Model fitting and regularization

(discrete optimization approach)

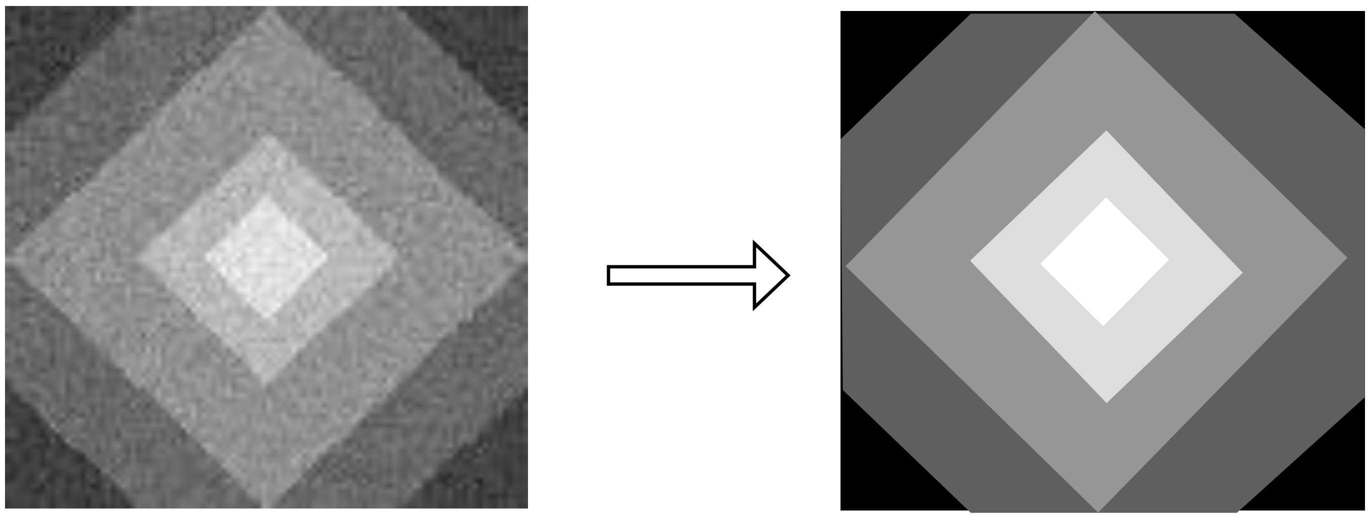
Yuri Boykov

Overview

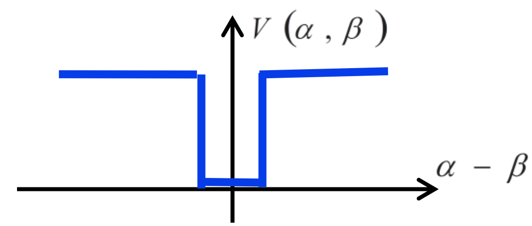
- Label costs (high-order sparsity prior)
- Model fitting
  - dealing with continuum of labels
  - K-means and EM + regularization
- Applications
  - unsupervised image segmentation, compression
  - geometric model fitting (lines, circles, planes, homographies, motion,…)

Potts model
(piece-wise constant labeling)

$$E(\mathbf{L}) = \sum_p (L_p - I_p)^2 + \sum_{(p,q)\in N} V(L_p, L_q)$$



$$V(\alpha,\beta) = w \cdot [\alpha \neq \beta]$$

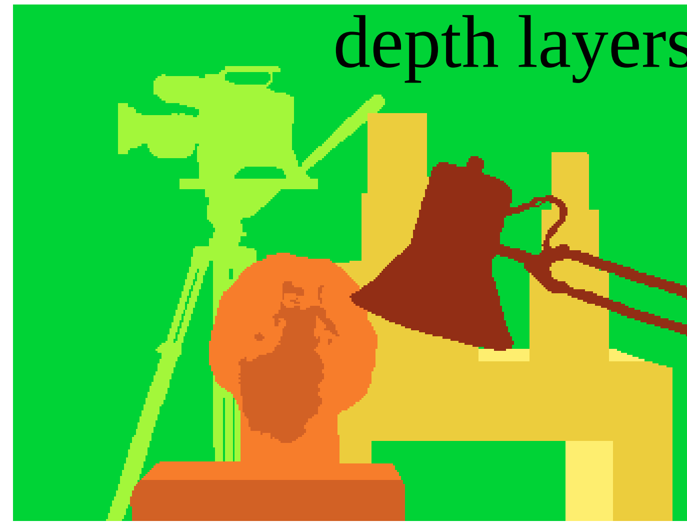


- Robust regularization
  - NP-hard, many local minima
  - provably good approximations (a-expansion)

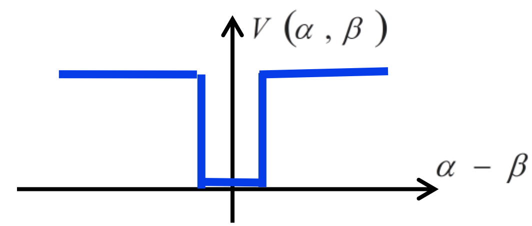
maxflow/mincut
combinatorial algorithms

Potts model
(piece-wise constant labeling)

$$E(\mathbf{L}) = \sum_p D_p(L_p) + \sum_{(p,q)\in N} V(L_p, L_q)$$



$$V(\alpha,\beta) = w \cdot [\alpha \neq \beta]$$



- Robust regularization
  - NP-hard, many local minima
  - provably good approximations (a-expansion)

maxflow/mincut
combinatorial algorithms

Adding label costs

$$E(\mathbf{L}) = \sum_p D_p(L_p) + \sum_{(p,q)\in N} V(L_p, L_q) + \sum_{L\in\Lambda} h_L \cdot \delta_L(\mathbf{L})$$

Λ - set of labels allowed at each point p

$$\delta_L(\mathbf{L}) = \begin{cases} 1, & \exists p : L_p = L \\ 0, & otherwise \end{cases}$$

- Leclerc [PAMI 89]
  - MDL framework, graduated non-convexity
- Zhu & Yuille [PAMI 96]
  - cont. framework (gradient descent + merging heuristics)
- Torr [PTRS 98], Li [CVPR 2007]
  - AIC/BIC framework, only 1st and 3rd terms
  - Seq. RANSAC heuristic (Torr) , LP relaxation w/o any guarantees (Li)
- Brox & Weikert [DAGM 04], Ayed & Mitiche [TIP'08]
  - Level-sets with merging heuristics (Brox)
  - Multi-level sets (Ayed)

Adding label costs

$$E(\mathbf{L}) = \sum_p D_p(L_p) + \sum_{(p,q)\in N} V(L_p, L_q) \quad + \sum_{L\subseteq\Lambda} h_L \cdot \delta_L(\mathbf{L})$$

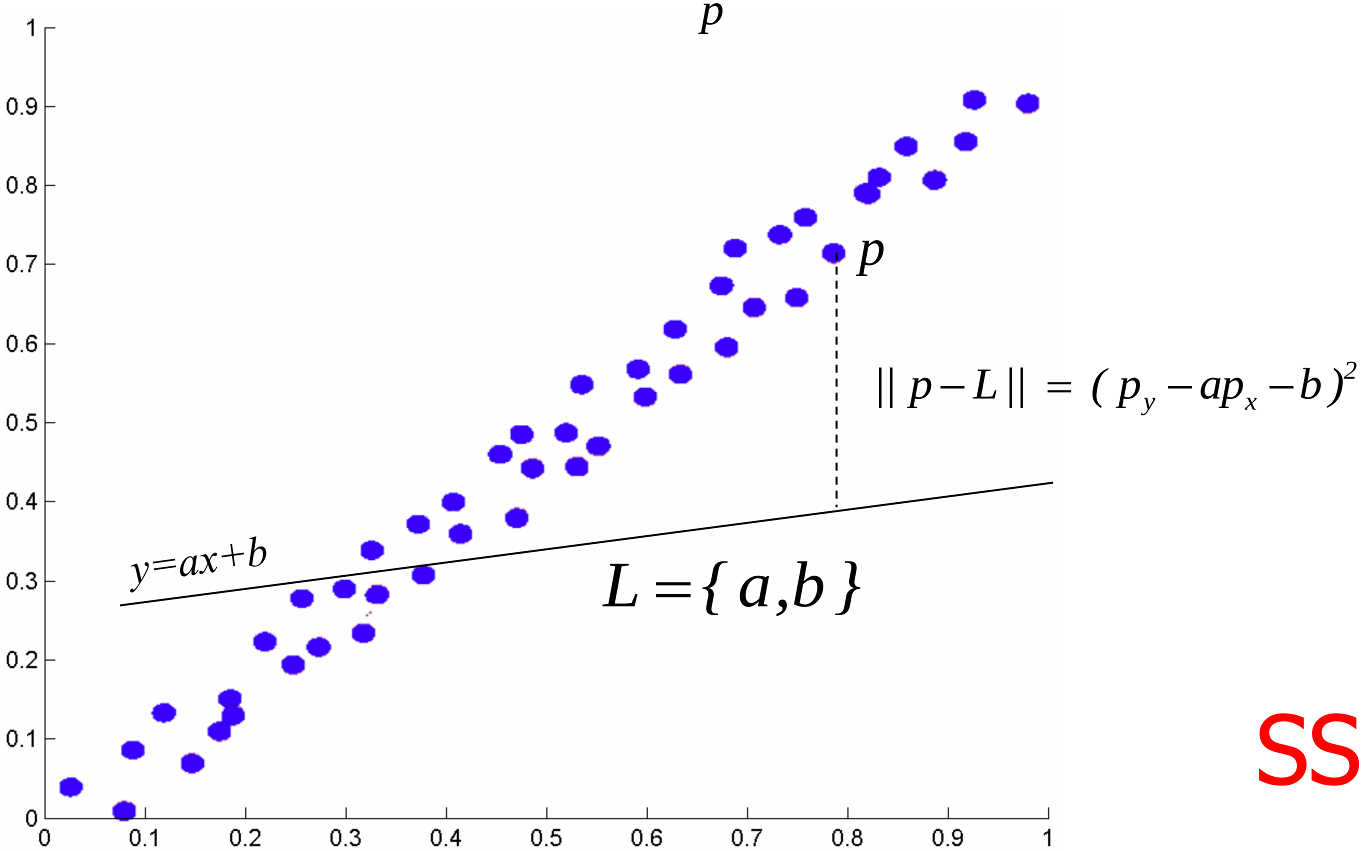
Λ - set of labels allowed at each point p

$$\delta_L(\mathbf{L}) = \begin{cases} 1, & \exists p : L_p \in L \\ 0, & otherwise \end{cases}$$

- Our work [CVPR 2010, IJCV 2011]
  - subsets of labels
  - multiple combinatorial algorithms w. optimality bounds
    - a-expansion++ (3rd term is a high-order clique)
    - UFL heuristics for 1st & 3rd term [Barinova et al., CVPR'10]
  - generic model fitting applications

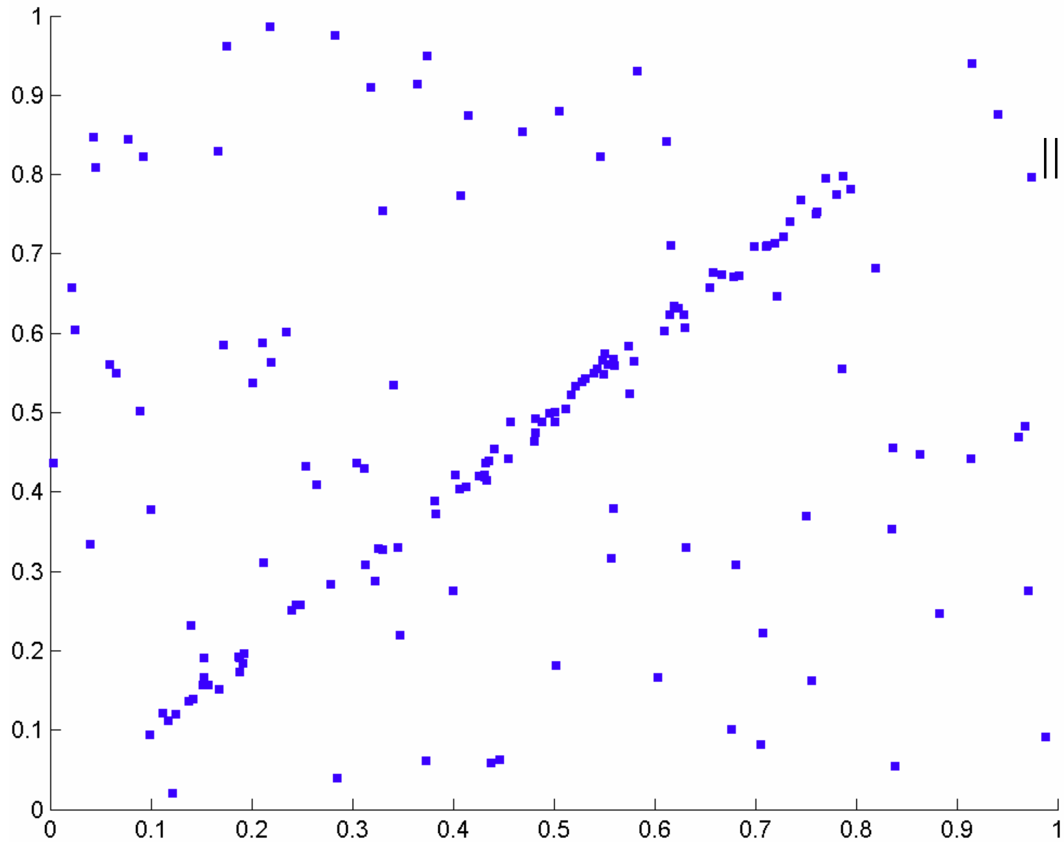
Model fitting

$$\hat{L} = \arg\min_{L} \sum_{p} \| p - L \|$$



SSD

many outliers



quadratic errors fail

use more robust error measures, e.g.

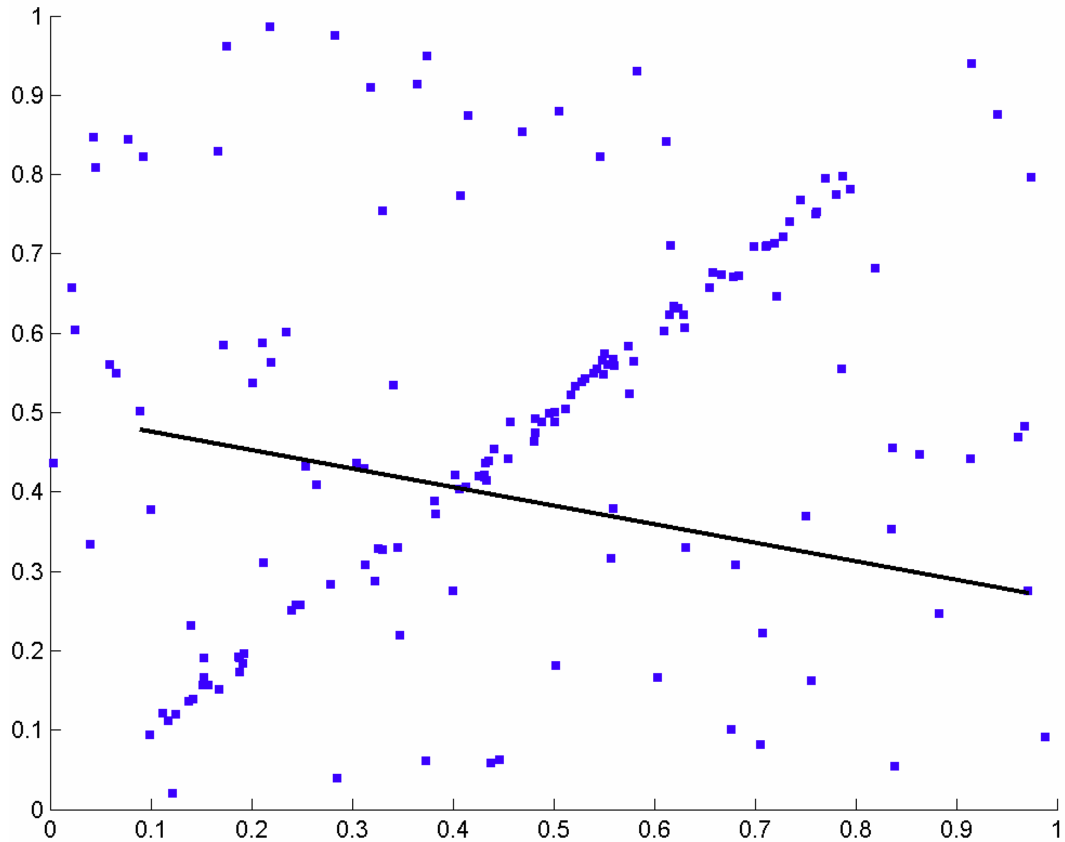
$$||\, p - L \,|| = |\, p_y - a p_x - b \,|$$

gives "MEDIAN" line

- more expensive computations (non-differentiable)
- **still fails if outliers exceed 50%**

