Weighted adjacency matrix for the data (i.e., the five lines with 92% outliers). (b) Normalized Cut clustering results on for the data.

Multi-Structure Fitting (KF)

To deal with the "overclustering" problem

- A structure pruning scheme is operated to fit the data with the least number of structures.
  - A model instance is first estimated from each point cluster with LMedS.
  - The algorithm then sequentially removes structures by testing if a structure is removed, whether the remaining structures can "explain" the data.
  - This proceeds until the condition of satisfactory explanation is violated.

Experimental Results (KF)

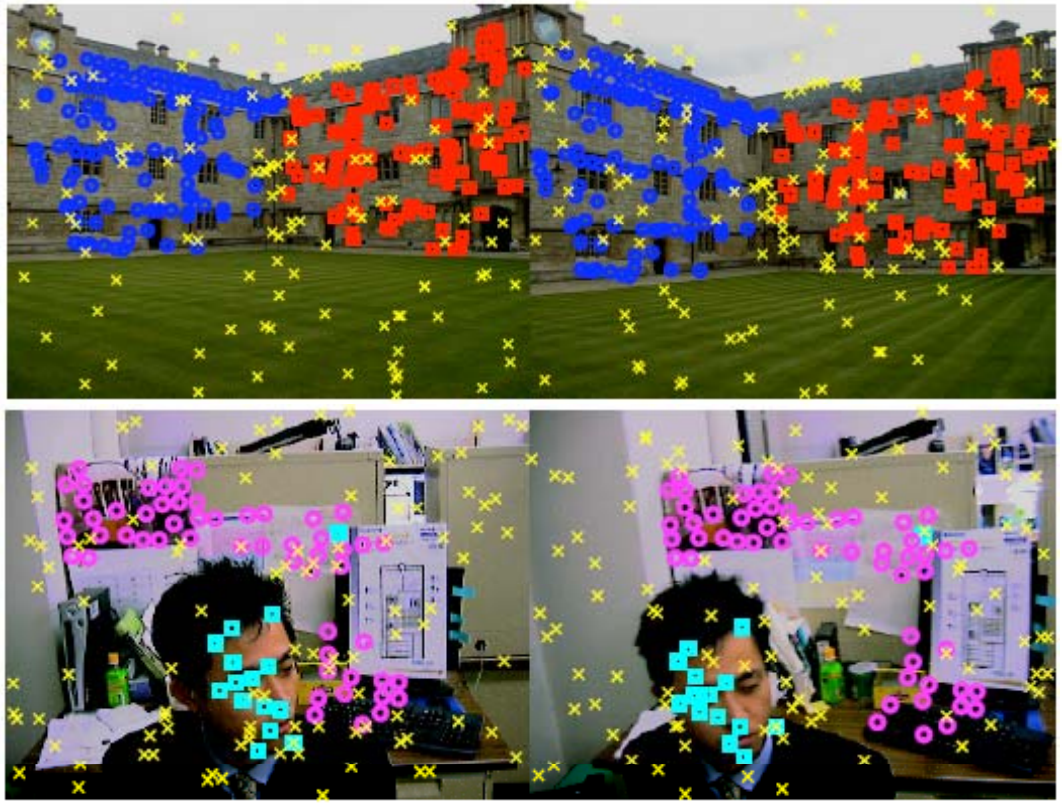


Figure 15. 1st row: Homography estimation results. 2nd row: Motion segmentation results . Yellow crosses indicate gross outliers determined by KF.



Multi-Structure Fitting (KF)

Advantages:

- ✓ The KF method can effectively remove gross outliers in the data and in parallel discover the multiple structures.
- ✓ It does not require a user to specify the inlier scale.

Disadvantages:

- ☐ It is computationally slow.
- ☐ The value of the step size h and the weighting ratio used in KF has a significant influence on determining the number of model instances.

Multi-Structure Fitting (AKSWH)

- AKSWH (Wang, et al, PAMI12) generates a set of *p*-subsets, and compute the model hypotheses using the *p*-subsets.
- It assigns each hypothesis a weight.

$$\mathcal{P} := \{\mathcal{P}_i\}_{i=1,2,\ldots} = \{(\hat{\boldsymbol{\theta}}_i, \hat{w}_i)\}_{i=1,2,\ldots}$$

- If we know some weighted hypotheses are associated with the *J*th structure, we can estimate the parameters of the *J*th structure by using:

$$\hat{\mathcal{P}} = \left\{ \hat{\mathcal{P}}^J \,\middle|\, \hat{\mathcal{P}}^J = (\hat{\boldsymbol{\theta}}^J, \hat{w}^J) = \arg\max_{w_i} \{\mathcal{P}_i^J \mid \mathcal{P}^J\} \right\}$$



Multi-Structure Fitting (AKSWH)

Weighting function:

It uses a variable-scale weight function which can be written as:

$$\hat{w}_j := \frac{\hat{f}_{\mathrm{KN},\hat{\theta}_j}(\mathbf{O})}{\hat{s}_K(\hat{\boldsymbol{\theta}}_j)} = \frac{1}{n}\sum_{i=1}^{n}\frac{\mathbf{KN}\left(r_i(\hat{\boldsymbol{\theta}}_j)/h(\hat{\boldsymbol{\theta}}_j)\right)}{\hat{s}_K(\hat{\boldsymbol{\theta}}_j)h(\hat{\boldsymbol{\theta}}_j)}$$

$$\text{where}\quad \mathbf{KN}_E(r) := \begin{cases} \frac{3}{4}(1-\|r\|^2) & \|r\| \le 1 \\ 0 & \|r\| > 1 \end{cases}$$

The fixed-scale weight is written as:

$$\hat{w}_j \propto \frac{1}{n}\sum_{i=1}^{n}\mathbf{KN}\left(r_i(\hat{\boldsymbol{\theta}}_j)/h\right)$$

Multi-Structure Fitting (AKSWH)

Selecting significant hypotheses: it employs the entropy thresholding approach, which can adaptively determines the threshold value.

Given a set of hypotheses with weights $\mathcal{W} := \{\hat{w}_i\}_{i=1,\ldots,n}$, we define:

$$\Psi_j := \left(\arg\max_{\hat{w}_i} \mathcal{W} \right)^2 - \hat{w}_j^2$$

The prior probability of component Ψ_j is:

$$p(\Psi_i) := \Psi_i \Big/ \sum_{j=1}^{n} \Psi_j$$

The significant hypotheses can be selected which satisfy the following condition:

$$\mathcal{P}^* = \{\mathcal{P}_i \mid \log p(\Psi_i) - \sum_{i=1}^{n} p(\Psi_i) \log p(\Psi_i) < 0\}$$



Multi-Structure Fitting (AKSWH)

Given a set of residuals $\{r_i(\hat{\boldsymbol{\theta}}_j)\}_{i=1,\ldots,n}$ for a model hypothesis $\hat{\boldsymbol{\theta}}_j$, it formulates the consensus set of residuals as:

$$\mathcal{C}(\hat{\boldsymbol{\theta}}_j) := \{\mathrm{L}(r_i(\hat{\boldsymbol{\theta}}_j))\}_{i=1,\ldots,n}, \text{ where } \mathrm{L}(r_i) = \begin{cases} 1 & \text{If } |r_i| \le \mathrm{E}\hat{s} \\ 0 & \text{Otherwise} \end{cases}$$

The J-distance between two consensus sets (corresponding to two hypotheses) is given by:

$$\mathrm{J}\left(\mathcal{C}(\hat{\boldsymbol{\theta}}_j), \mathcal{C}(\hat{\boldsymbol{\theta}}_k)\right) := 1 - \frac{\mathcal{C}(\hat{\boldsymbol{\theta}}_j) \cap \mathcal{C}(\hat{\boldsymbol{\theta}}_k)}{\mathcal{C}(\hat{\boldsymbol{\theta}}_j) \cup \mathcal{C}(\hat{\boldsymbol{\theta}}_k)}$$



Multi-Structure Fitting (AKSWH)

To solve the over-clustering problem, **it uses Mutual Information Theory (MIT) to fuse clusters** belonging to the same model instance.