RANSAC

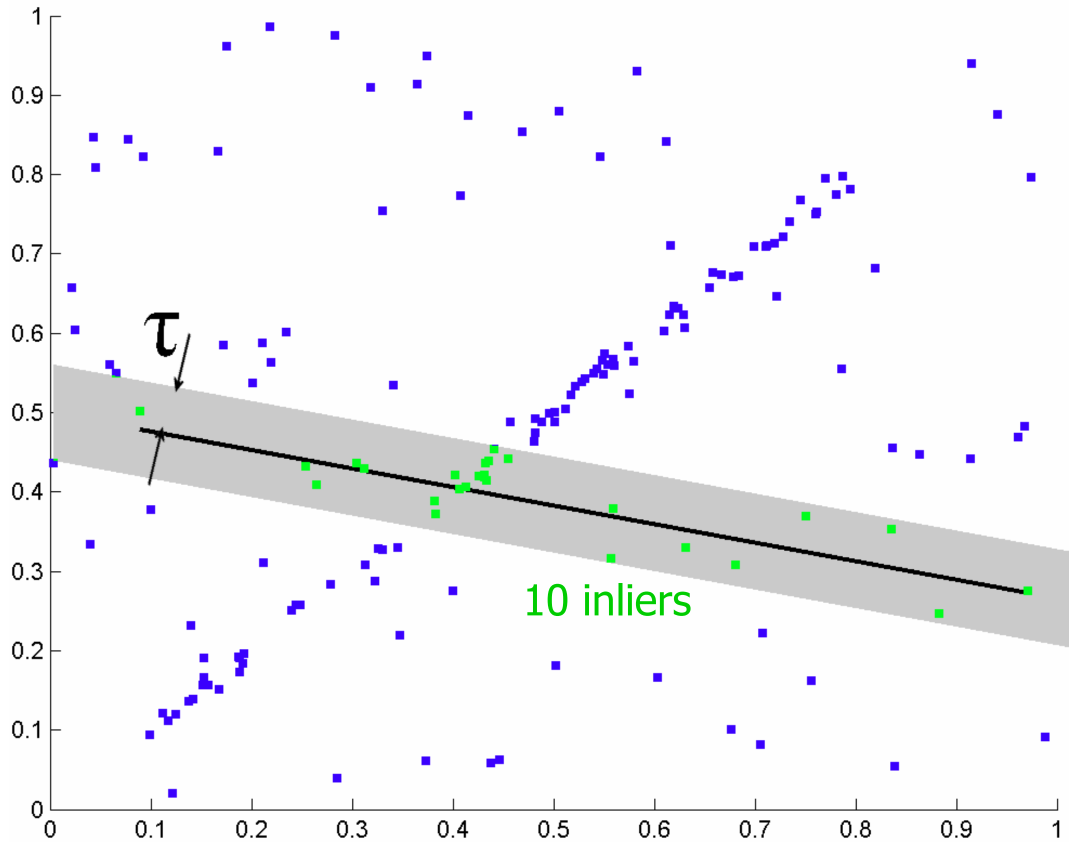
many outliers



1. sample randomly two points, get a line

RANSAC

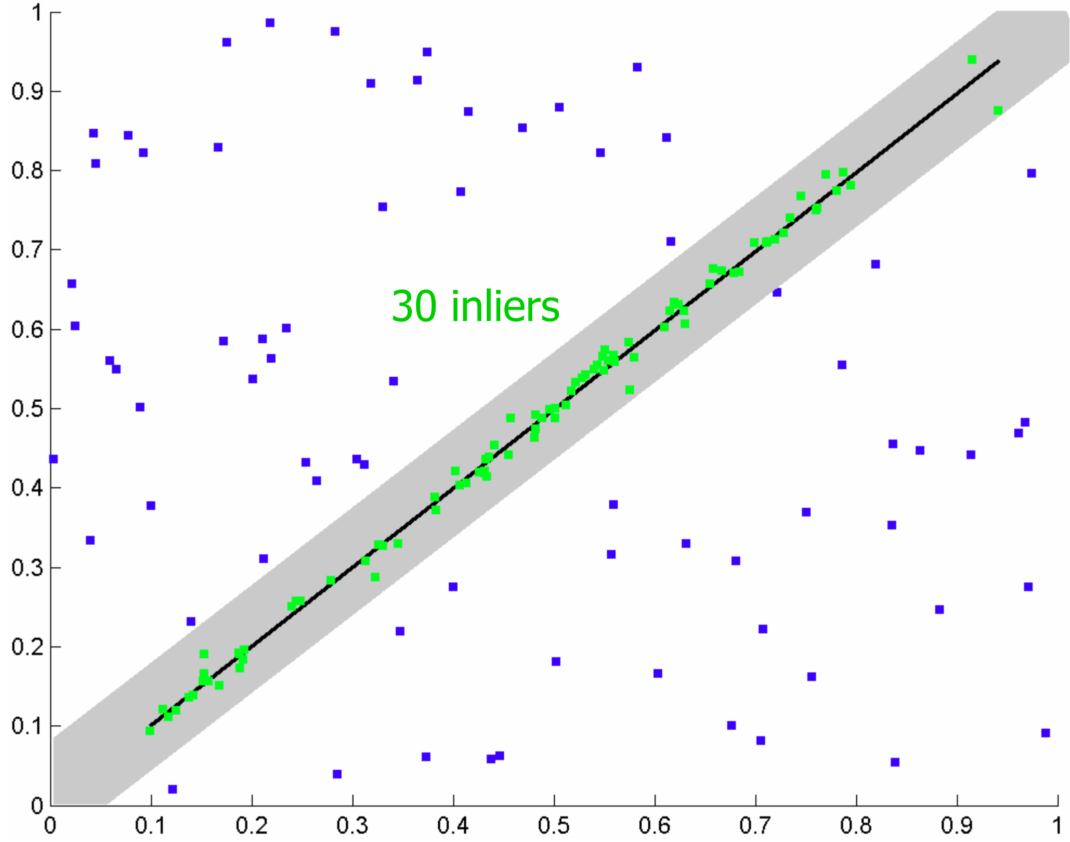
many outliers



1. sample randomly two points, get a line
2. count inliers for threshold T

RANSAC

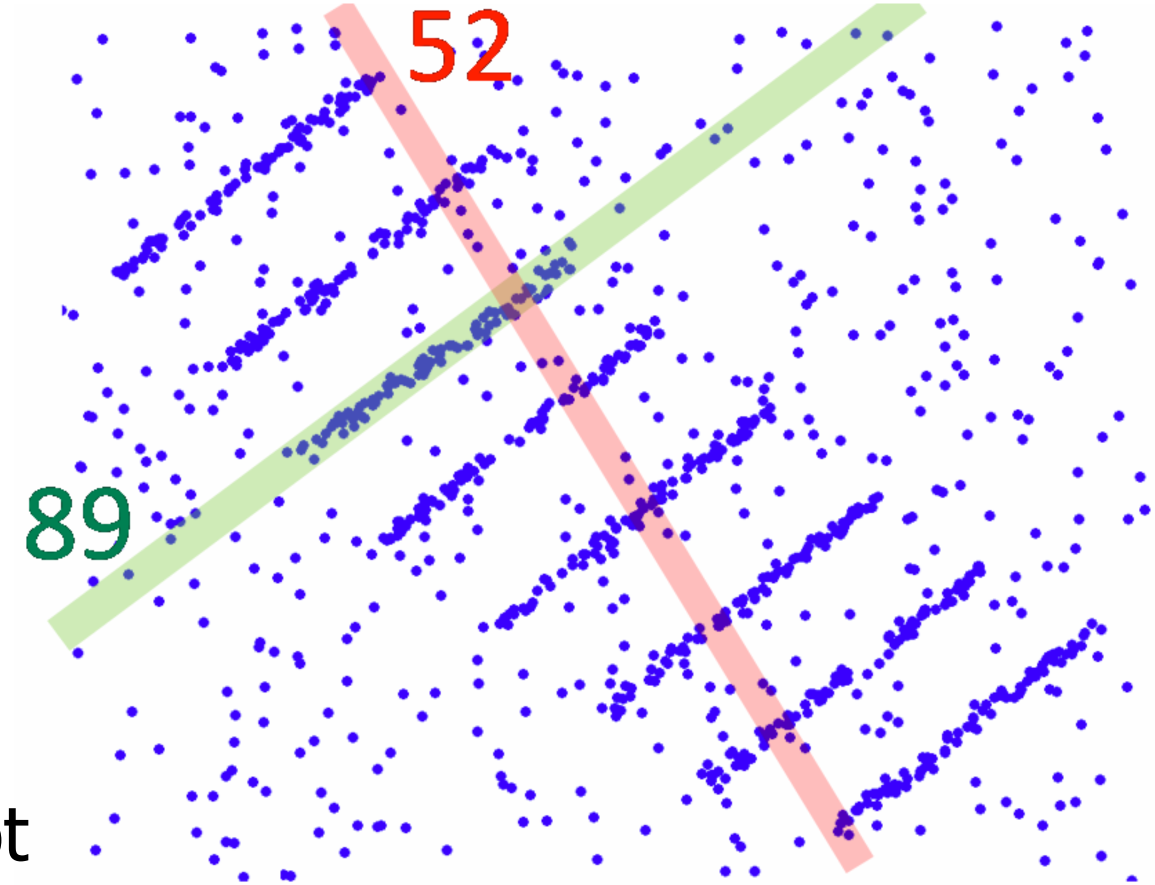
many outliers



1. sample randomly two points, get a line
2. count inliers for threshold *T*
3. repeat N times and select model with most inliers

RANSAC

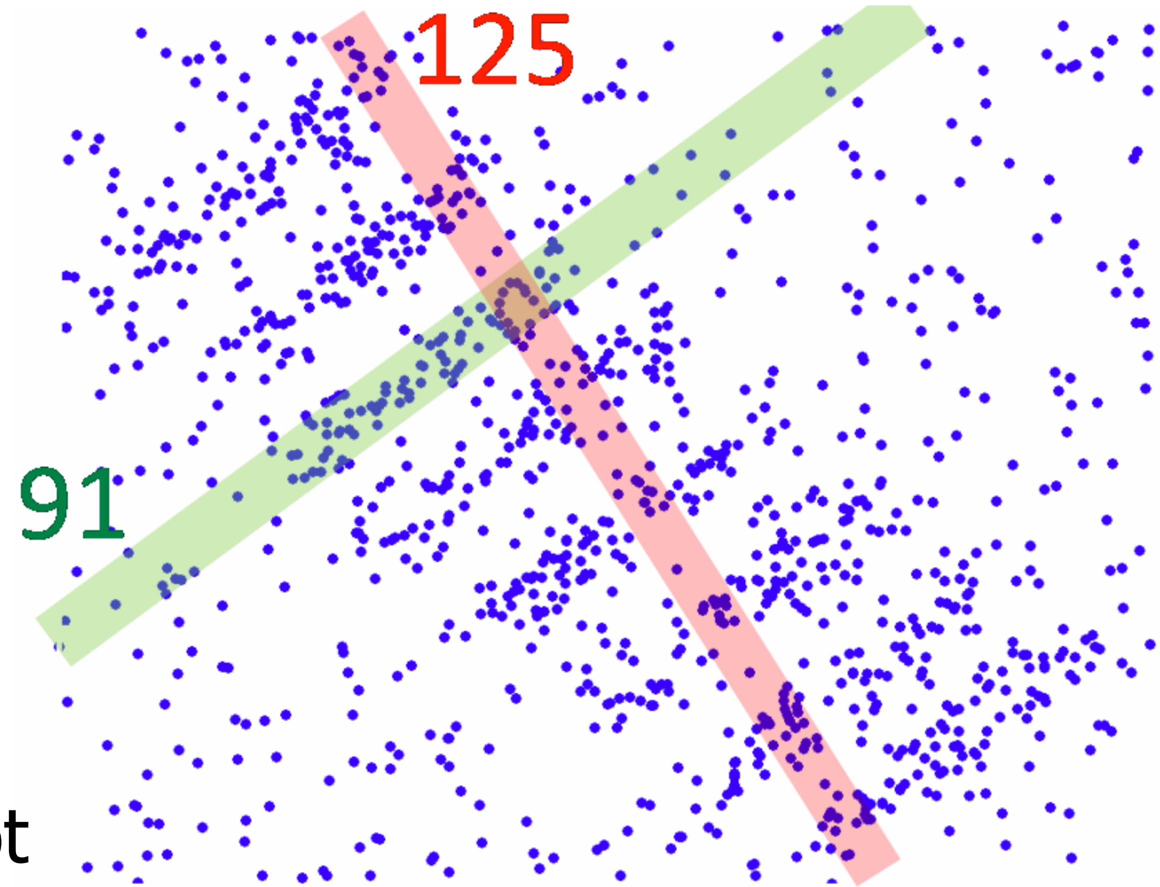
Multiple models and many outliers



Why not
RANSAC
again?

sequential RANSAC (Torr 98)

Multiple models and many outliers



Higher noise

Why not
RANSAC
again?

In general, maximization of inliers
does not work for
outliers + **multiple** models

Energy-based approach

$$E(L) = \sum_p \| p - L \|$$

energy-based interpretation
of RANSAC criteria for
single model fitting:

- <u>find optimal **label**</u> L
for one very specific
error measure

$$\| dist \| = \begin{cases} 0, & if \ \ dist \le T \\ 1, & if \ \ dist > T \end{cases}$$

Energy-based approach

$$E(\boldsymbol{L}) = \sum_p \| p - L_p \|$$

If **multiple** models

- assign different models (labels L_p) to every point p

- find optimal **labeling**

$\boldsymbol{L} = \{ L_1, L_2, \ldots, L_n \}$

Need regularization!

Energy-based approach

$$E(\boldsymbol{L}) \;=\; \sum_{p} \| p - L_p \| \;+\; \sum_{L \in \Lambda} h_L \cdot \delta_L(\boldsymbol{L})$$

If **multiple** models

- assign different models (labels L_p) to every point p

- find optimal **labeling**

$\boldsymbol{L} = \{ L_1, L_2, \ldots, L_n \}$

Λ - set of labels allowed at each point p

$$\delta_L(\mathbf{L}) = \begin{cases} 1, & \exists p : L_p = L \\ 0, & otherwise \end{cases}$$

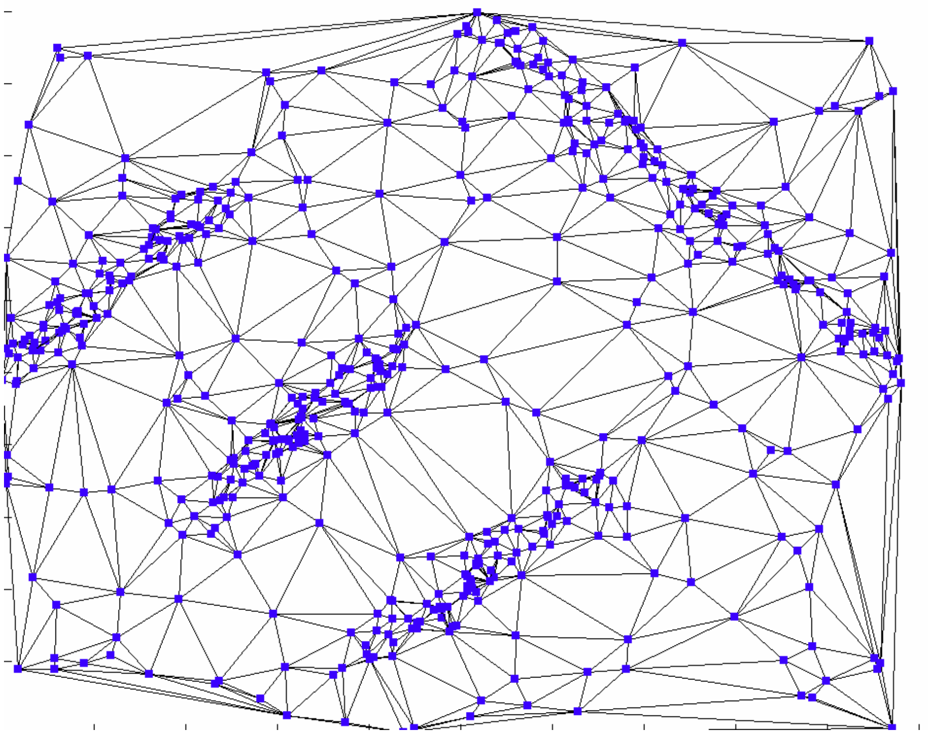
Energy-based approach

$$E(\mathbf{L}) = \sum_{p} || p - L_p || + \sum_{pq \in N} w \cdot [L_p \neq L_q]$$

If **multiple** models

- assign different models (labels L_p) to every point p

- find optimal **labeling** $\boldsymbol{L} = \{ L_1, L_2, \ldots, L_n \}$



Energy-based approach

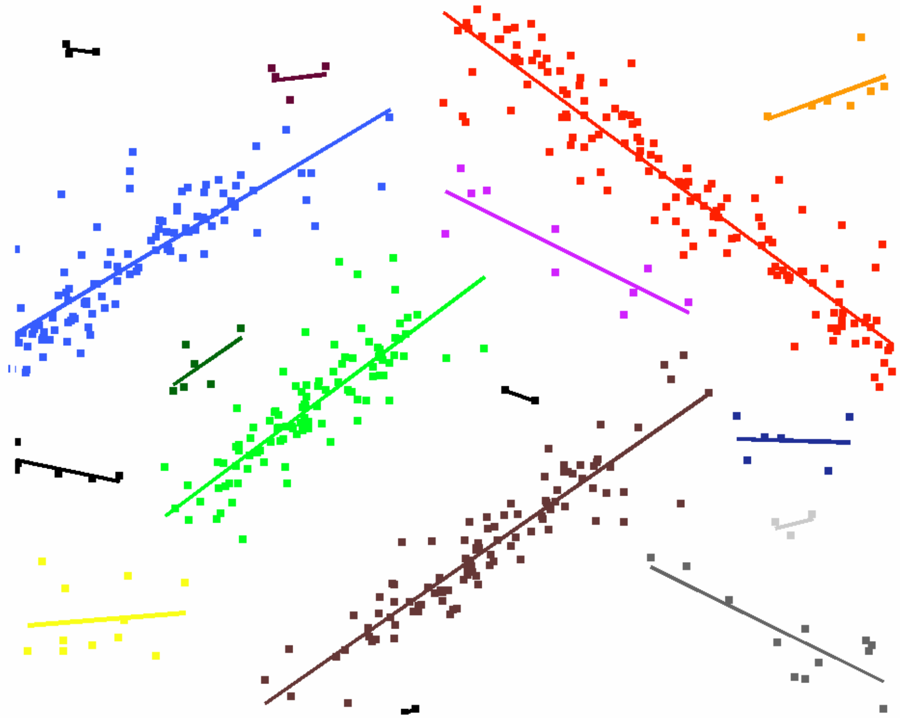
$$E(\mathbf{L}) = \sum_p || p - L_p || + \sum_{pq \in N} w \cdot [L_p \neq L_q] + \sum_{L \in \Lambda} h_L \cdot \delta_L(\mathbf{L})$$

If **multiple** models

- assign different models (labels L_p) to every point p

- find optimal **labeling**

$\boldsymbol{L} = \{ L_1, L_2, \ldots, L_n \}$



Practical problem: number of potential labels (models) is huge, how can we use a-expansion designed for a finite set of labels?

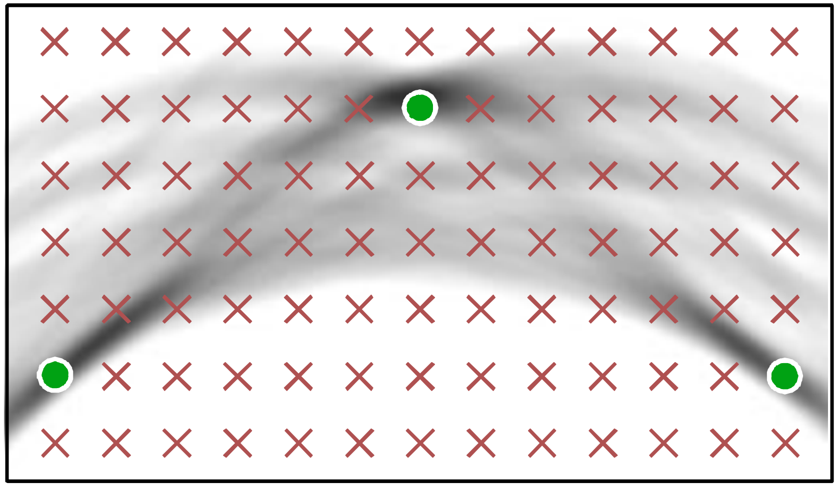
Discrete optimization for continuum of labels?

example: line detection

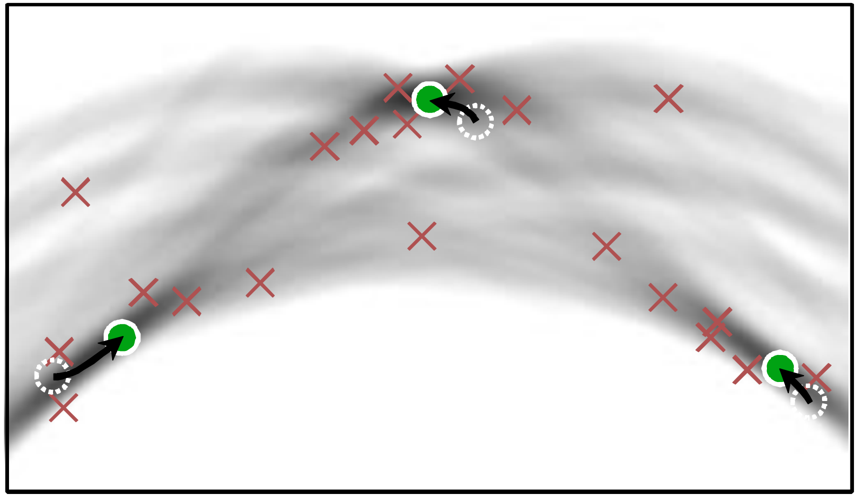
Discrete optimization for continuum of labels?

example: line detection

Hough transform (that is, **space of lines**)



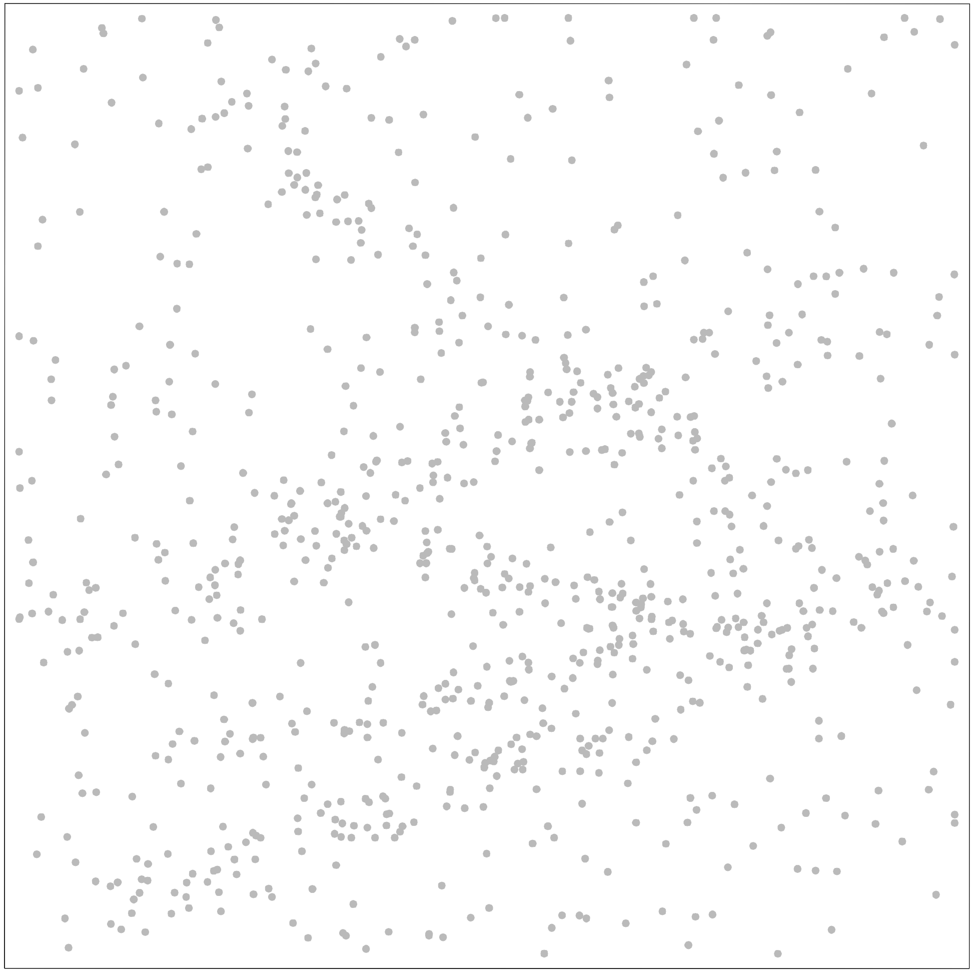
Uniform discretization of label space



Adaptive exploration of label space (PEARL)

PEARL

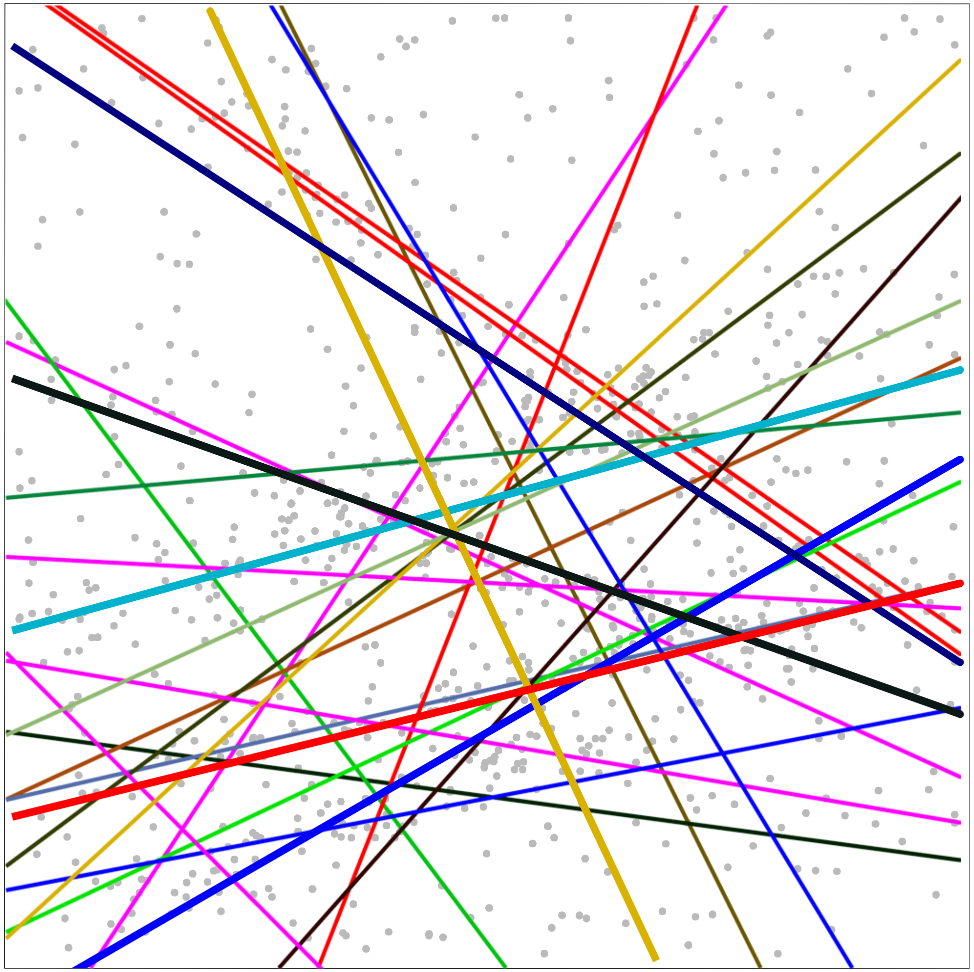
Propose
Expand
And
Reestimate
Labels



data points

PEARL

Propose
Expand
And
Reestimate
Labels



data points + randomly sampled models

sample data to generate a **finite set of initial labels**

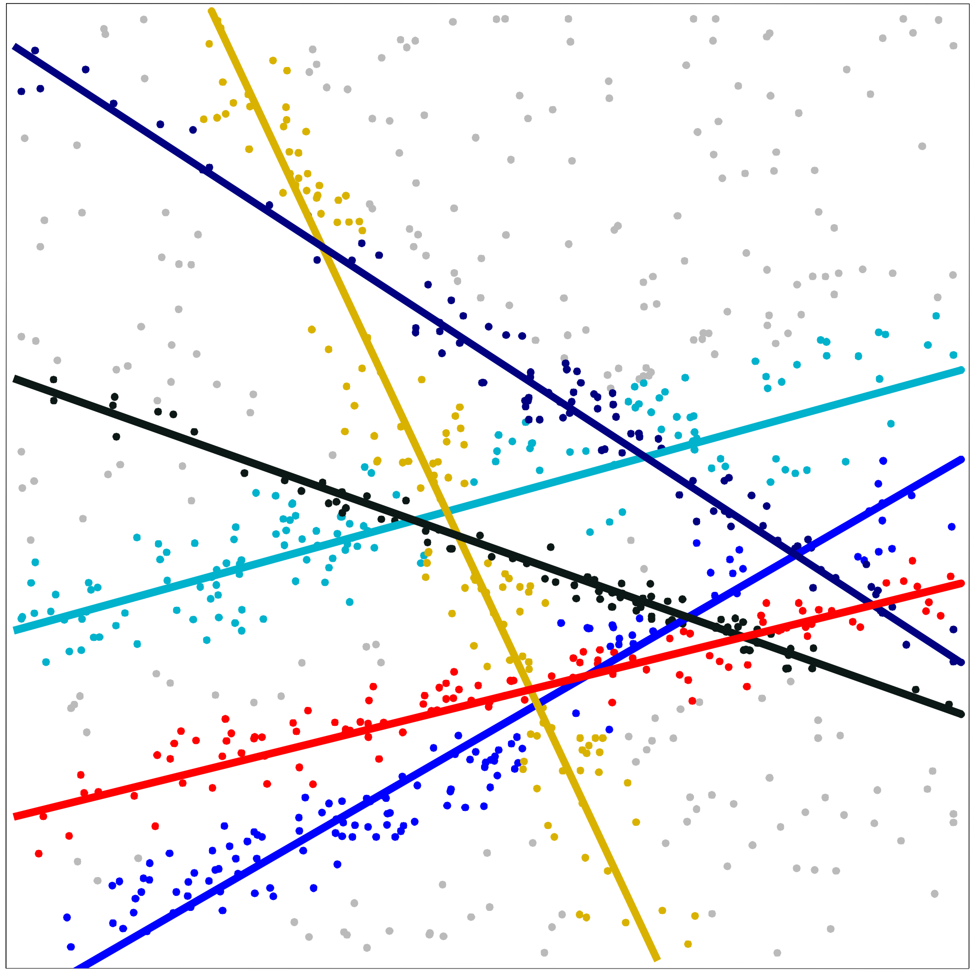
Λ

$$E(\mathbf{L}) = \sum_{p} ||p - L_p|| + \sum_{pq\in N} w\cdot[L_p \neq L_q] + \sum_{L\in\Lambda} h_L \cdot \delta_L[\mathbf{L}]$$

PEARL

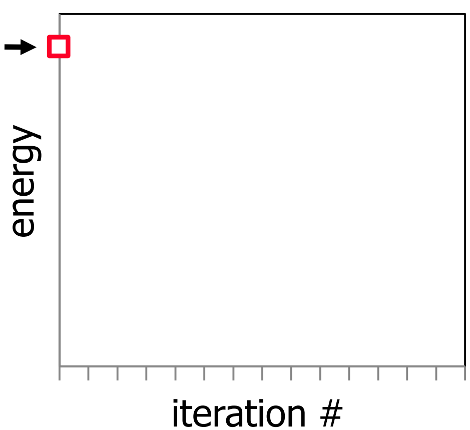
$$L_p \in \Lambda$$

Propose
Expand
And
Reestimate
Labels



iteration 1: optimize labeling L

a-expansion:
minimize $E(\mathbf{L})$
over a fixed
set of labels
Λ

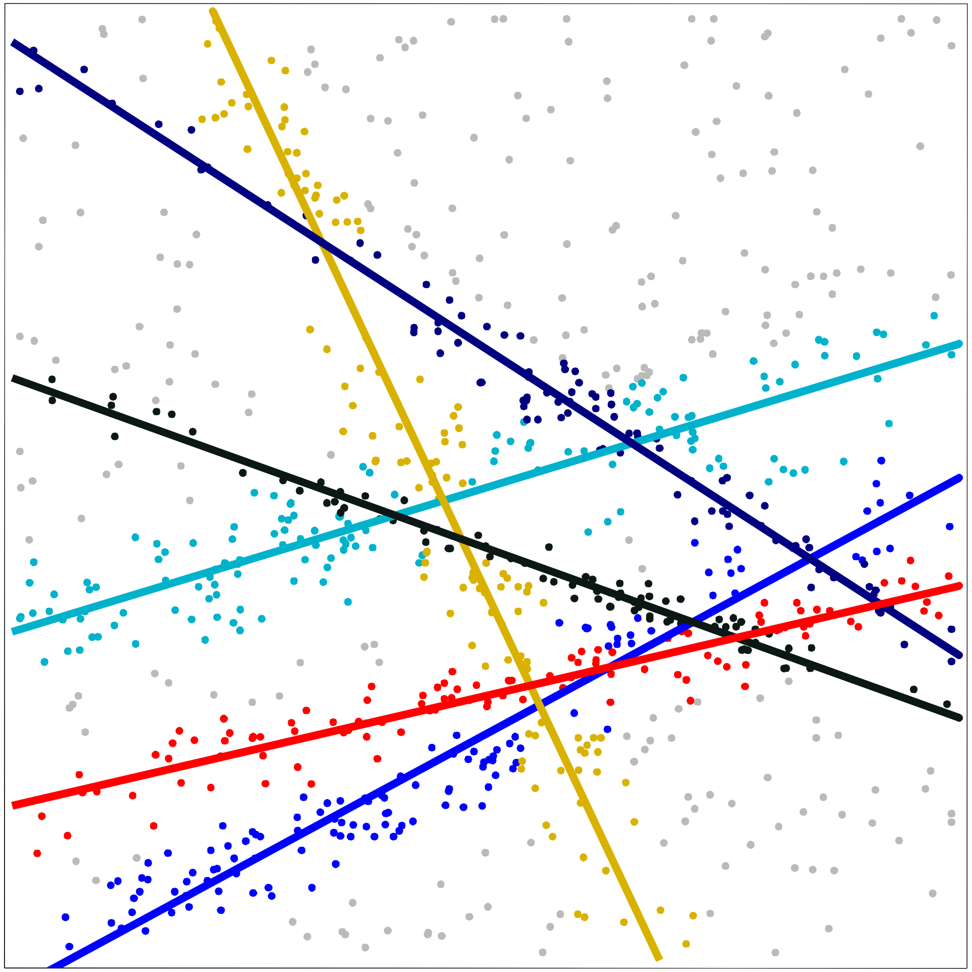


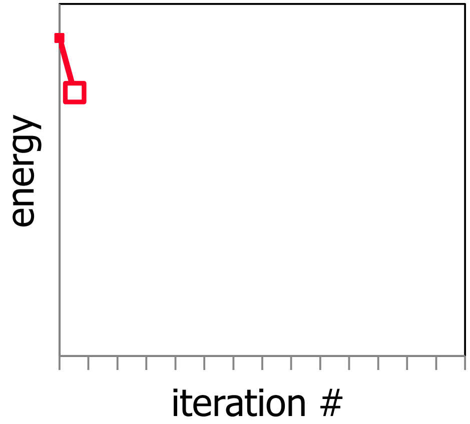
$$E(\mathbf{L}) = \sum_p ||p - L_p|| + \sum_{pq\in N} w\cdot[L_p \neq L_q] + \sum_{L\in\Lambda} h_L \cdot \delta_L[\mathbf{L}]$$