- The mutual information between two hypotheses can be written as:

$$\mathbb{M}(\mathcal{P}_i^{\dagger}, \mathcal{P}_j^{\dagger}) \equiv \mathbb{M}(\boldsymbol{\theta}_i^{\dagger}, \boldsymbol{\theta}_j^{\dagger}) := \mathrm{p}(\boldsymbol{\theta}_i^{\dagger}, \boldsymbol{\theta}_j^{\dagger}) \log \frac{\mathrm{p}(\boldsymbol{\theta}_i^{\dagger}, \boldsymbol{\theta}_j^{\dagger})}{\mathrm{p}(\boldsymbol{\theta}_i^{\dagger})\mathrm{p}(\boldsymbol{\theta}_j^{\dagger})}$$

$$\text{where} \quad \frac{\mathrm{p}(\boldsymbol{\theta}_i^{\dagger}, \boldsymbol{\theta}_j^{\dagger})}{\mathrm{p}(\boldsymbol{\theta}_i^{\dagger})\mathrm{p}(\boldsymbol{\theta}_j^{\dagger})} = \frac{n\sum_{l=1}^{n} \mathrm{p}(\boldsymbol{x}_l \mid \boldsymbol{\theta}_i^{\dagger})\mathrm{p}(\boldsymbol{x}_l \mid \boldsymbol{\theta}_j^{\dagger})}{\sum_{l=1}^{n} \mathrm{p}(\boldsymbol{x}_l \mid \boldsymbol{\theta}_i^{\dagger}) \sum_{l=1}^{n} \mathrm{p}(\boldsymbol{x}_l \mid \boldsymbol{\theta}_j^{\dagger})} \quad \text{and} \quad \mathrm{p}(\boldsymbol{x}_l \mid \boldsymbol{\theta}) \propto \frac{1}{s} \exp\left(-\frac{F(\boldsymbol{x}_i, \boldsymbol{\theta})^2}{2s^2}\right)$$

Multi-Structure Fitting (AKSWH)

The differences between AKSWH and J-linkage:

- A consensus set in J-linkage is a set of classifications of the parameter hypothesis with respect to one data point. The J-distances of all possible pairs of the data points must be calculated.
- In contrast, the consensus set in AKSWH is the inlier/outlier binary classification of all data points with respect to one model hypothesis. One can calculate the J-distances only for the significant hypotheses, by which the computational efficiency can be greatly improved.

Multi-Structure Fitting (AKSWH)

The differences between AKSWH and J-linkage:

- J-linkage clusters the pairs of data points, and selects as the inliers the data points belonging to one cluster when the number of the data points is larger than a user-specified threshold.
- AKSWH directly clusters the pairs of hypotheses in parameter space and does not use any threshold to determine the number of clusters.

Multi-Structure Fitting (AKSWH)

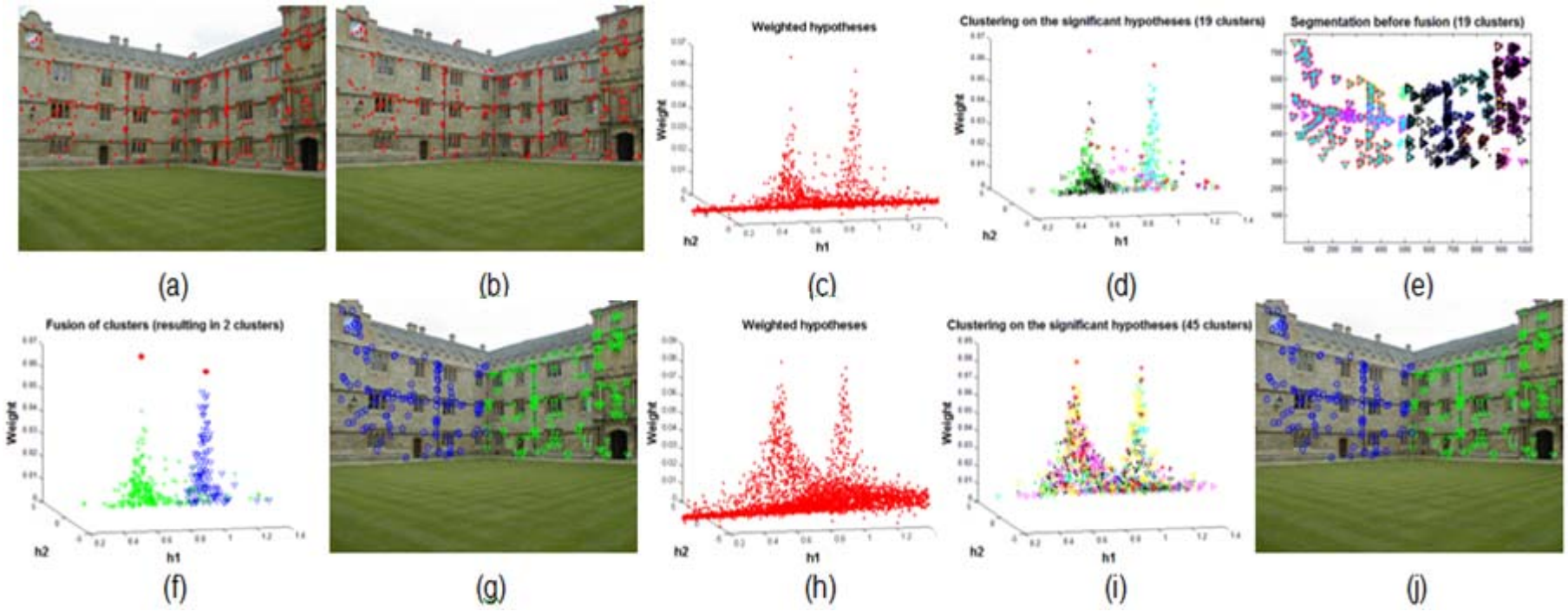


Figure 16. An example illustrating the main steps of the AKSWH method. (a) and (b) The input image pair with the matched SIFT feature points. (c) to (g) some results obtained by AKSWH2 (using the variable-scale weight function). (h) to (j) some results obtained by AKSWH1 (using the fixed-scale weight function). 384 /806 significant hypotheses are selected from 5000 hypotheses and 19 /45 clusters are obtained by AKSWH2/AKSWH1, respectively.