fixed

PEARL

Propose
Expand
And
Reestimate
Labels





iteration 1: reestimate models

reestimating labels in Λ for given inliers

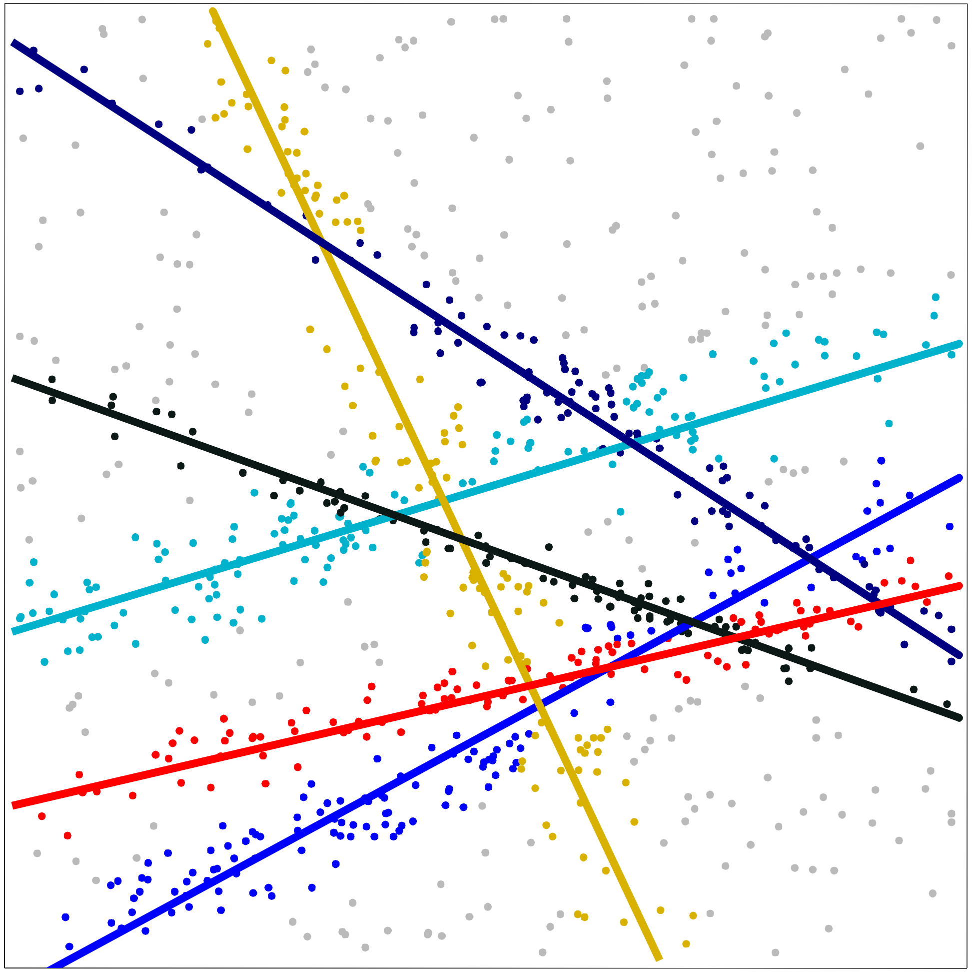
minimizing the first term of energy $E(L)$

$$E(\mathbf{L}) = \sum_{p} ||p - L_p|| + \sum_{pq \in N} w \cdot [L_p \neq L_q] + \sum_{L \in \Lambda} h_L \cdot \delta_L[\mathbf{L}]$$

PEARL

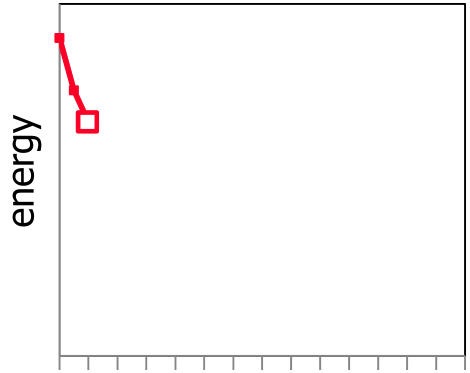
$$L_p \in \Lambda$$

Propose
Expand
And
Reestimate
Labels



iteration 2: optimize labeling L

a-expansion:
minimize $E(\mathbf{L})$ over a fixed set of labels Λ

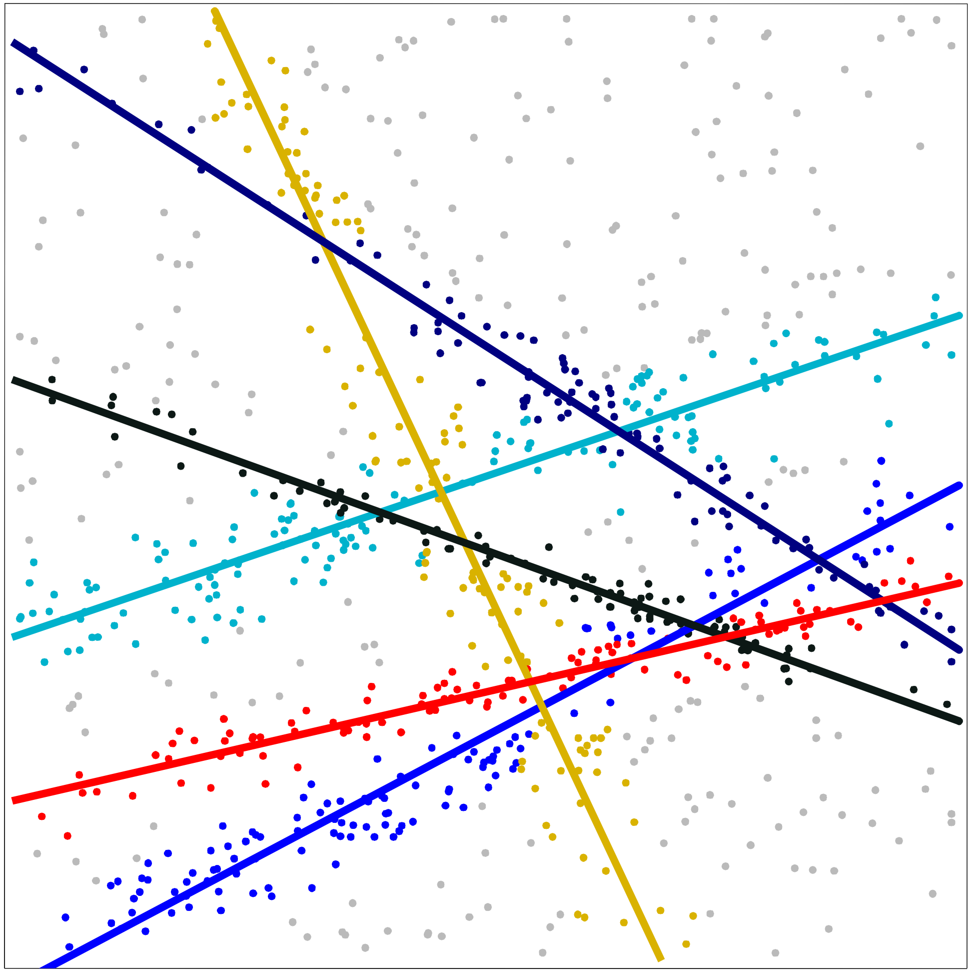


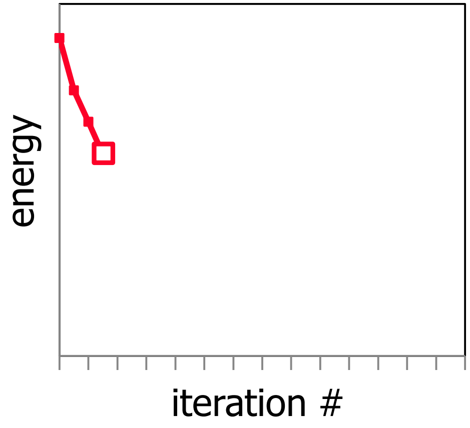
$$E(\mathbf{L}) = \sum_{p} ||p - L_p|| + \sum_{pq \in N} w \cdot [L_p \neq L_q] + \sum_{L \in \Lambda} h_L \cdot \delta_L[\mathbf{L}]$$

fixed

PEARL

Propose
Expand
And
Reestimate
Labels





reestimating labels in for given inliers

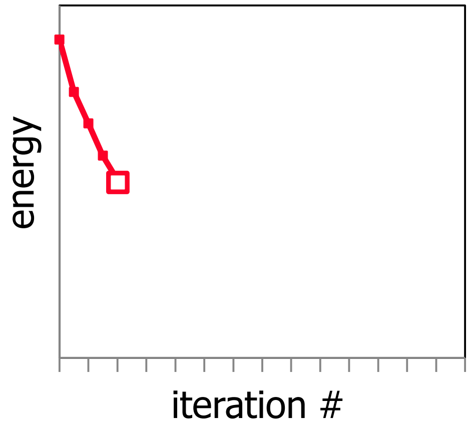
minimizing the first term of energy $E(L)$

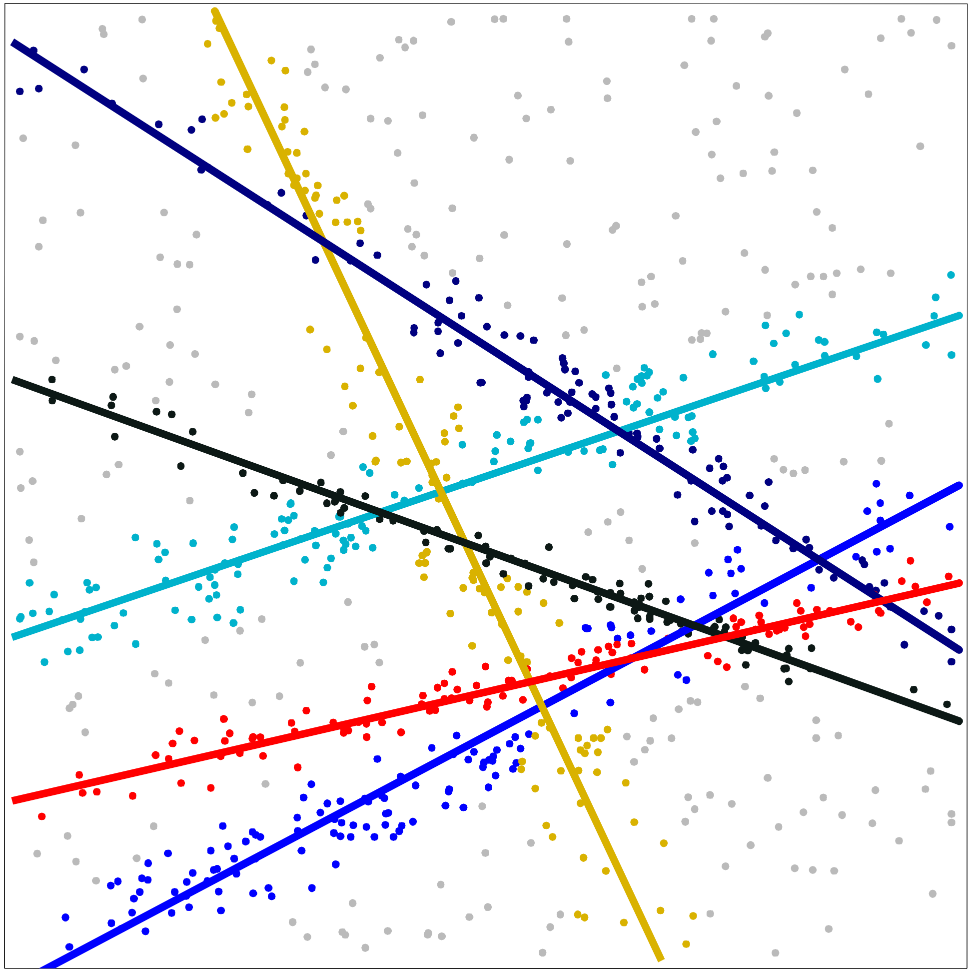
iteration 2: reestimate models

$$E(\mathbf{L}) = \sum_{p} ||p - L_p|| + \sum_{pq \in N} w \cdot [L_p \neq L_q] + \sum_{L \in \Lambda} h_L \cdot \delta_L[\mathbf{L}]$$

PEARL

Propose
Expand
And
Reestimate
Labels



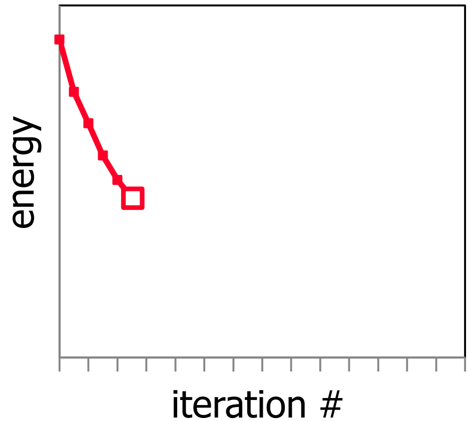


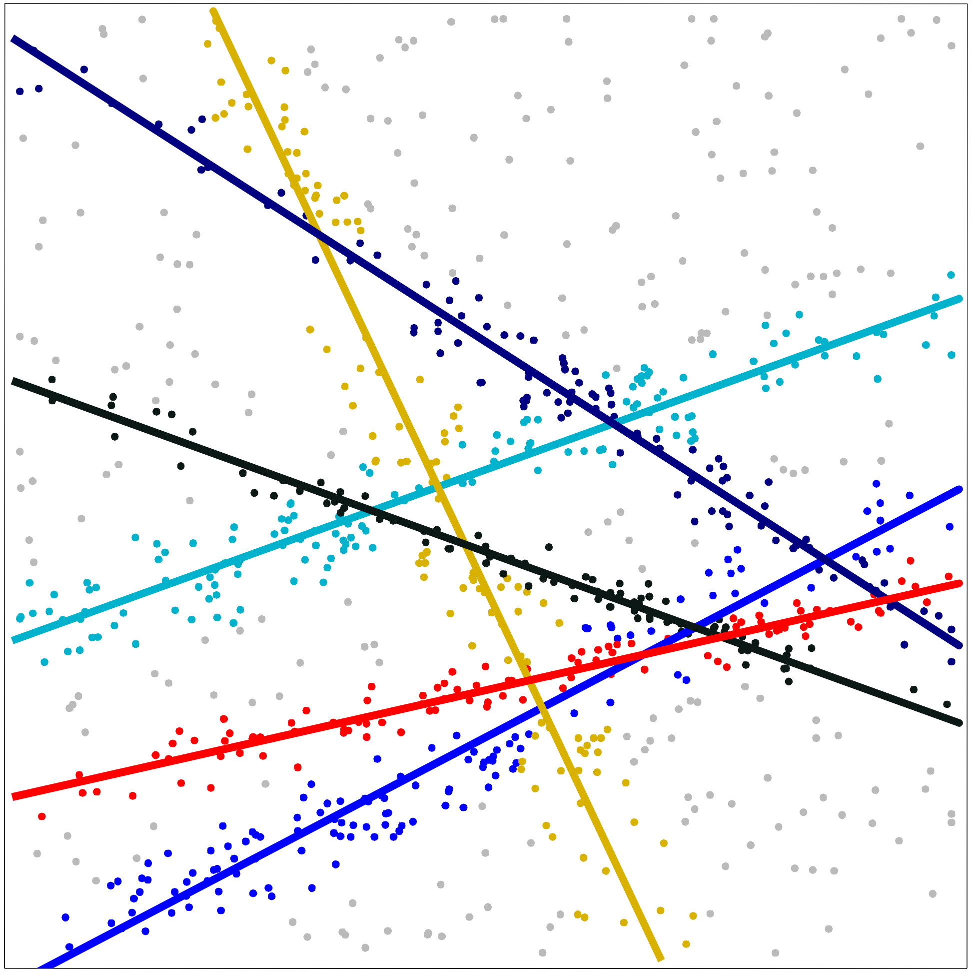
iteration 3: optimize labeling *L*

$$E(\mathbf{L}) = \sum_{p} ||p - L_p|| + \sum_{pq \in N} w \cdot [L_p \neq L_q] + \sum_{L \in \Lambda} h_L \cdot \delta_L[\mathbf{L}]$$