Experimental Results

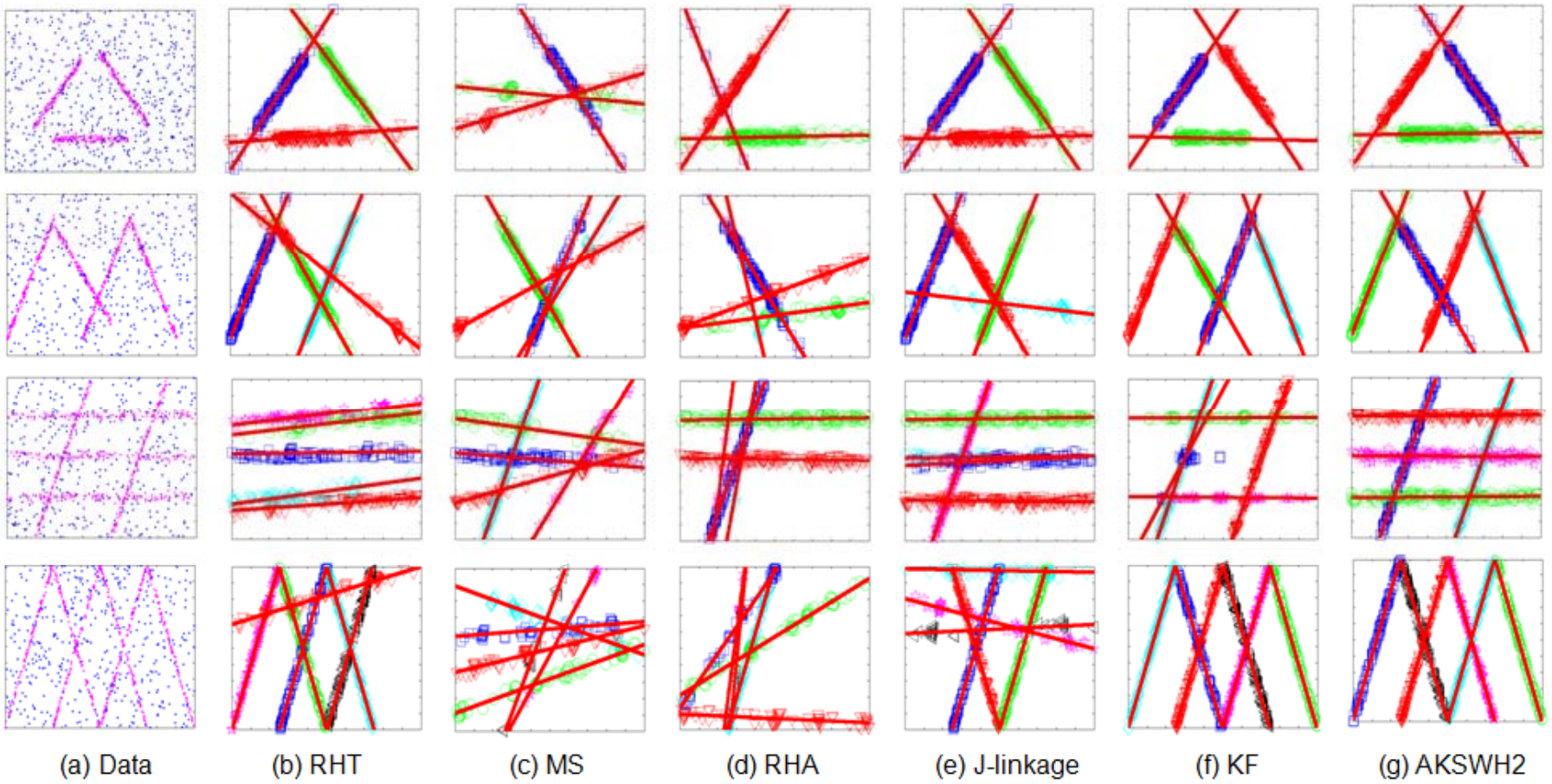


Figure 17. Examples for line fitting and segmentation. 1st to 4th rows respectively fit three, four, five and six lines. The corresponding outlier percentages are respectively 85%, 85%, 87% and 90%. The inlier scale is 1.5. (a) The original data. (b) to (g) The results obtained by RHT, MS, RHA, J-linkage, KF and AKSWH2, respectively.

Experimental Results

TABLE 2

THE FITTING ERRORS IN PARAMETER ESTIMATION (AND THE CPU TIME IN SECONDS).

	M1	M2	M3	M4	M5	M6	M7	M8
3 lines	1.28 (0.80)	1.21 (6.72)	23.7 (7.56)	25.1 (8.12)	1.17 (25.2)	0.99 (177)	1.18 (15.4)	1.13 (4.52)
4 lines	13.3 (0.84)	1.16 (8.10)	24.2 (5.87)	48.5 (7.31)	8.67 (19.5)	1.06 (223)	1.16 (17.7)	1.12 (4.02)
5 lines	19.9 (0.79)	1.40 (8.46)	7.35 (7.26)	116 (7.86)	10.9 (23.7)	9.27 (116)	1.25 (14.2)	1.29 (4.12)
6 lines	15.7 (1.21)	1.21 (11.0)	47.0 (5.53)	178 (9.06)	27.6 (23.2)	1.20 (685)	1.17 (20.2)	1.15 (3.43)

(M1-RHT; M2-ASKC; M3-MS; M4-RHA; M5-J-LINKAGE; M6-KF; M7-AKSWH1; M8-AKSWH2. WE RUN THE APPAROACHES ON A LAPTOP WITH AN I7 2.66GHZ CPU IN WINDOW 7 PLATFORM)

Experimental Results

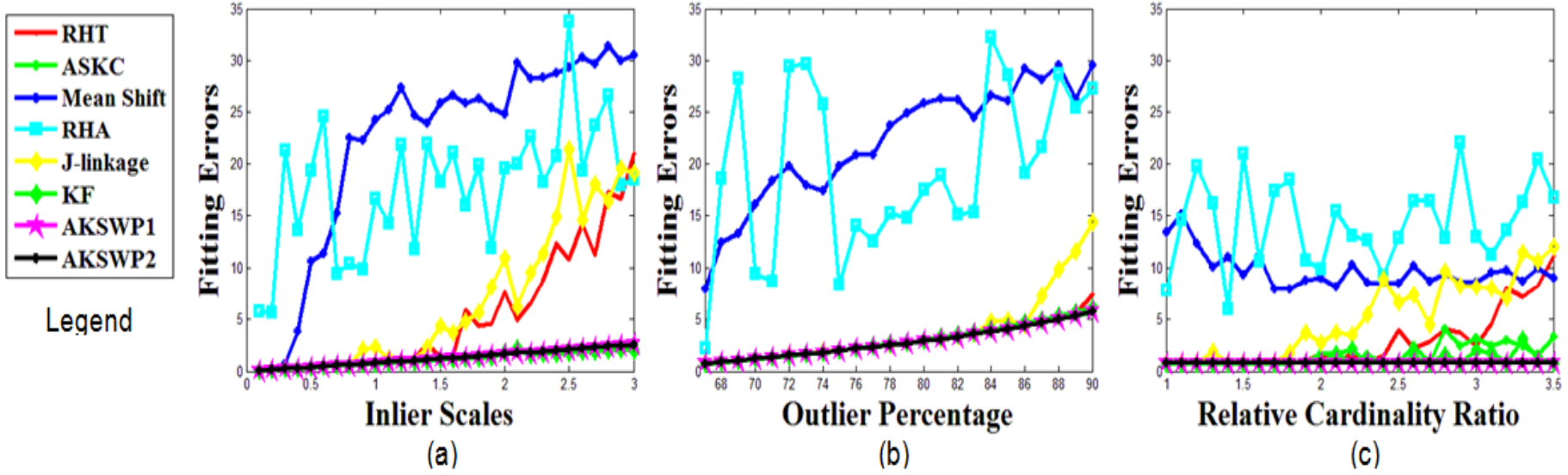


Figure 18. The average results obtained by the eight approaches. (a-c) respectively shows the influence of inlier scale, outlier percentage, and the relative cardinality radio of outliers to inliers.

Experimental Results

TABLE 3
THE FITTING ERRORS IN PARAMETER ESTIMATION.

	Inlier scale		Outlier percentage		Cardinality ratio	
	Std.Var.	Max.Err.	Std.Var.	Max.Err.	Std.Var.	Max.Err.
RHT	4.75	36.1	0.36	31.2	2.90	26.7
ASKC	0.06	3.01	0.10	6.23	2.30	18.5
MS	5.03	46.3	7.70	49.6	3.85	32.8
RHA	19.3	333	27.0	334	14.3	143
J-linkage	5.60	41.0	2.06	45.6	5.95	22.6
KF	0.19	8.14	0.11	6.41	1.07	20.4
AKSWH1	0.04	2.88	0.09	6.01	0.02	0.86
AKSWH2	0.06	2.75	0.09	6.09	0.03	0.86

Experimental Results

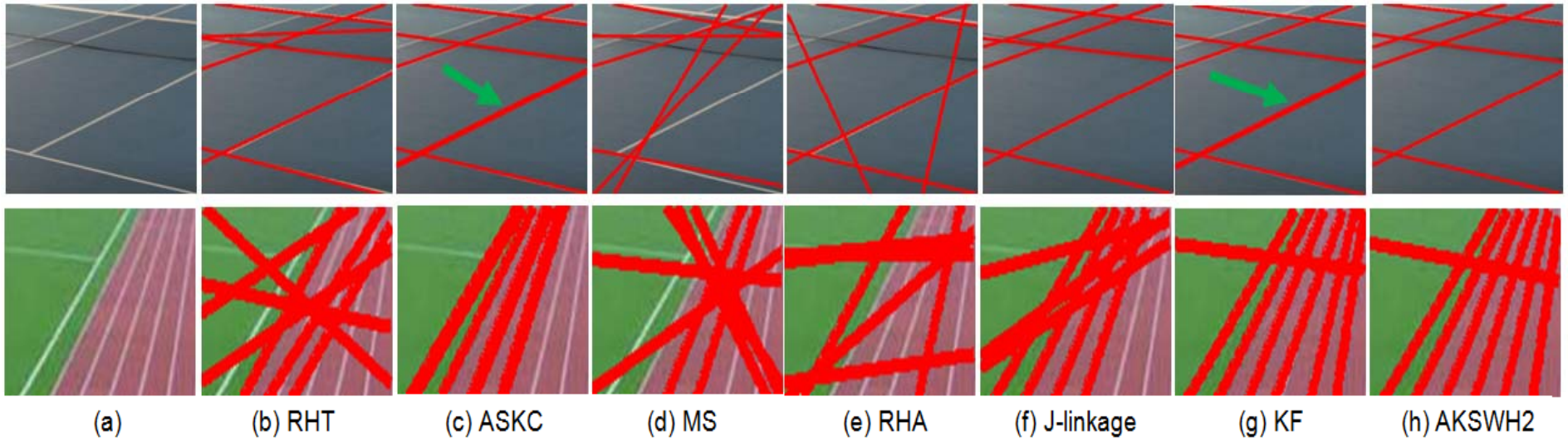


Figure 19. Examples for line fitting with real images. 1st ("tennis court") and 2th ("tracks") rows respectively fit six and seven lines. (a) The original images; (b) to (g) are the results obtained by RHT, ASKC, MS, RHA, J-linkage, KF and AKSWH2, respectively.

Experimental Results

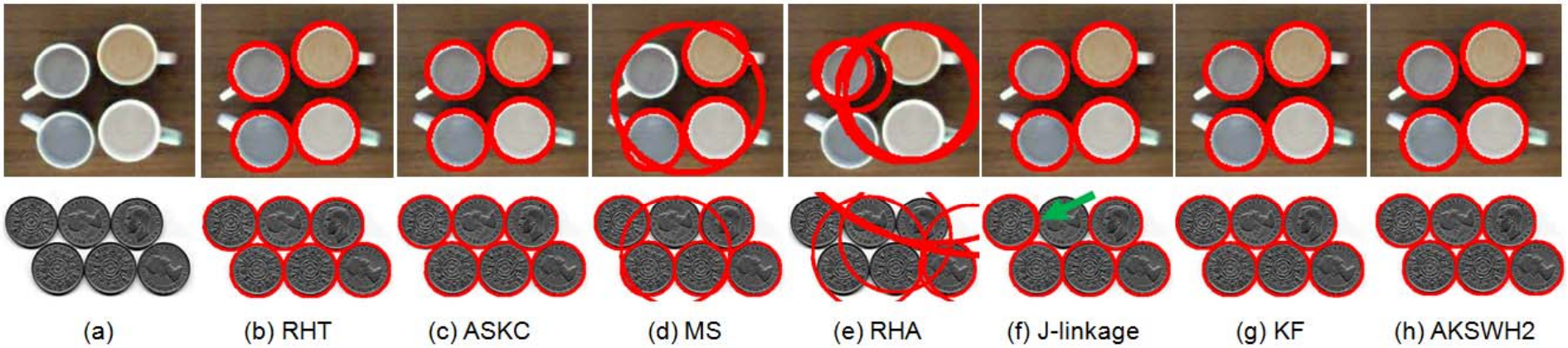


Figure 20. Examples for circle fitting. 1st ("cups") to 2th ("coins") rows respectively fit four and six circles. (a) The original images; (b) to (h) The results obtained by RHT, ASKC, MS, RHA, J-linkage, KF and AKSWH2, respectively.

Experimental Results

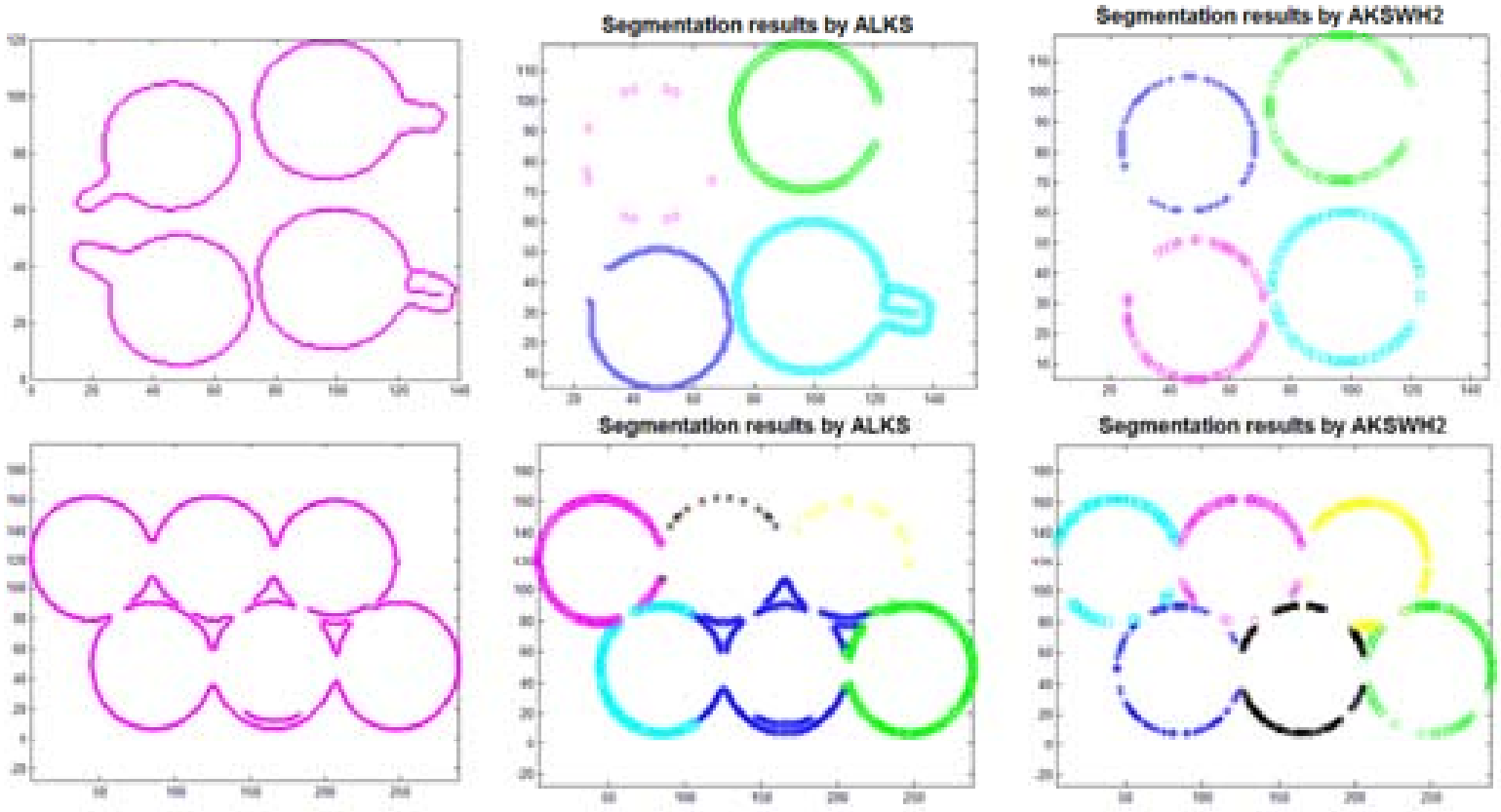


Figure 21. The segmentation results by ALKS and AKSWH2.