PEARL

Propose
Expand
And
Reestimate
Labels

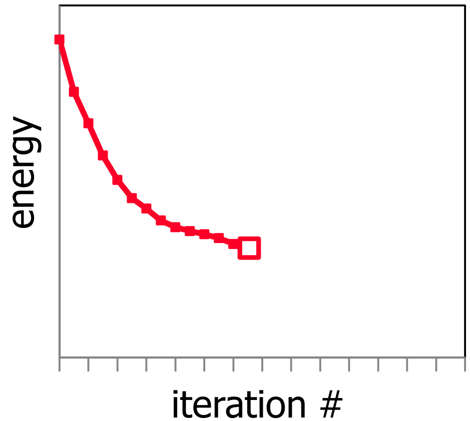


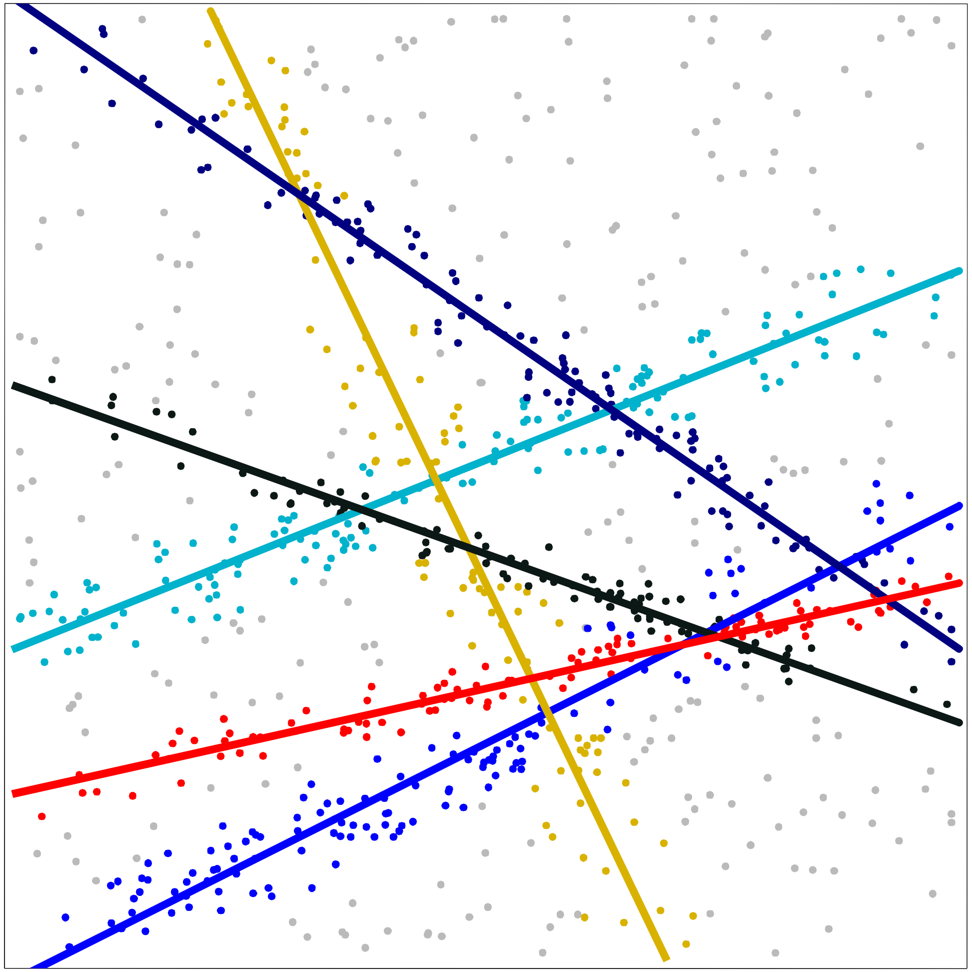


iteration 3: reestimate models

$$E(\mathbf{L}) = \sum_{p} || p - L_p || + \sum_{pq \in N} w \cdot [L_p \neq L_q] + \sum_{L \in \Lambda} h_L \cdot \delta_L[\mathbf{L}]$$

PEARL

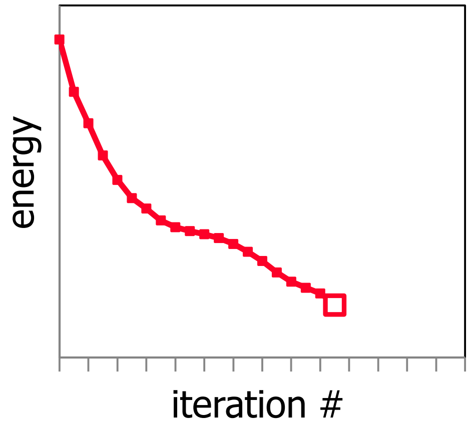


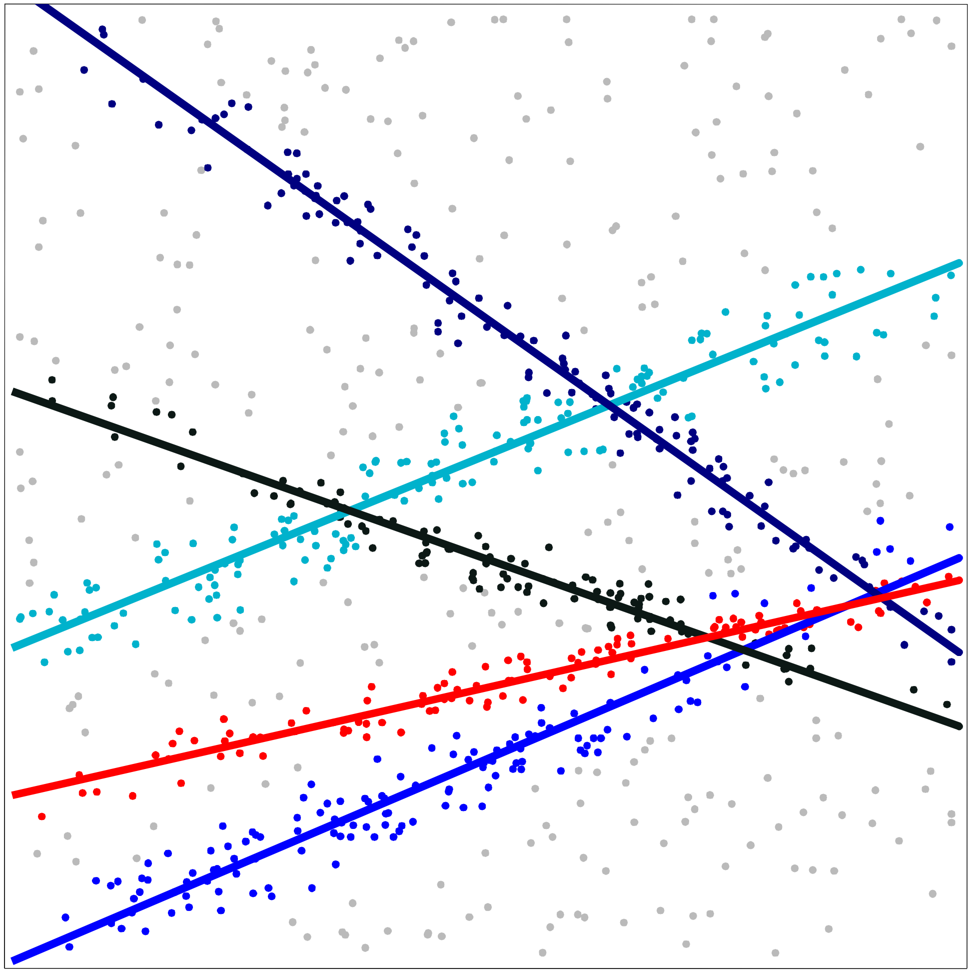


iteration 7...

$$E(\mathbf{L}) = \sum_{p} ||\, p - L_p \,|| + \sum_{pq \in N} w \cdot [L_p \neq L_q] + \sum_{L \in \Lambda} h_L \cdot \delta_L[\mathbf{L}]$$

PEARL

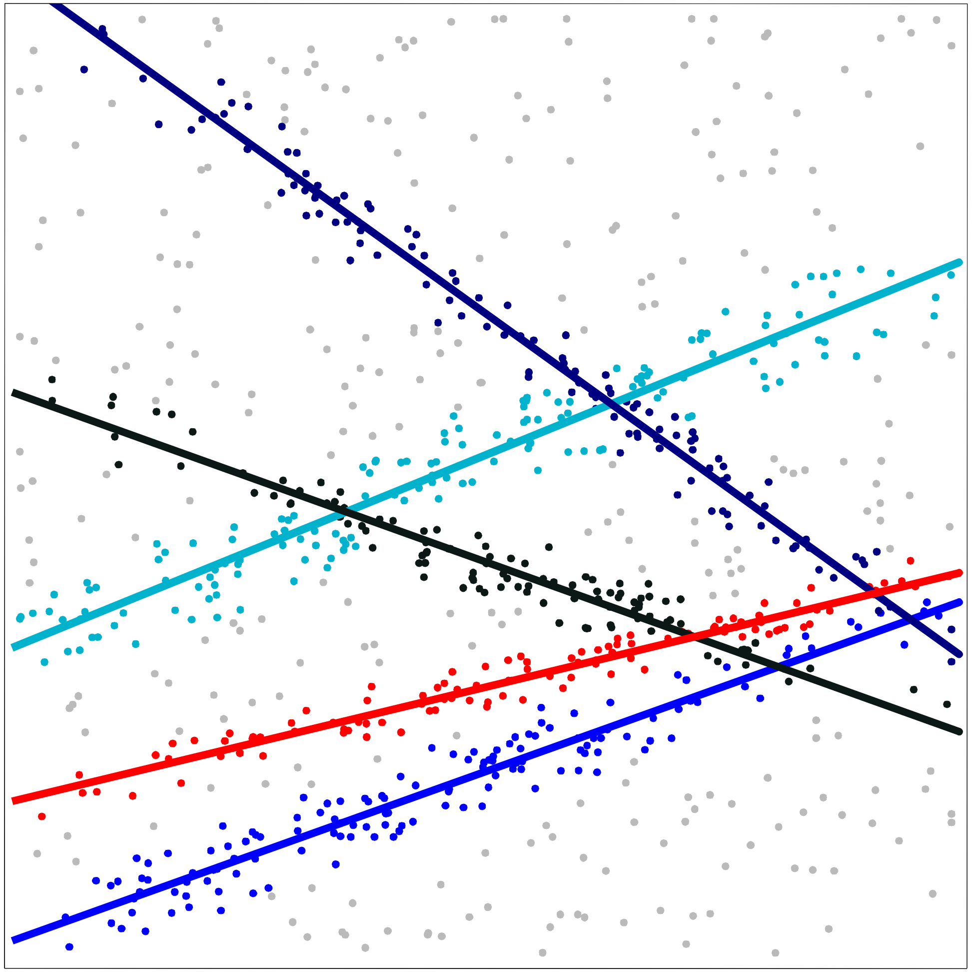


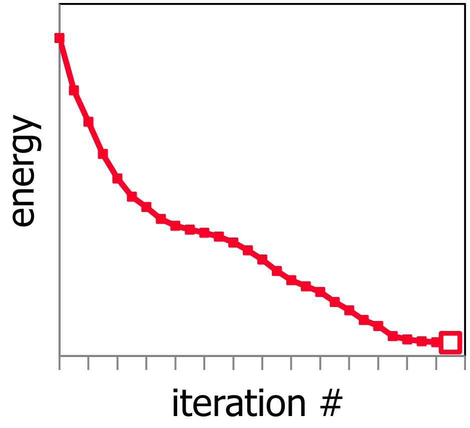


iteration 10...

$$E(\boldsymbol{L}) = \sum_{p} ||p - L_p|| + \sum_{pq \in N} w \cdot [L_p \neq L_q] + \sum_{L \in \Lambda} h_L \cdot \delta_L[\boldsymbol{L}]$$

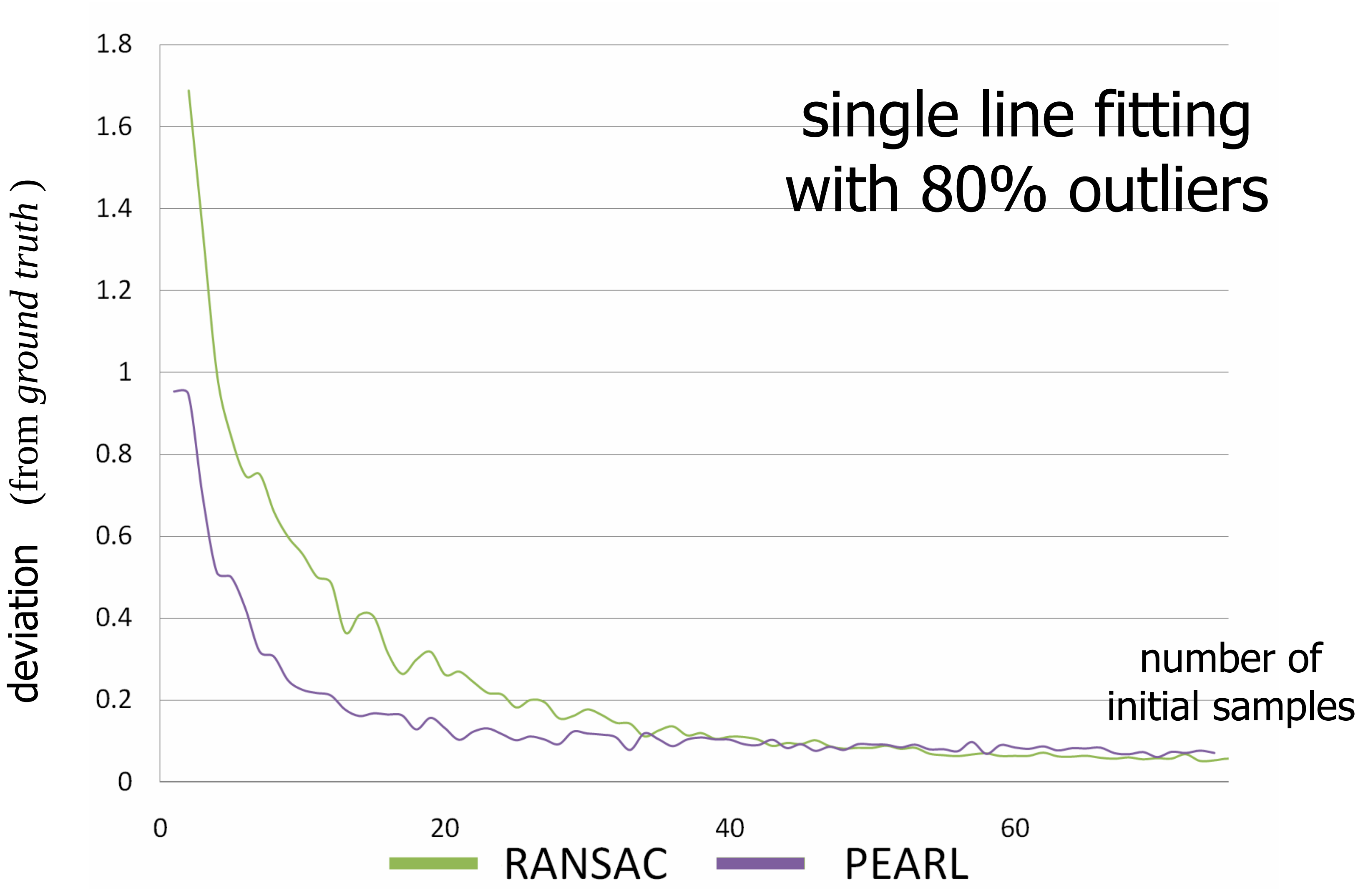
PEARL



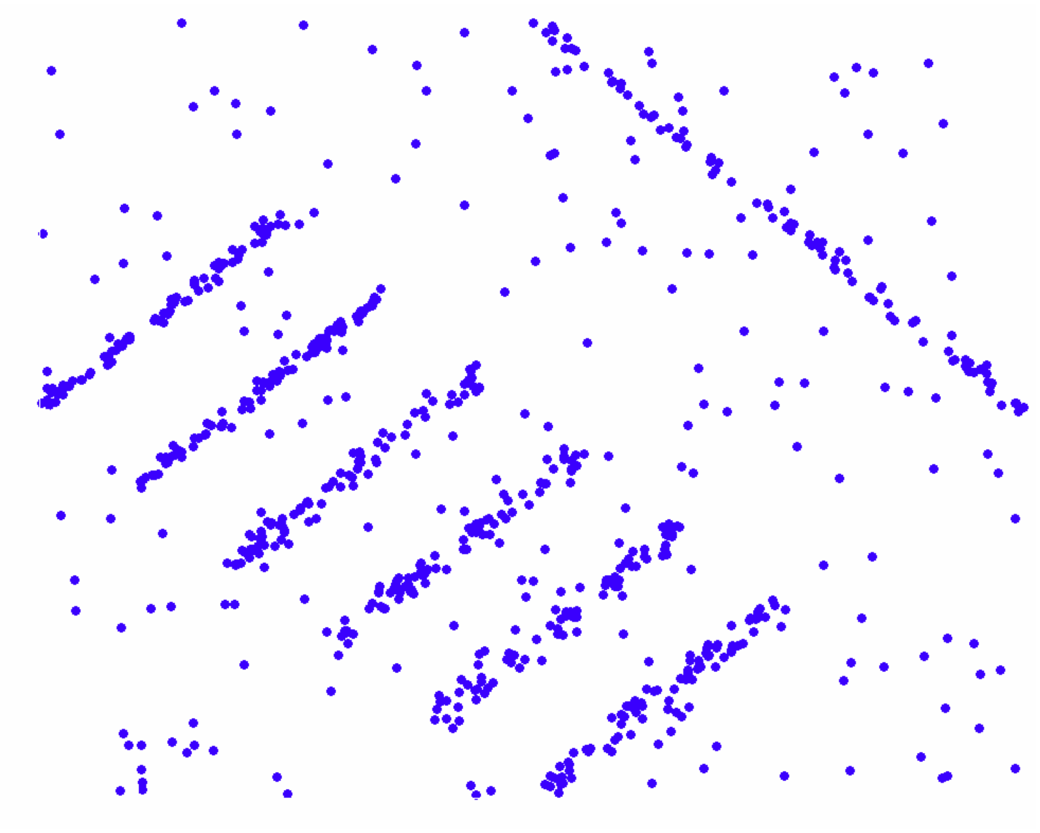


iteration 15... converged.