Experimental Results

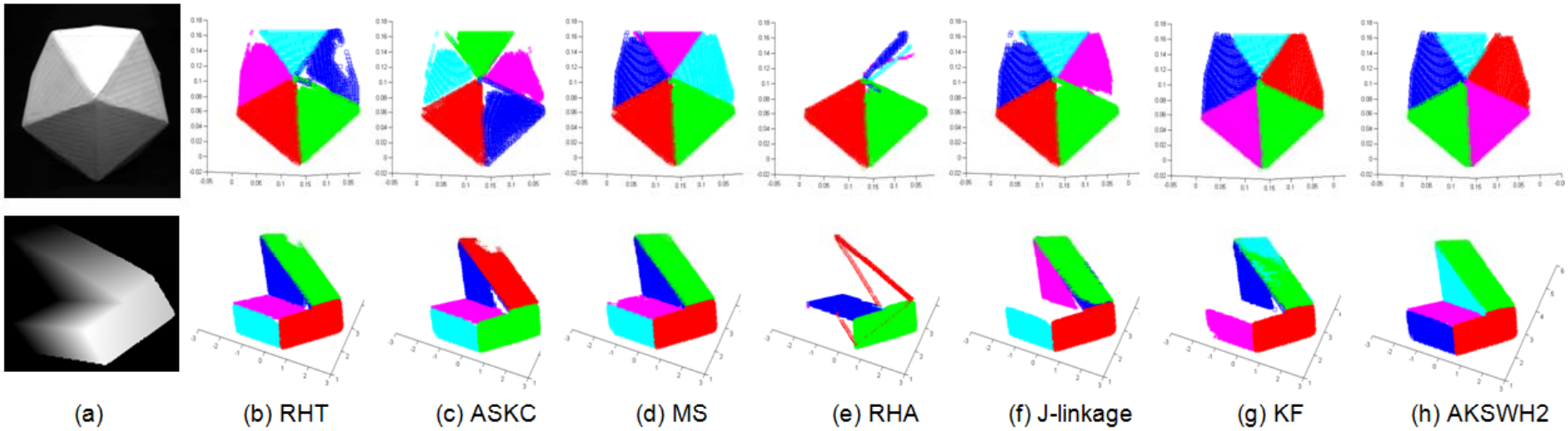


Figure 21. Examples for range image segmentation. 1st ("five planes" including 10842 data points) to 2th ("block" having 12069 data points) rows fit five planes. (a) The original images; (b) to (g) The segmentation results obtained by RHT, ASKC, MS, RHA, J-linkage, KF and AKSWH2, respectively.

For the computational efficiency, AKSWH1/2 are **more than one order faster** than J-linkage, and **more than two order faster** than KF !

Experimental Results

TABLE 4

THE CPU TIME USED BY THE APPROACHES(IN SECONDS)

	M1	M2	M3	M4	M5	M6	M7	M8
court	0.74	18.3	7.98	15.2	273	4600*	22.1	10.8
tracks	0.85	9.92	6.57	22.5	31.2	167	24.5	4.23
cups	2.19	14.9	11.1	10.8	20.2	51.4	9.86	3.27
coins	1.25	26.7	3.91	12.5	28.8	1272	6.59	4.10
5planes	3.40	164	10.1	29.6	295*	5931*	11.9	15.1
blocks	3.12	164	19.7	27.0	310*	3445*	11.3	19.8

(M1-RHT; M2-ASKC; M3-MS; M4-RHA; M5-J-LINKAGE; M6-KF; M7-AKSWH1; M8-AKSWH2.'*' MEANS THE APPROACH USES THE RE-SAMPLED DATA POINTS)

Experimental Results

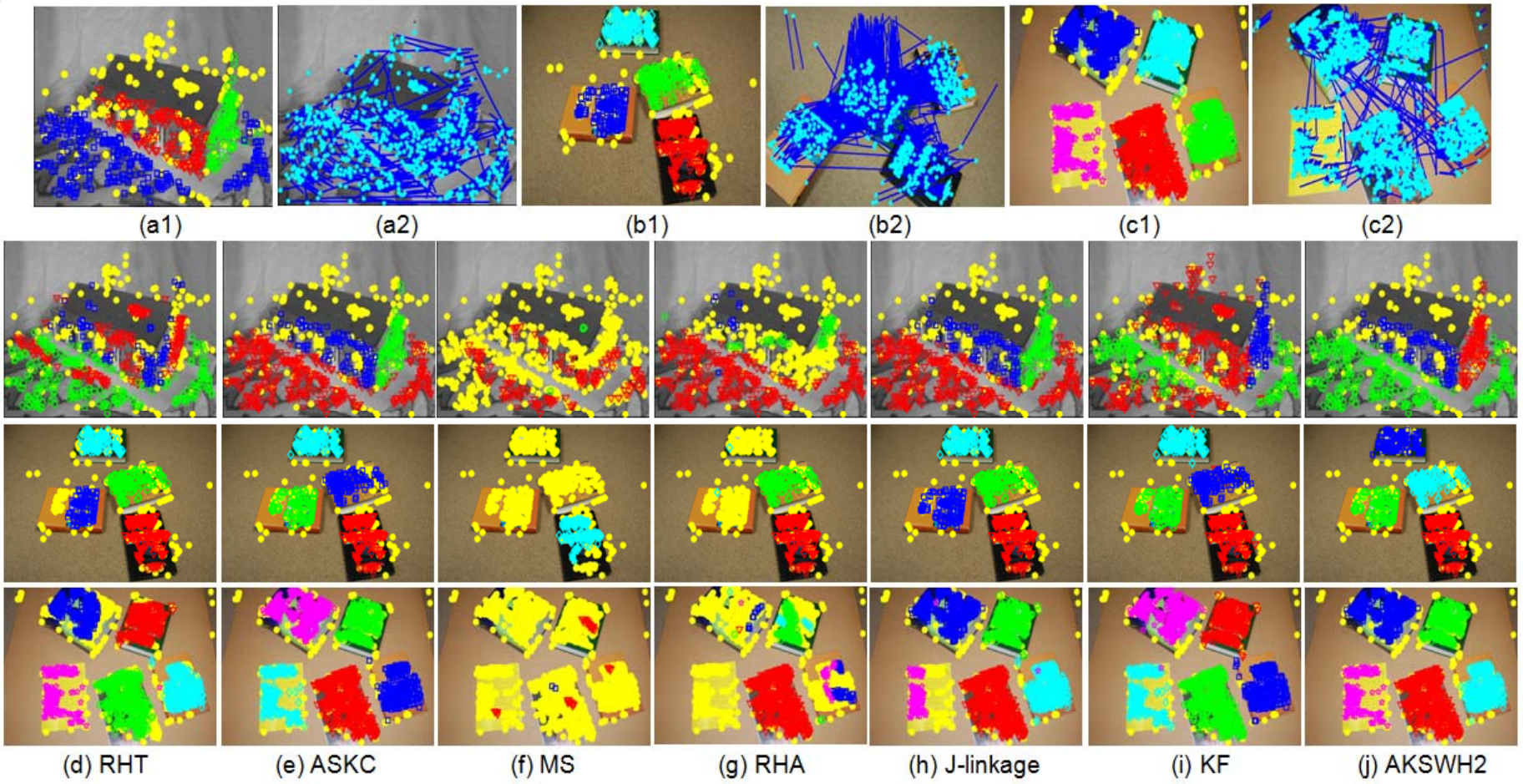


Figure 22. Estimating homographies and segmenting multiple-structure data with the image pairs of "Model House" ("MH"), four books ("4B") and five books ("5B"). (a1), (b1) and (c1) show the left images with the ground truth segmentation results superimposed. The yellow dots are the outliers. (a2), (b2) and (c2) show the right images with the disparities of corresponding points superimposed. 2nd to 4th rows are the segmentation results obtained by RHT, ASKC, MS, RHA, J-linkage, KF and AKSWH2, respectively.

Experimental Results

TABLE 5

THE FITTING ERRORS OBTAINED BY THE EIGHT APPROACHES AND THE CPU TIME USED (IN SECONDS).

	MinN	MaxN	TN	M1	M2	M3	M4	M5	M6	M7	M8
MC2	131	200	347	1.89 (14.8)	1.24 (30.8)	44.6 (147)	4.47 (9.02)	1.29 (15.4)	1.87 (302)	1.29 (28.3)	1.28 (16.8)
MC3	136	303	799	1.36 (16.0)	0.49 (54.7)	507 (121)	361 (13.7)	0.66 (26.5)	0.66 (85.8)	0. 52 (29.6)	0.50 (12.5)
MH	105	294	702	39.5 (39.5)	1.57 (55.3)	365 (136)	151 (11.6)	2.43 (29.0)	15.6 (64.5)	1.85 (29.3)	1.56 (15.1)
4B	122	231	777	1.79 (33.0)	0.86 (123)	467 (189)	350 (17.8)	0.95 (33.6)	2.11 (989)	0.92 (20.7)	0.87 (14.9)
5B	257	577	2394	1.27 (27.4)	0.43 (168)	132 (187)	244 (29.4)	0.85 (157)	0.62 (8996)	0.50 (36.0)	0.44 (24.3)

(M1-RHT; M2-ASKC; M3-MS; M4-RHA; M5-J-linkage; M6-KF; M7-AKSWH1; M8-AKSWH2.)

Experimental Results

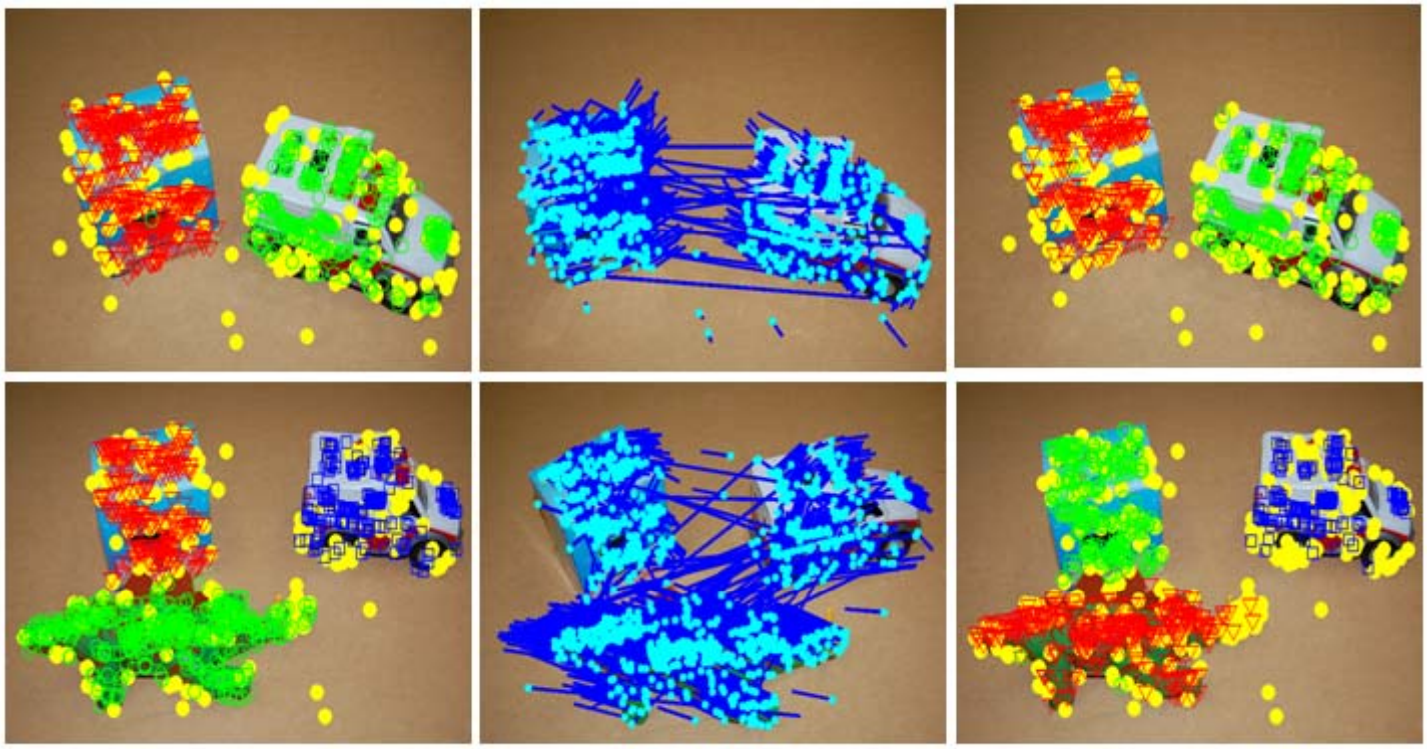


Figure 23. The segmentation results obtained by AKSWH2 for the image pairs of “Box-Car” (“BC”) and “Box-Car-Dinosaur” (“BCD”) respectively. (a) shows the left image with the ground truth segmentation superimposed. (b) shows the right image with the disparities of corresponding points superimposed. (c) shows the results obtained by AKSWH2.

Experimental Results

TABLE 6

THE FITTING ERRORS OBTAINED BY THE EIGHT APPROACHES AND THE CPU TIME USED (IN SECONDS).

	MinN	MaxN	TN	M1	M2	M3	M4	M5	M6	M7	M8
BC	378	500	1116	0.39 (12.6)	0.17 (145)	5.92 (139)	14.5 (149)	0.27 (116)	0.15 (626)	0.14 (198)	0.15 (64.5)
BCD	232	460	1227	29.7 (13.4)	0.53 (197)	26.5 (229)	36.4 (168)	27.6 (123)	0.51 (519)	0. 51 (274)	0.26 (61.5)

(M1-RHT; M2-ASKC; M3-MS; M4-RHA; M5-J-linkage; M6-KF; M7-AKSWH1; M8-AKSWH2)

Multi-Structure Fitting (AKSWH)

Advantages:

- ✓ It can simultaneously estimate not only the parameters of and the scales of model instances in data, but also the number of model instances in the data.
- ✓ It does not require to specify the inlier scale, which is adaptively estimated.
- ✓ It is computationally efficient.

Disadvantages:

- ☐ It uses a fixed K value.
- ☐ The IKOSE scale estimator requires the residuals of inliers are Gaussian-like distributed.



Higher Order Constraints for Multi-Structure Fitting (PEARL)

- Isack and Yuri [IJCV12] formulate geometric multi-model fitting as an optimal labeling problem.
- It optimizes the labels of the data with a global energy function, which balances geometric errors and regularity of inlier clusters.
- Regularization is based on spatial coherence (on some near-neighbor graph).
- PEARL converges to a local minimum of the energy and automatically selects a small number of models that best explain the whole data set.

ACCV 2012

Higher Order Constraints for Multi-Structure Fitting (PEARL)