PEARL can significantly improve initial models



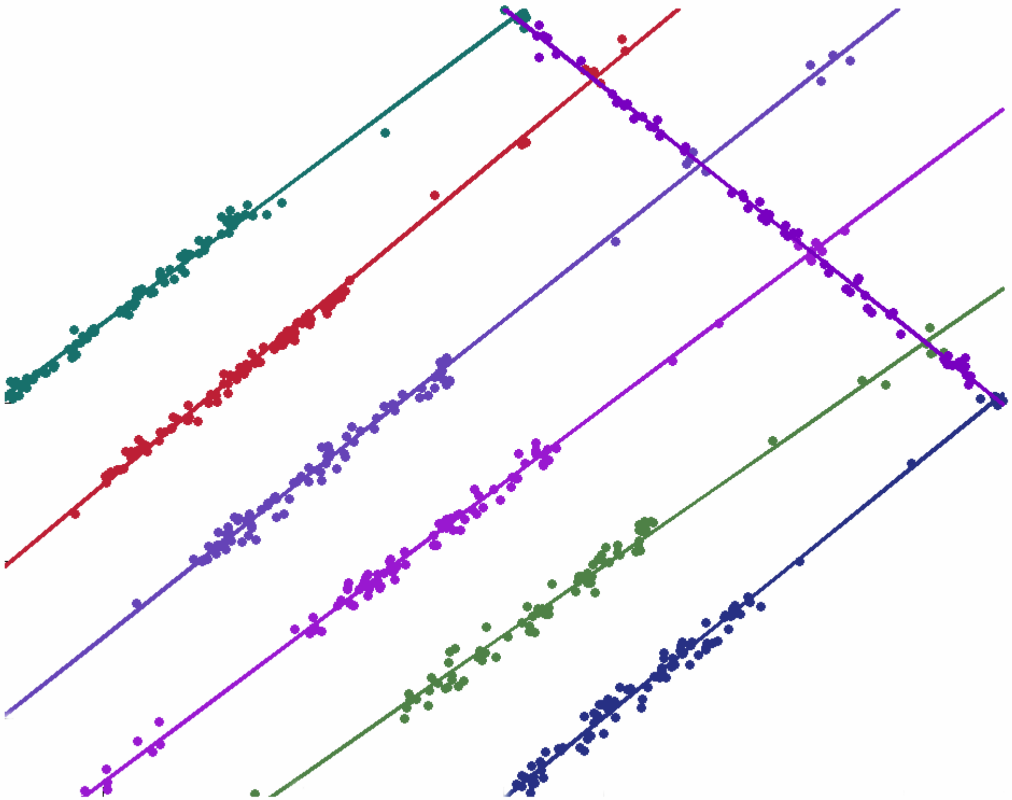
Comparison for multi-model fitting



Low noise

original data points

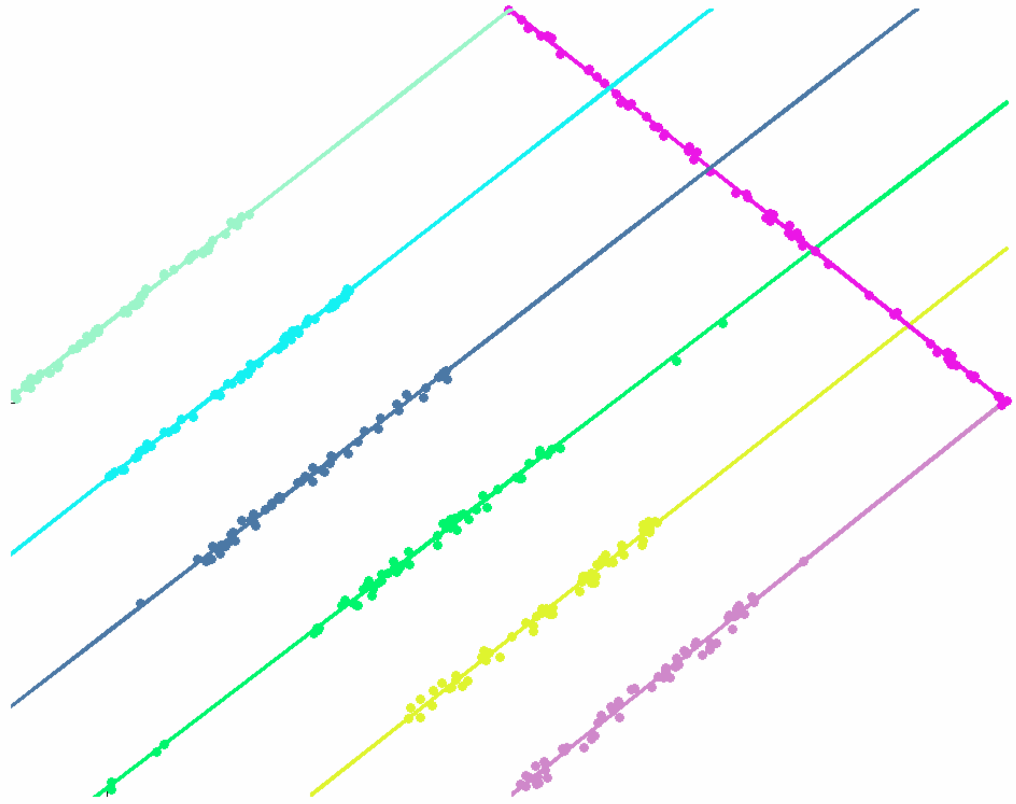
Comparison for multi-model fitting



Low noise

sequential RANSAC

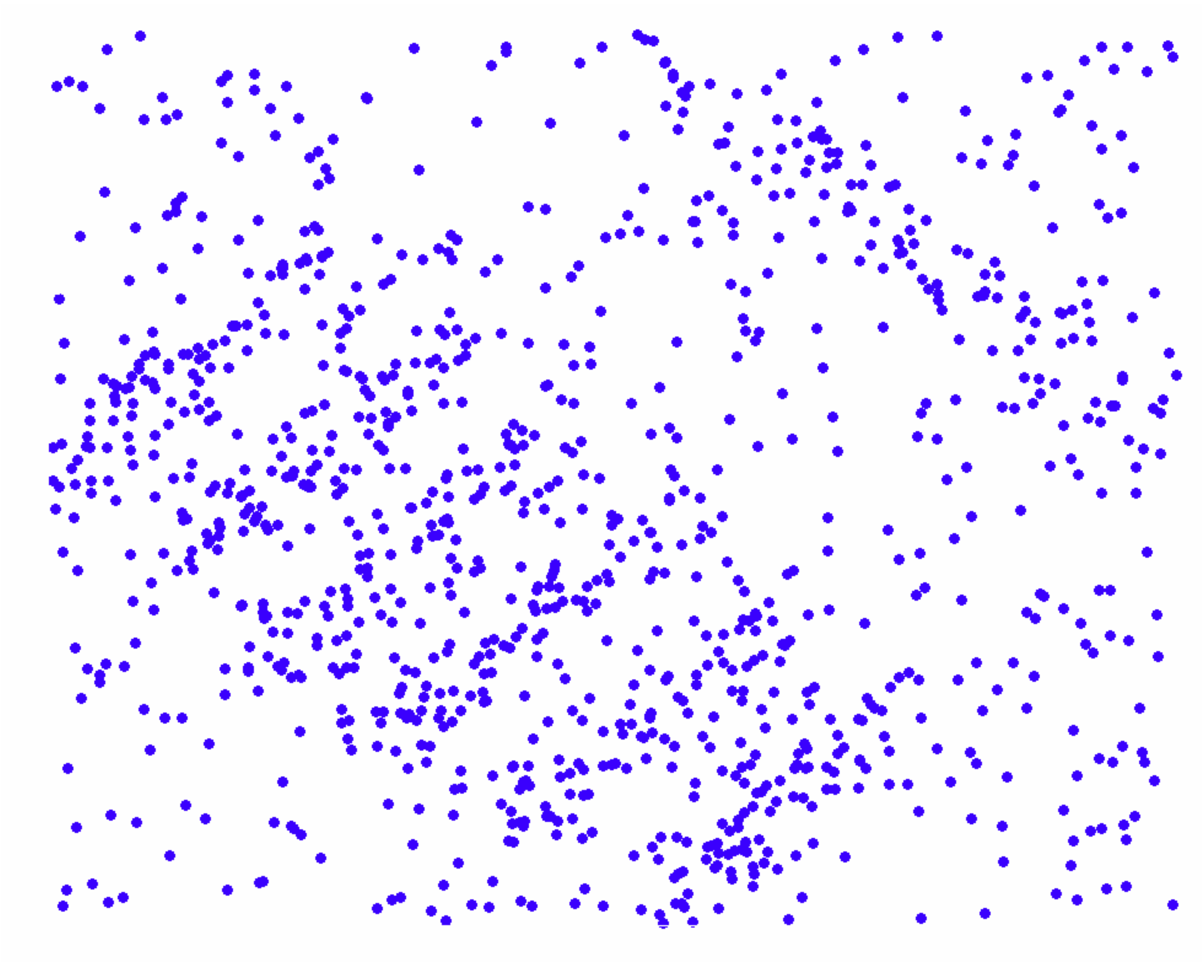
Comparison for multi-model fitting



Low noise

PEARL

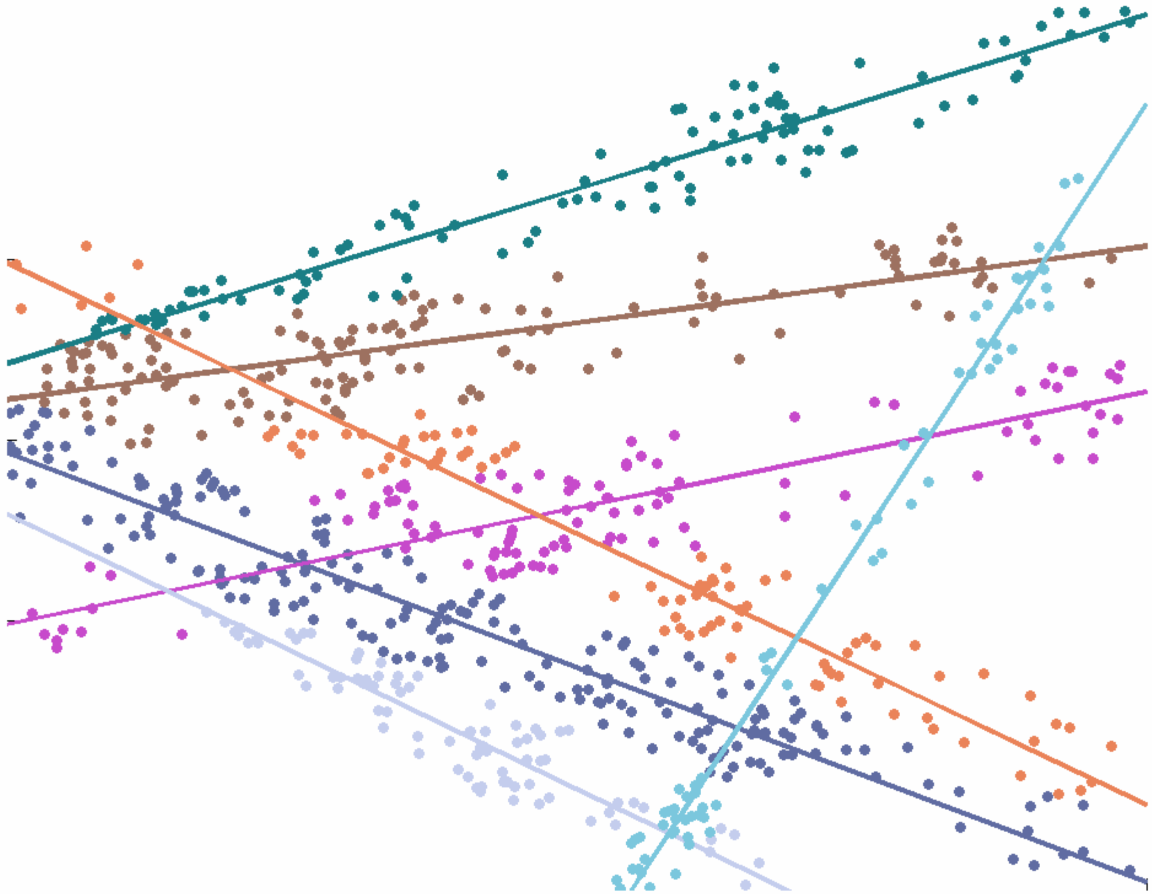
Comparison for multi-model fitting



original data points

High noise

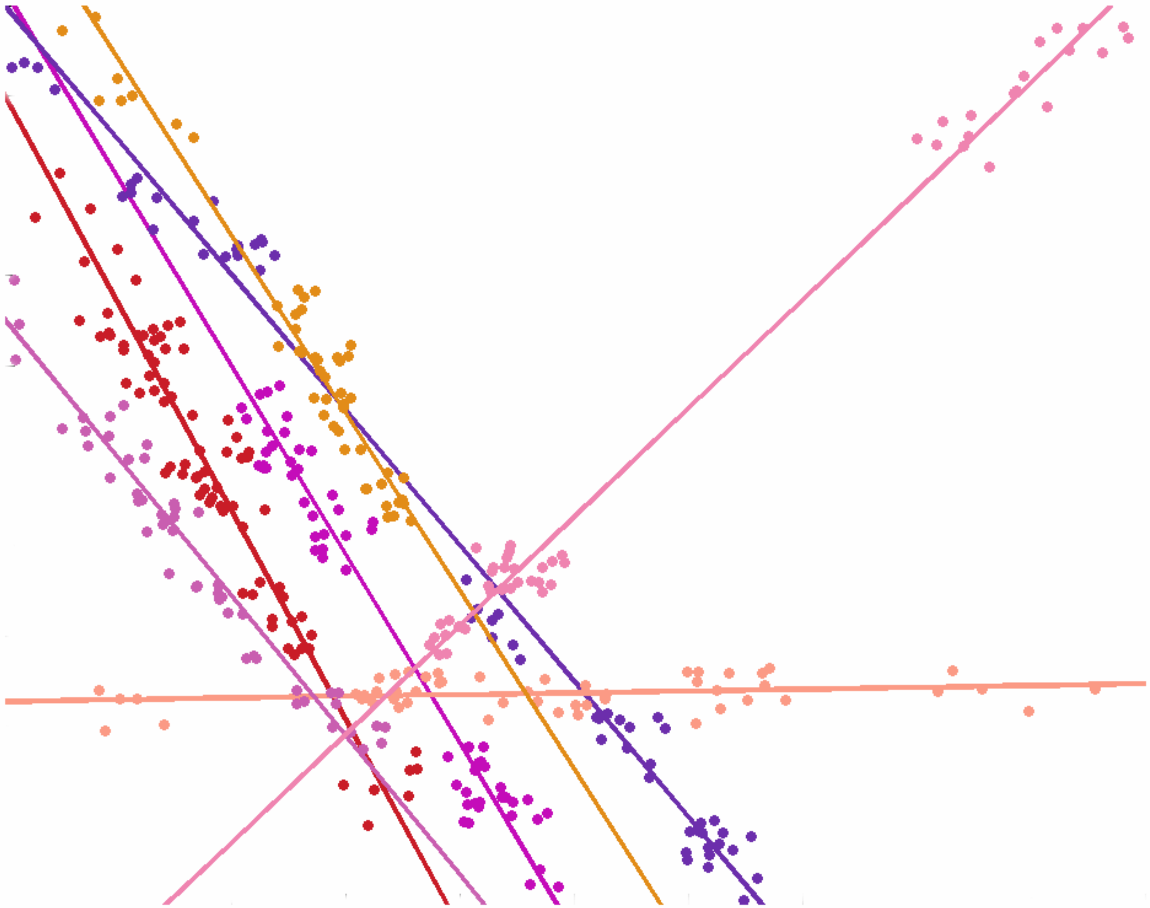
Comparison for multi-model fitting



High noise

sequential RANSAC

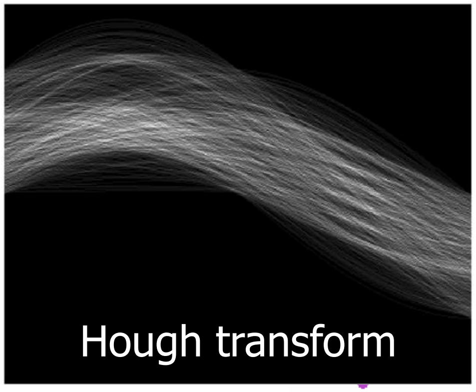
Comparison for multi-model fitting

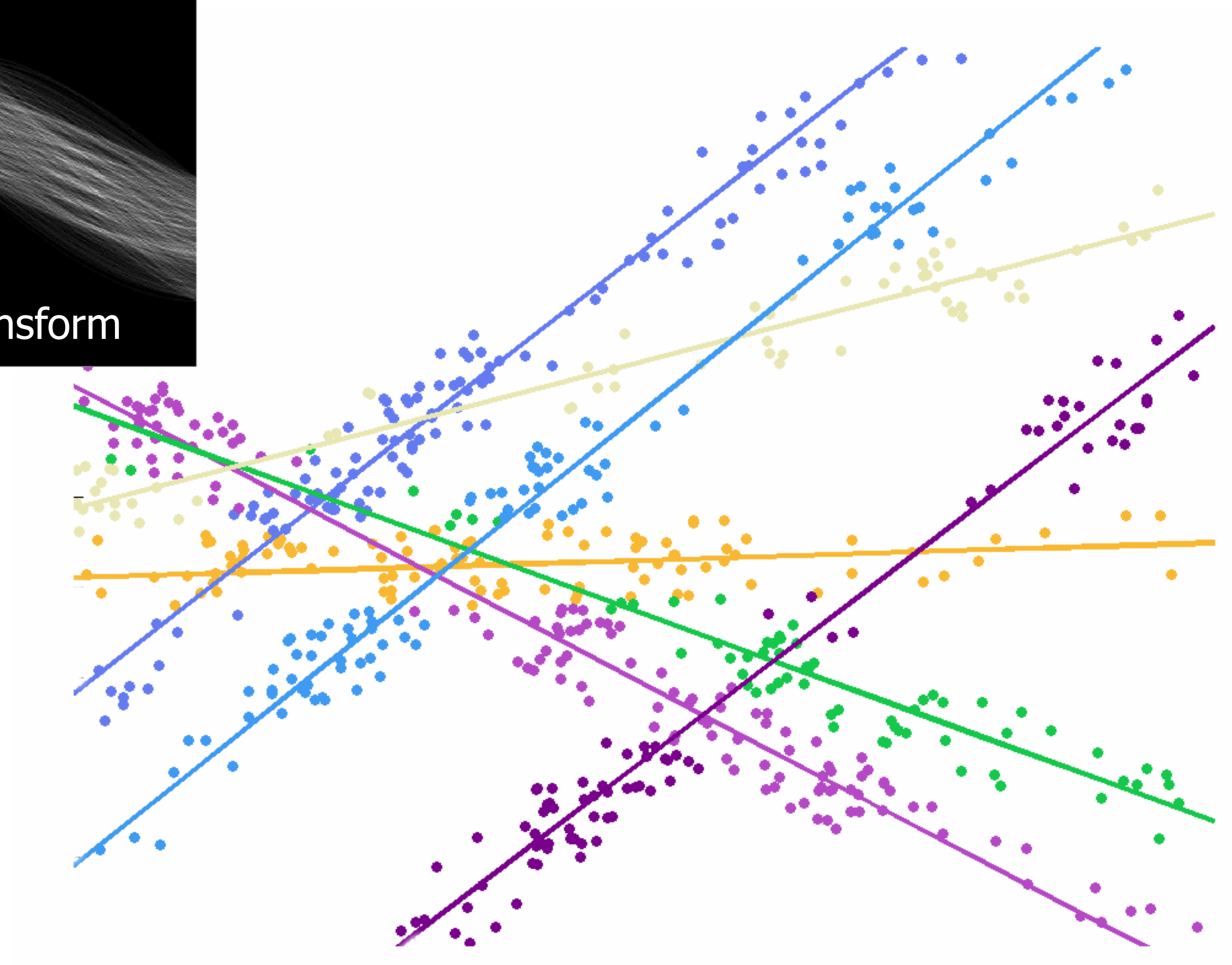


High noise

Other generalization of RANSAC (J-linkage, Toldo & Fusiello, ECCV'08)

Comparison for multi-model fitting

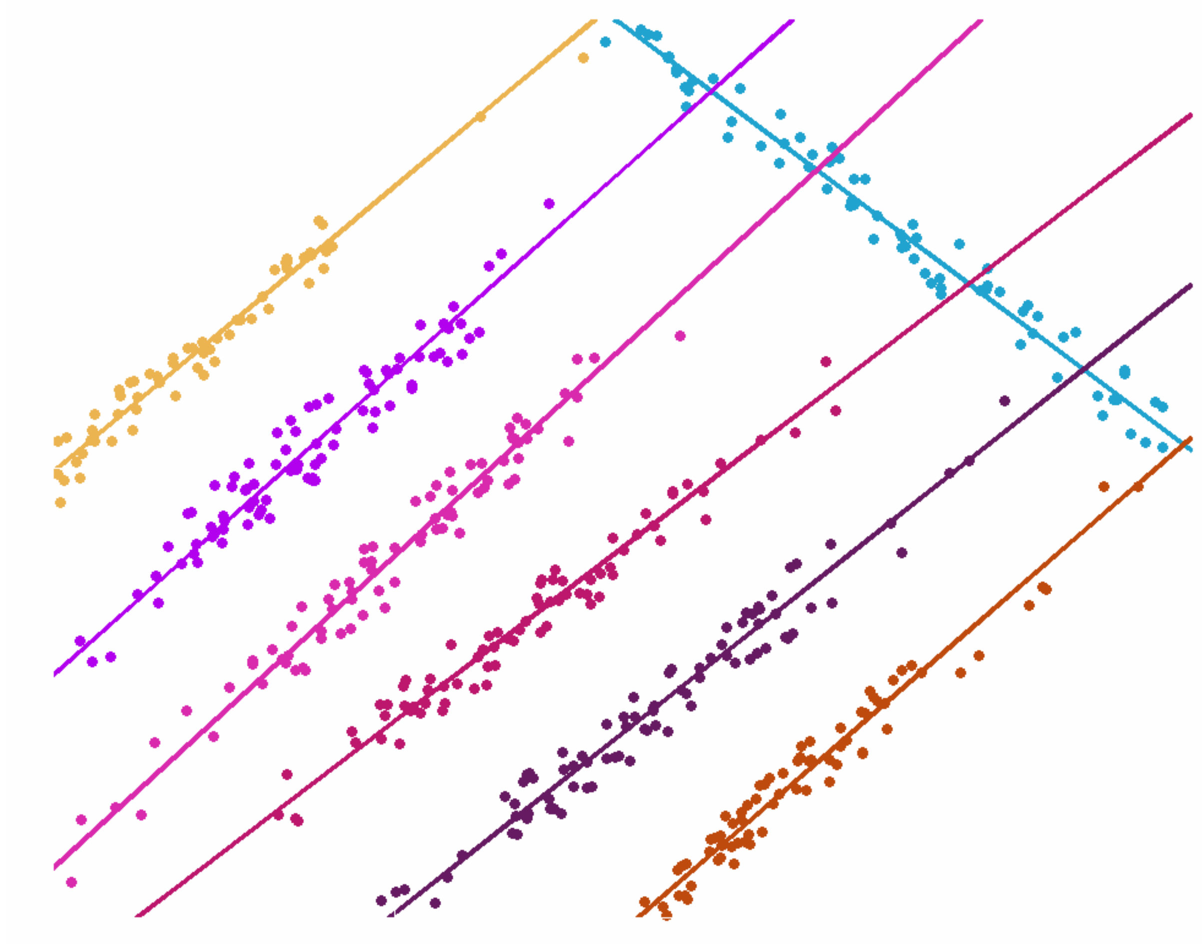




High noise

Finding modes in Hough-space, e.g. via **mean-shift**
(also maximizes the number of inliers)