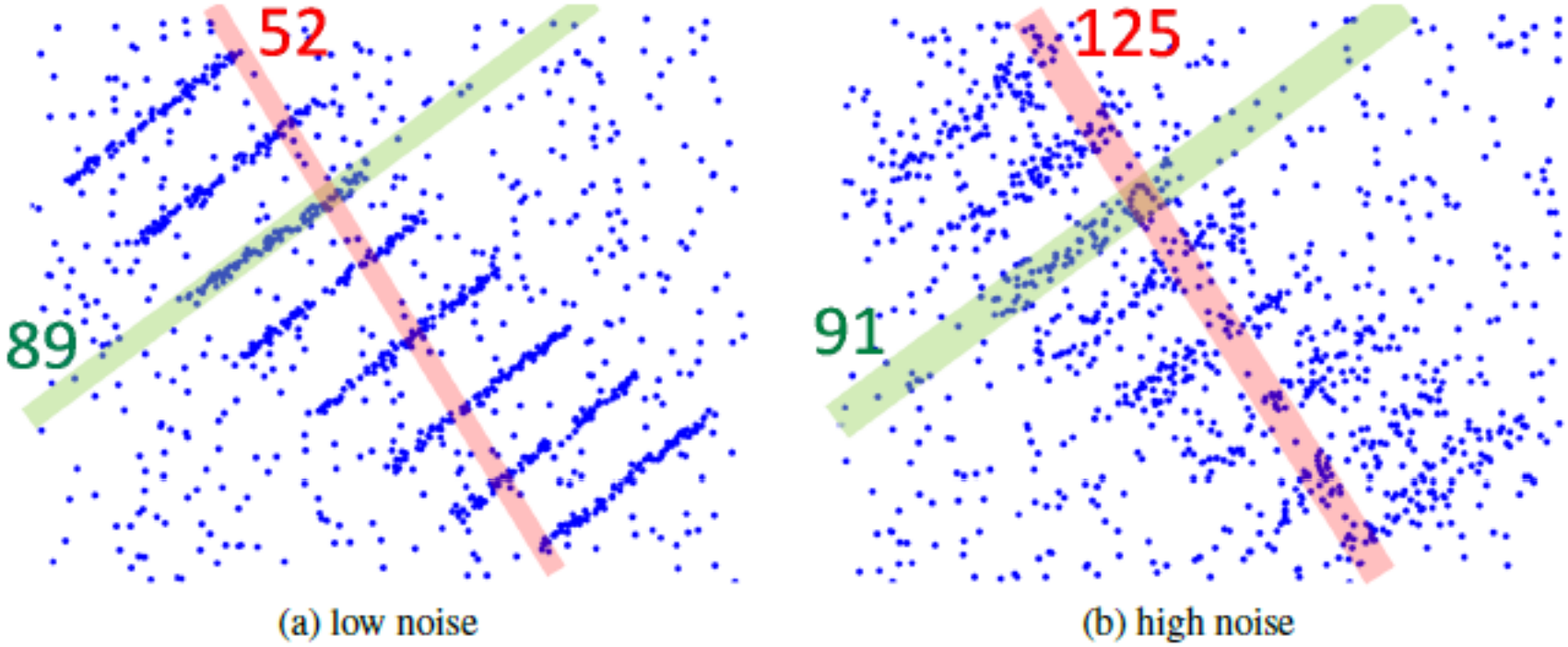


Figure 24. In multi-model cases, the criteria of maximizing the number of inliers may work for low levels of noise (a). However, higher noise levels may cause a failure case as some random model (red) may have more inliers than the true model (green). (from Isack and Yuri, IJCV12)



Higher Order Constraints for Multi-Structure Fitting (PEARL)

- The energy-based interpretation of the basic RANSAC algorithm is represented as the minimization of energy:

$$E(L) = \sum_{p} ||p - L||$$

where

$$||p - L|| = \begin{cases} 0 \text{ if } dist(p, L) < T \\ 1 \text{ otherwise} \end{cases}$$



Higher Order Constraints for Multi-Structure Fitting (PEARL)

- Let $\mathbf{L} = \{L_p | p \in P\}$ be the labeling of points in data set P. Model fitting could be formulated as minimization of energy $E(\mathrm{L})$ over labeling L.
- If the goal is only to minimize the fitting errors, the energy function is written as:

$$E(\mathbf{L}) = \sum_{p} ||p - L_p||. \qquad (1)$$

This function corresponds to overfitting as every point is assigned some perfectly fit model and there are no outliers.



Higher Order Constraints for Multi-Structure Fitting (PEARL)

- The above-mentioned energy function must be combined with some energy term to regularize the labeling.
- One form of regularization is to combine geometric errors with the label count penalty:

$$E(\mathbf{L}) = \sum_{p} ||p - L_p|| + \beta \cdot |\mathcal{L}_{\mathbf{L}}| \qquad (2)$$

where $\mathcal{L}_{\mathbf{L}}$ is the set of distinct models (labels) assigned to points.



Higher Order Constraints for Multi-Structure Fitting (PEARL)

- Considering the spatial regularization, the energy function (3) can be written as:

$$E(\mathbf{L}) = \sum_{p} ||p - L_p|| + \lambda \cdot \sum_{(p,q)\in\mathcal{N}} w_{pq} \cdot \delta(L_p \neq L_q), \qquad (3)$$

where N is some neighborhood (e.g. edges on some near-neighbor graph).

- The second term is a smoothness prior. Weights set discontinuity penalties for each pair of "neighboring" data points.

ACCV 2012

Higher Order Constraints for Multi-Structure Fitting (PEARL)

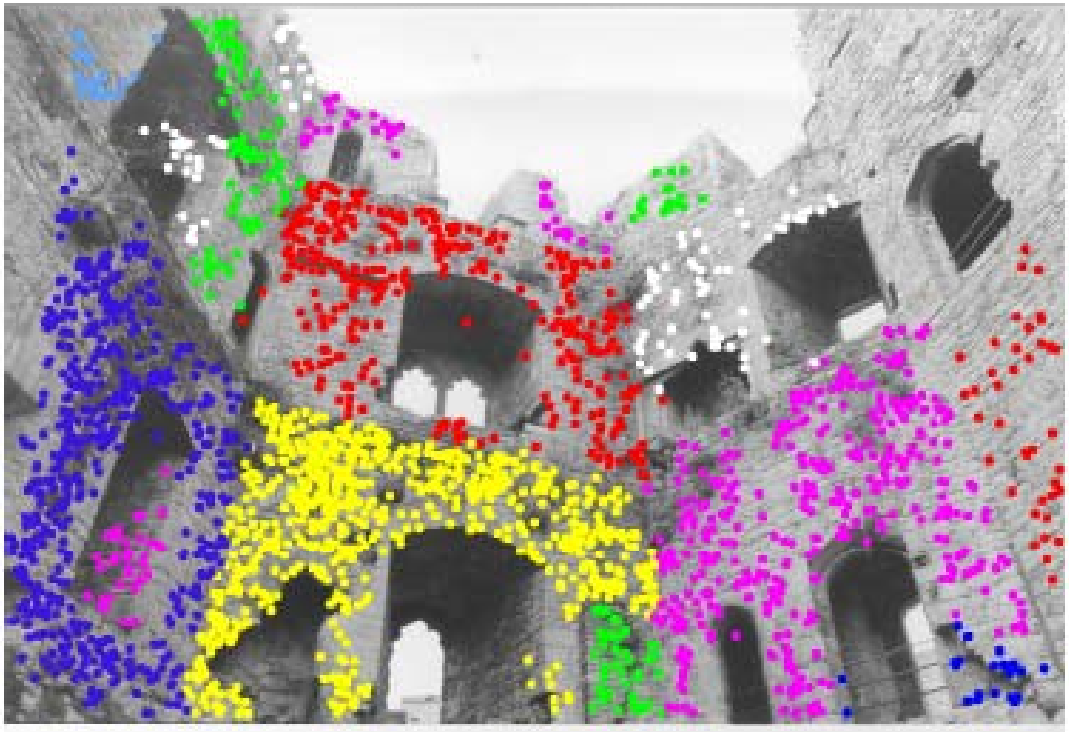
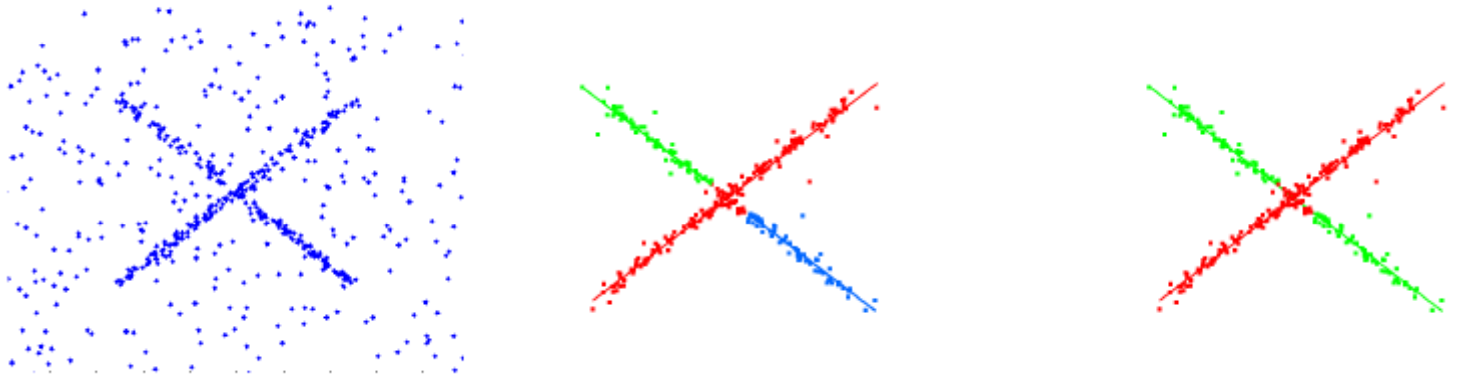


Figure 25. One example of fitting homographies. It motivates spatial regularization in geometric model fitting. (from Isack and Yuri, IJCV12)



Higher Order Constraints for Multi-Structure Fitting (PEARL)

- The optimal solution for the energy function (3) can not merge two models with very similar parameters:



(a) data (300 outliers) (b) minimum of energy (3) (c) *merging*, energy (4)

Figure 26. Optimization of energy (3) may leave spatially isolated groups of inliers assigned to 2 models even if their parameters are infinitesimally close (b). Per-label costs in energy (4) solves this problem (c). (from Isack and Yuri, IJCV12)



Higher Order Constraints for Multi-Structure Fitting (PEARL)

- Considering a more general combination of spatial smoothness with label counts, the energy function (4) can be written as:

$$E(\mathbf{L}) = \sum_{p} ||p - L_p|| + \lambda \cdot \sum_{(p,q)\in\mathcal{N}} w_{pq} \cdot \delta(L_p \neq L_q) + \beta \cdot |\mathcal{L}_{\mathbf{L}}|. \quad (4)$$

Some Results

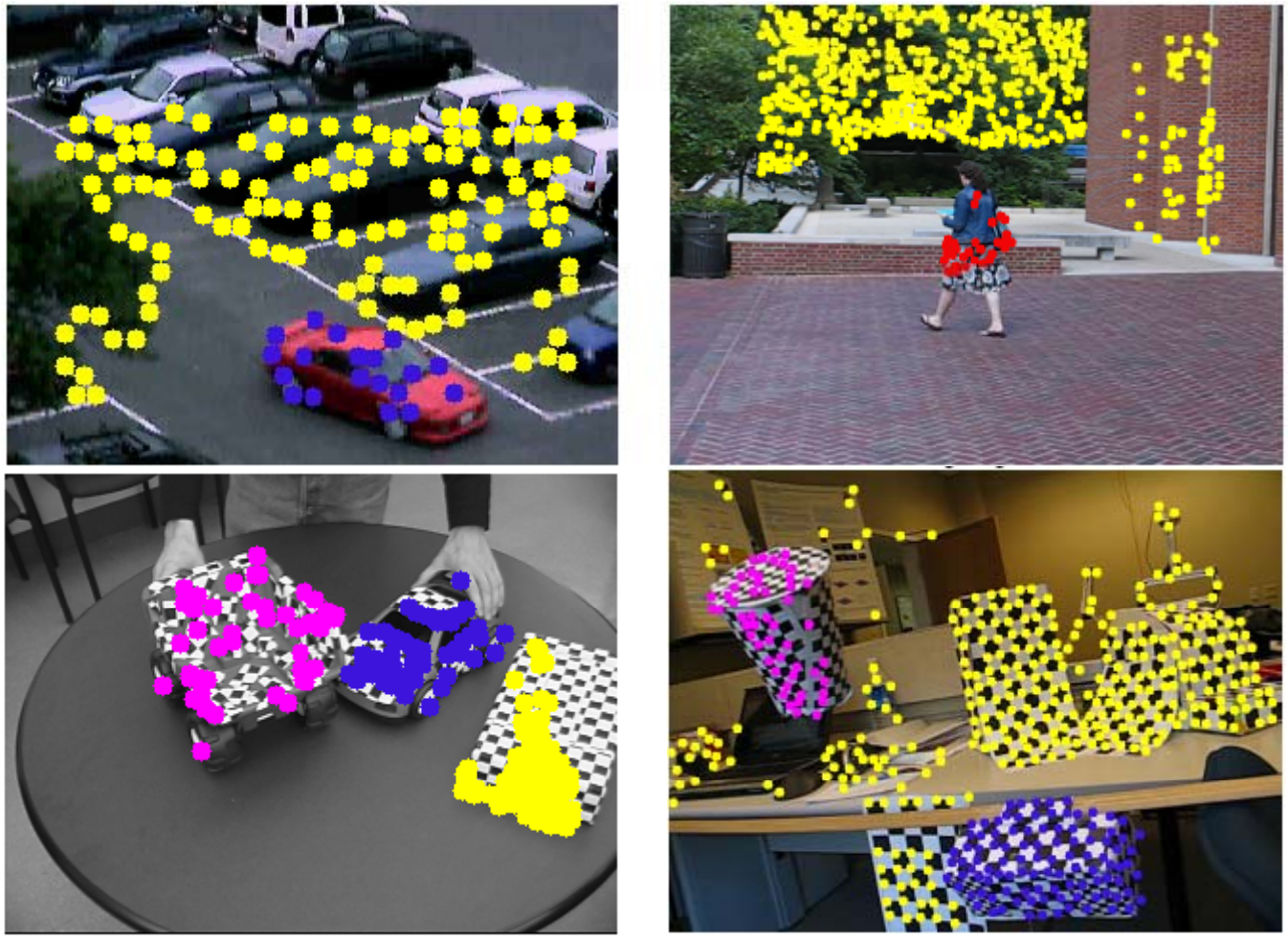


Figure 27. Representative PEARL's results. (from Isack and Yuri, IJCV12)



Higher Order Constraints for Multi-Structure Fitting (PEARL)

- ✓ It provides an effective way to handle multi-structure fitting from the minimization of energy point of view.

- ☐ However, it requires some thresholds in the energy function and in classifying weak models (e.g. those with small number of inliers).

Opening Questions

- How many structures do data exactly contain?
- How to robustly estimate the inlier scale of a model if the inliers do not satisfy a Gaussian-like distribution?
- How to effectively guide sampling data ?
- etc……

Some Source Codes and Datasets

- Kernel Fitting (KF): http://cs.adelaide.edu.au/~tjchin/doku.php?id=code
- J-Linkage: http://www.diegm.uniud.it/fusiello/demo/jlk/
- AKSWH (some datasets with ground truth): http://cs.adelaide.edu.au/~dsuter/code_and_data/index.html

……

Some References

- Hough, P.V.C., *Methods and means for recognising complex patterns*, in *U.S. Patent 3 069 654*. 1962.
- XU, L., E. OJA, P. KULTANEN, *A New Curve Detection Method: Randomized Hough Transform (RHT).* Pattern Recognition Letters, 1990. 11(5). 331-338.
- Lee, K.-M., P. Meer, R.-H. Park, *Robust Adaptive Segmentation of Range Images.* IEEE Trans. Pattern Analysis and Machine Intelligence, 1998. 20(2). 200-205.
- Bab-Hadiashar, A., D. Suter, *Robust segmentation of visual data using ranked unbiased scale estimate.* ROBOTICA, International Journal of Information, Education and Research in Robotics and Artificial Intelligence, 1999. 17. 649-660.
- Wang, H., D. Suter, *Robust Adaptive-Scale Parametric Model Estimation for Computer Vision.* IEEE Trans. Pattern Analysis and Machine Intelligence, 2004. 26(11). 1459-1474.
- Zuliani, M., C.S. Kenney, B.S. Manjunath. *The multiRANSAC Algorithm and Its Application to Detect Planar Homographies*. *IEEE International Conference on Image Processing*. 2005. 153-156.
- Tuzel, O., R. Subbarao, P. Meer. *Simultaneous Multiple 3D Motion Estimation via Mode Finding on Lie Groups*. *IEEE International Conference on Computer Vision*. 2005. 18-25.
- Zhang, W., J. Kosecka, *Nonparametric Estimation of Multiple Structures with Outliers.* LNCS, DVW, 2006. 4358. 60-74.
- Toldo, R., A. Fusiello. *Robust Multiple Structures Estimation with J-Linkage*. *European Conference on Computer Vision*. 2008. 537-547.
- Subbarao, R., P. Meer, *Nonlinear Mean Shift over Riemannian Manifolds.* International Journal on Computer Vision, 2009. 84(1). 1-20.
- Chin, T.-J., H. Wang, D. Suter. *Robust Fitting of Multiple Structures: The Statistical Learning Approach*. *IEEE International Conference on Computer Vision*. 2009. 413-420.
- Wang, H., T.-J. Chin, D. Suter. Simultaneously Fitting and Segmenting Multiple-Structure Data with Outliers IEEE Trans. Pattern Analysis and Machine Intelligence. 2012. 34(6), 1177 - 1192.
- Isack, H., Y. Boykov. Energy-based Geometric Multi-Model Fitting. International Journal of Computer Vision. 2012. 97(2), 123-147.