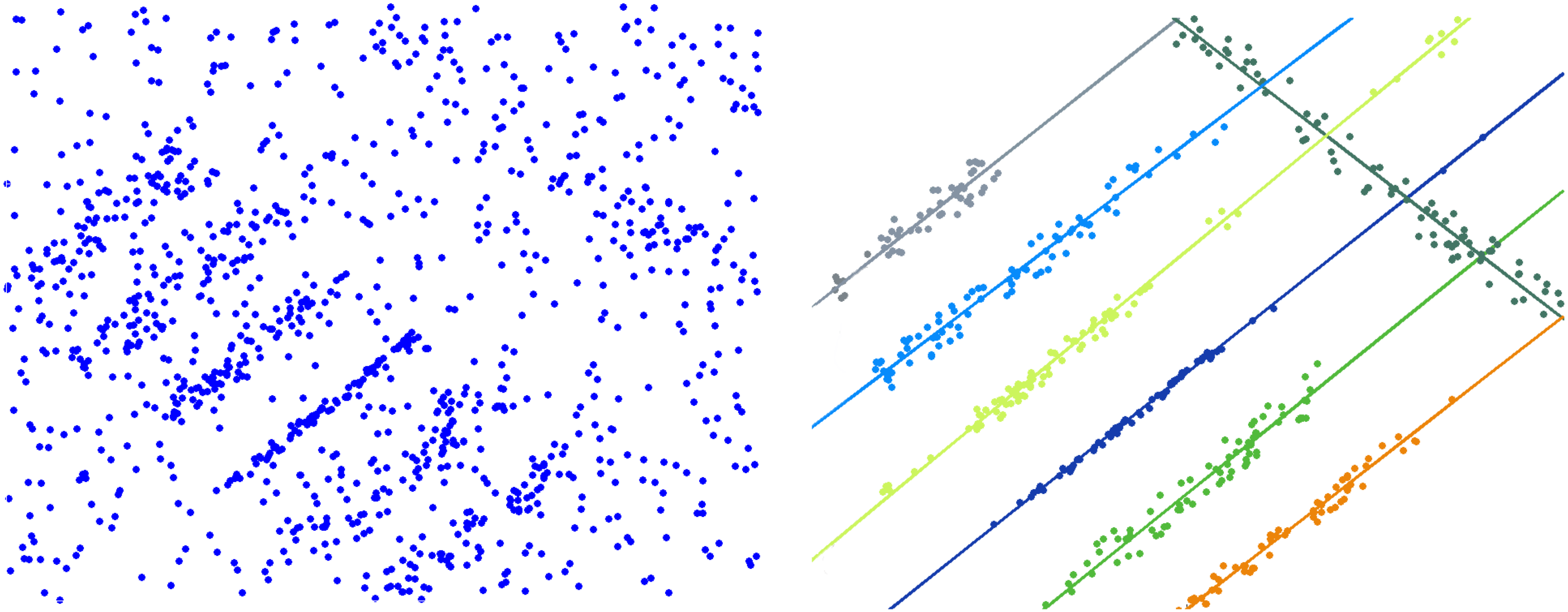
Comparison for multi-model fitting



High noise

PEARL

Automatic noise level estimation by fitting models $L=(a,b,\sigma)$



various level of noise

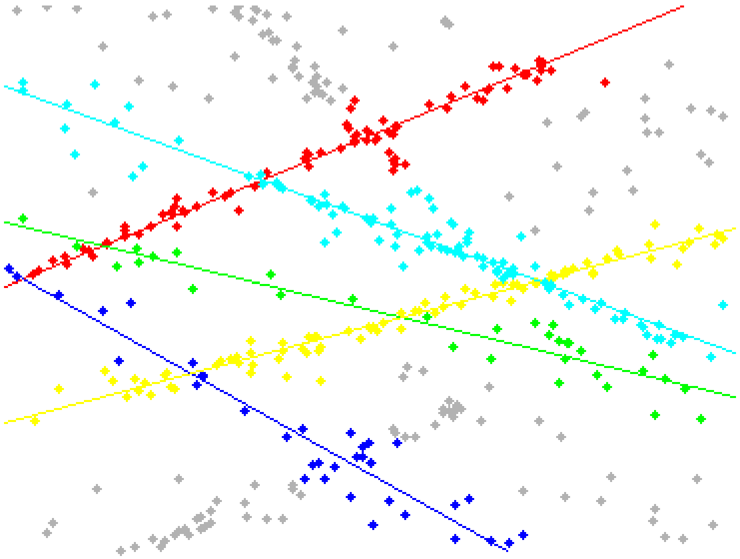
our result

Each model L_k gets its own σ_κ

K-means vs. PEARL

$$E(\mathbf{L}) = \sum_{p} \| p - L_p \| + \text{hard constraint on number of models}$$

K-means



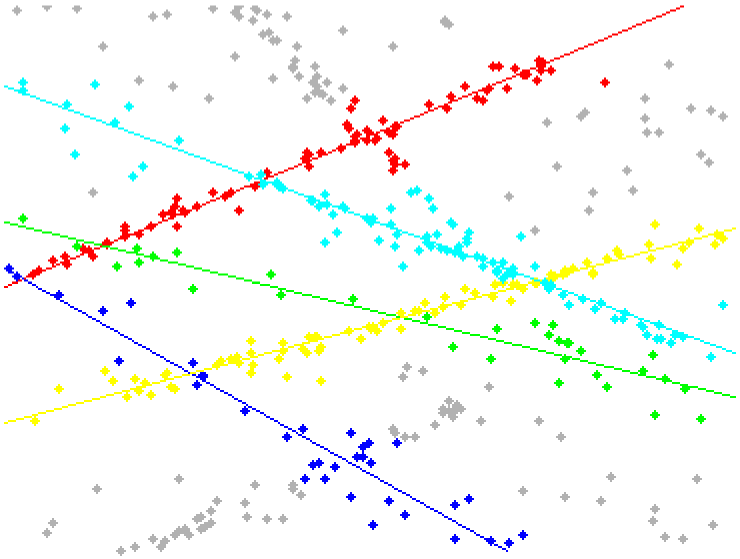
5 random initial lines + outlier model

gets stuck in local minima

K-means vs. PEARL

$$E(\mathbf{L}) = \sum_p \| p - L_p \| + \sum_{L \in \Lambda} h_L \cdot \delta_L(\mathbf{L})$$

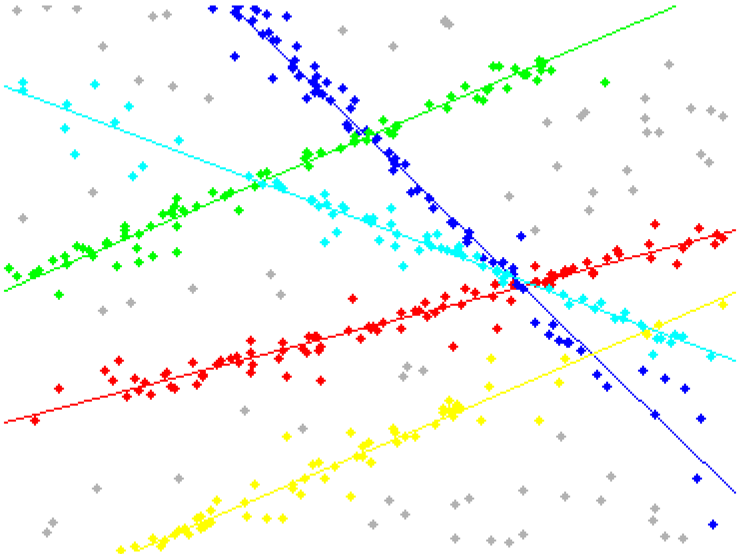
K-means



5 random initial lines + outlier model

gets stuck in local minima

PEARL $h_L = 1000$



1000 initial lines + outlier model

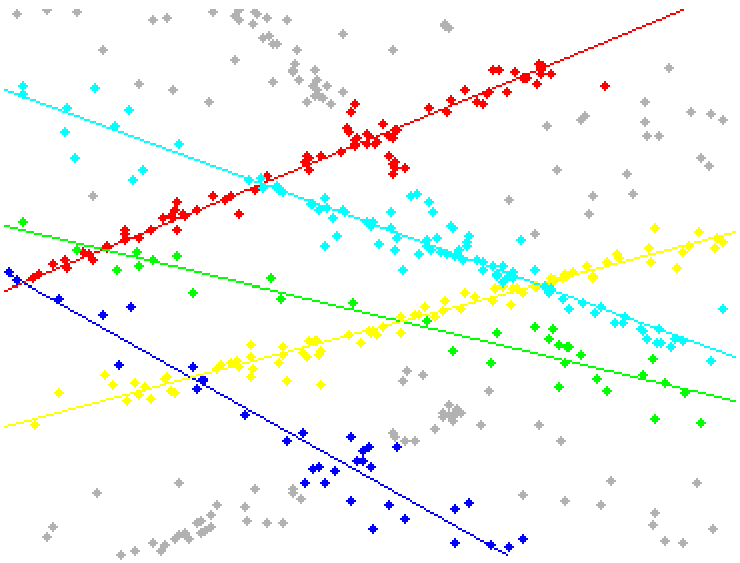
better explores label space

K-means vs. PEARL

$$E(\mathbf{L}) = \sum_p \| p - L_p \| + \sum_{L \in \Lambda} h_L \cdot \delta_L(\mathbf{L})$$

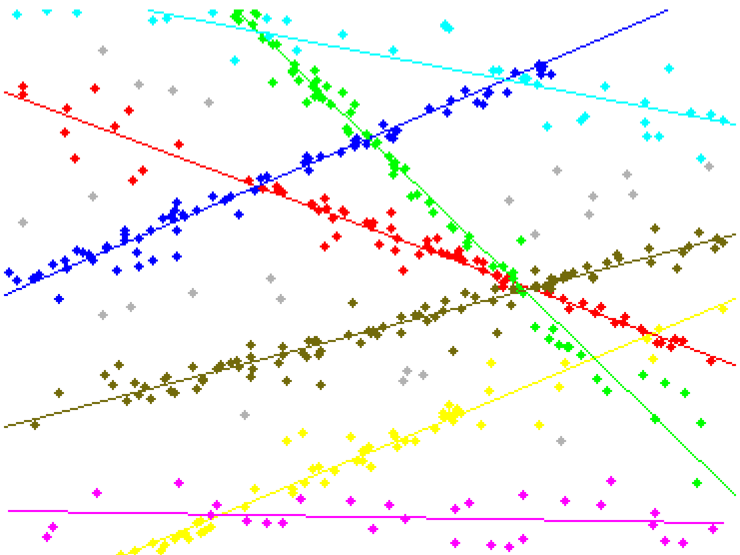
K-means

PEARL $h_L = 500$



5 random initial lines + outlier model

gets stuck in local minima



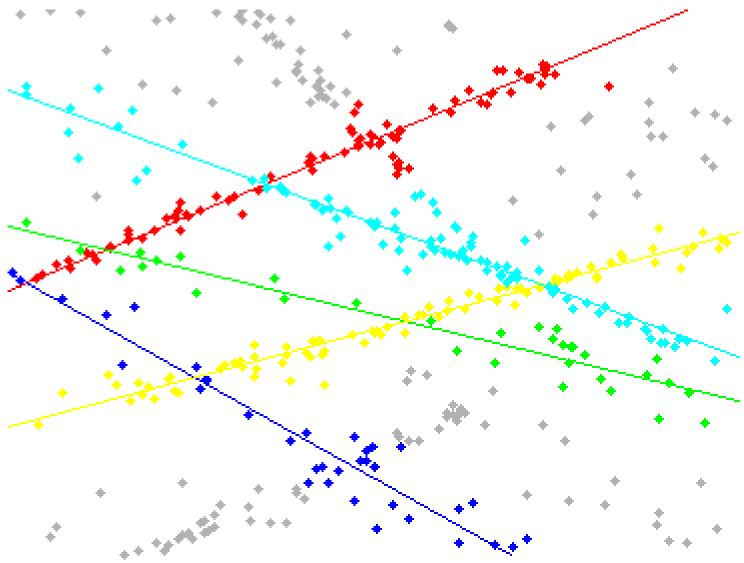
1000 initial lines + outlier model

better explores label space

K-means vs. PEARL

$$E(\mathbf{L}) = \sum_p \| p - L_p \| + \sum_{L \in \Lambda} h_L \cdot \delta_L(\mathbf{L})$$

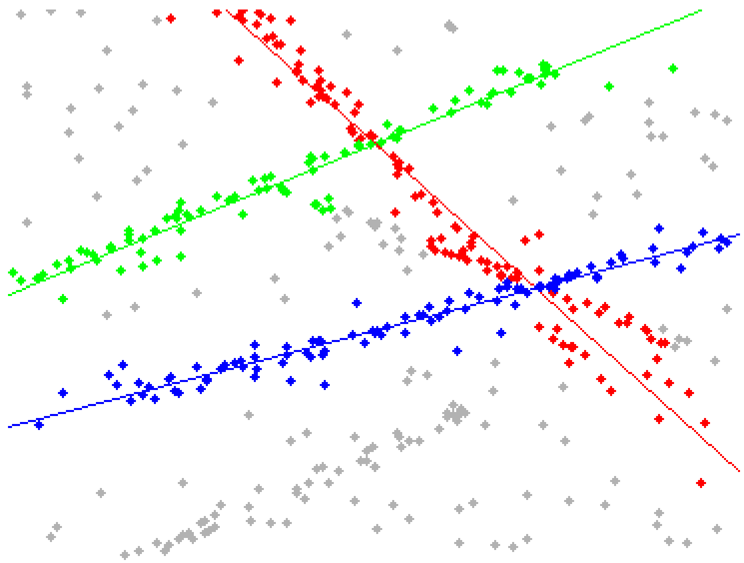
K-means



5 random initial lines + outlier model

gets stuck in local minima

PEARL $h_L = 2000$

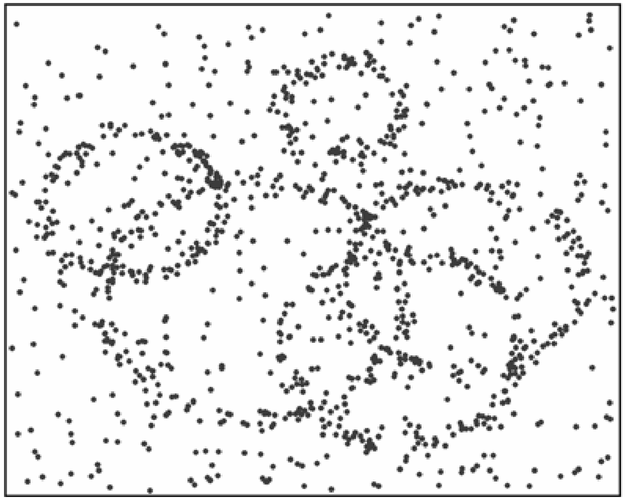


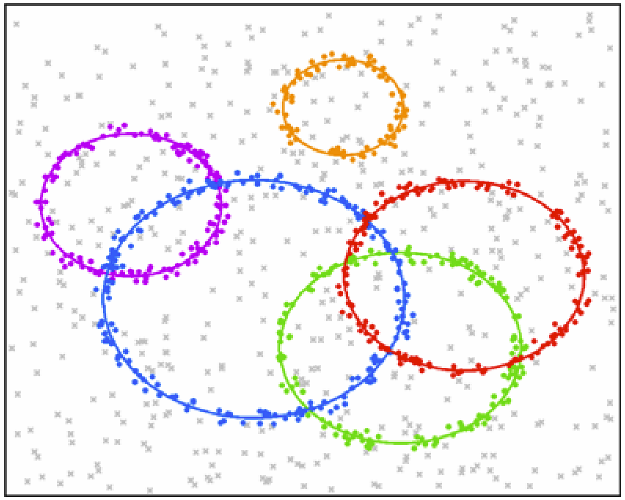
1000 initial lines + outlier model

better explores label space

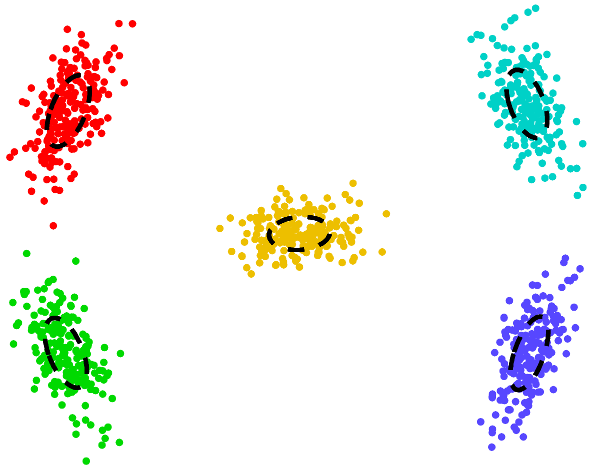
Fitting circles

$$E(\mathbf{L}) = \sum_p \| p - L_p \| + \sum_{L \in \Lambda} h_L \cdot \delta_L(\mathbf{L})$$

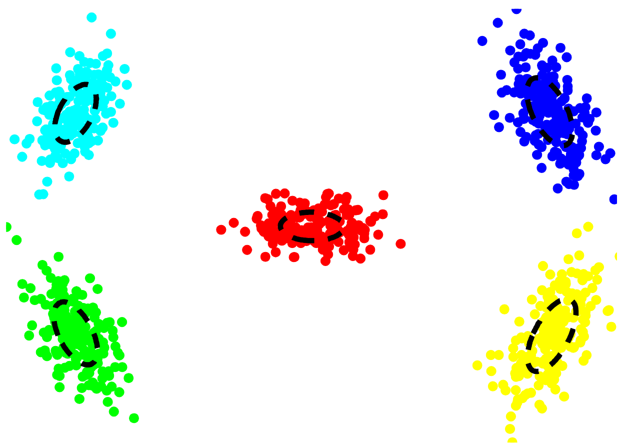




EM vs K-means

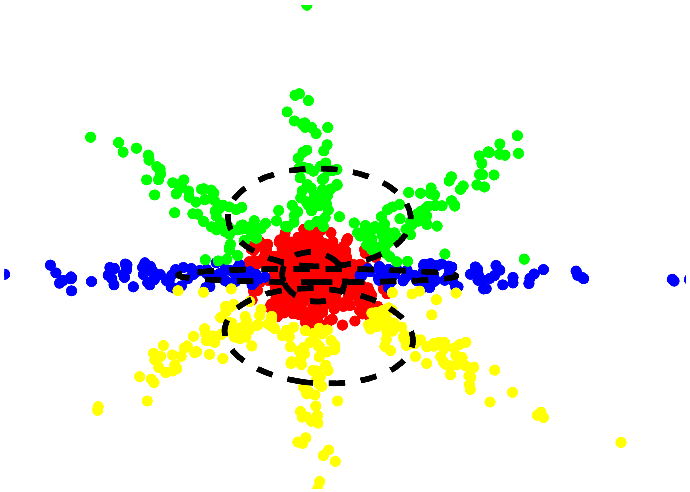


EM, with 5 models

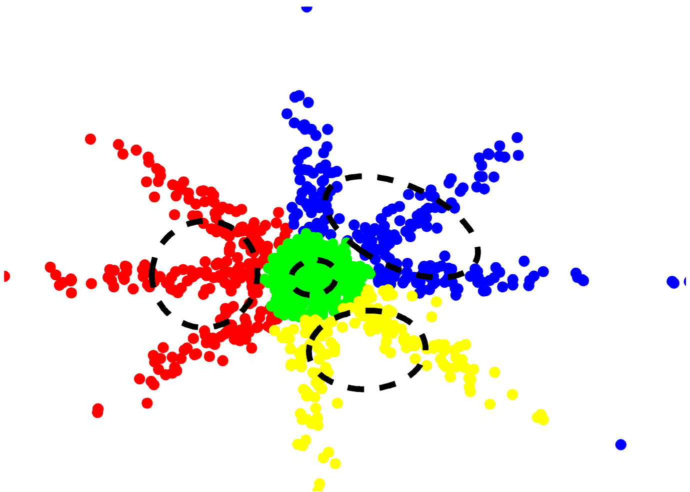


K-means, with 5 models

EM vs K-means

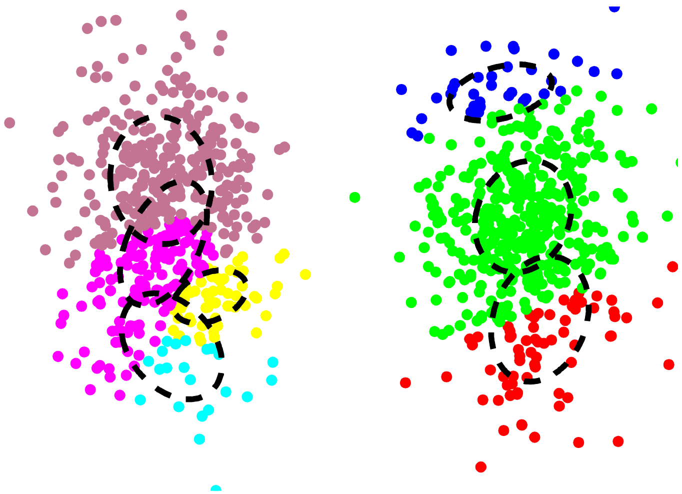


EM, with 4 models

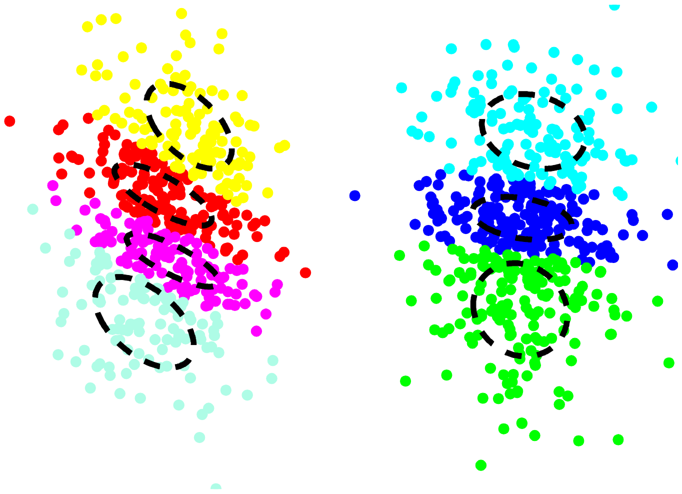


K-means, with 4 models

EM vs K-means

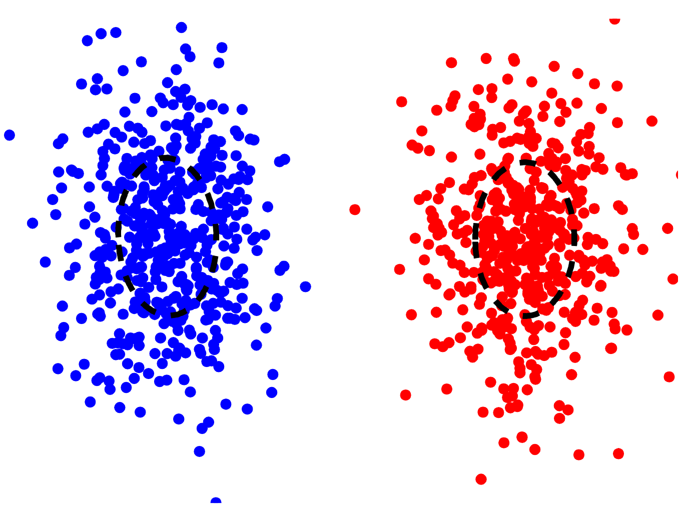


EM, with 7 models



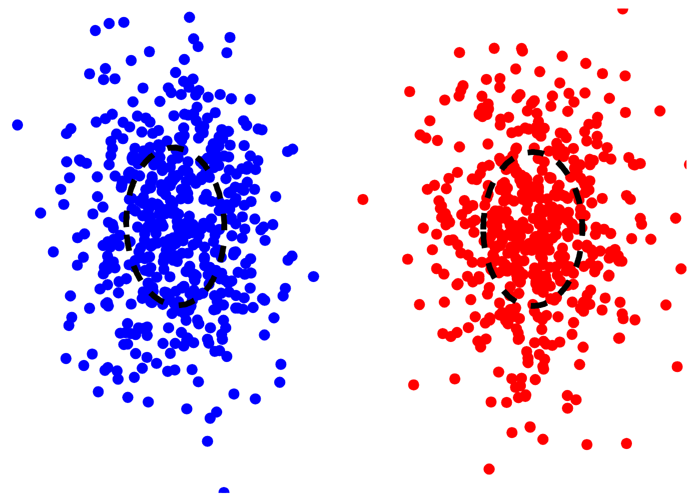
K-means, with 7 models

EM vs K-means + sparsity + many proposals



EM + dirichlet, with 50 models

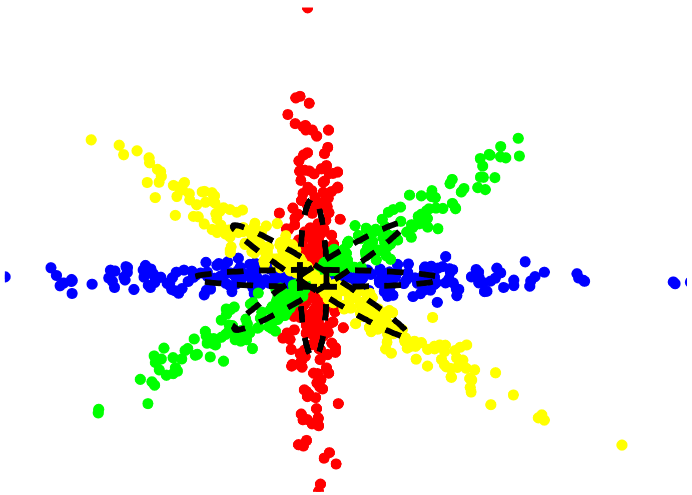
[Figueiredo & Jain, PAMI 2002]



K-means + label cost, with 50 models

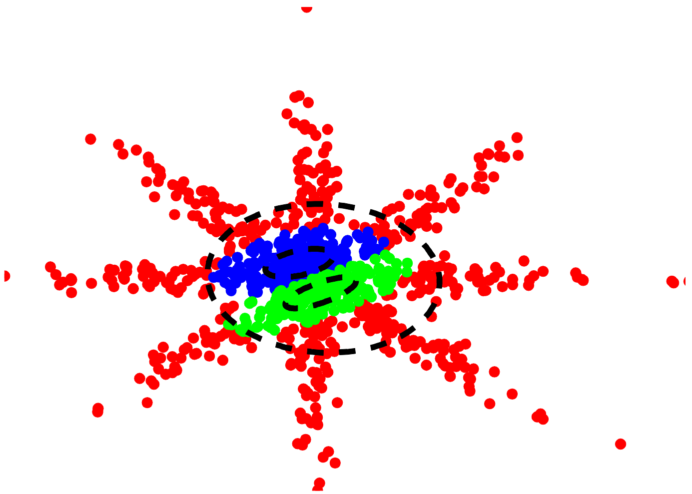
[Delong, Osokin, Isack, Boykov, IJCV 2012]
(PEARL)

EM vs K-means + sparsity + many proposals



EM + dirichlet, with 50 models

[Figueiredo & Jain, PAMI 2002]



K-means + label cost, with 50 models

[Delong, Osokin, Isack, Boykov, IJCV 2012]
(PEARL)

EM vs PEARL

Models in vision have non-overlapping support
(since non-transparent models occlude each other)

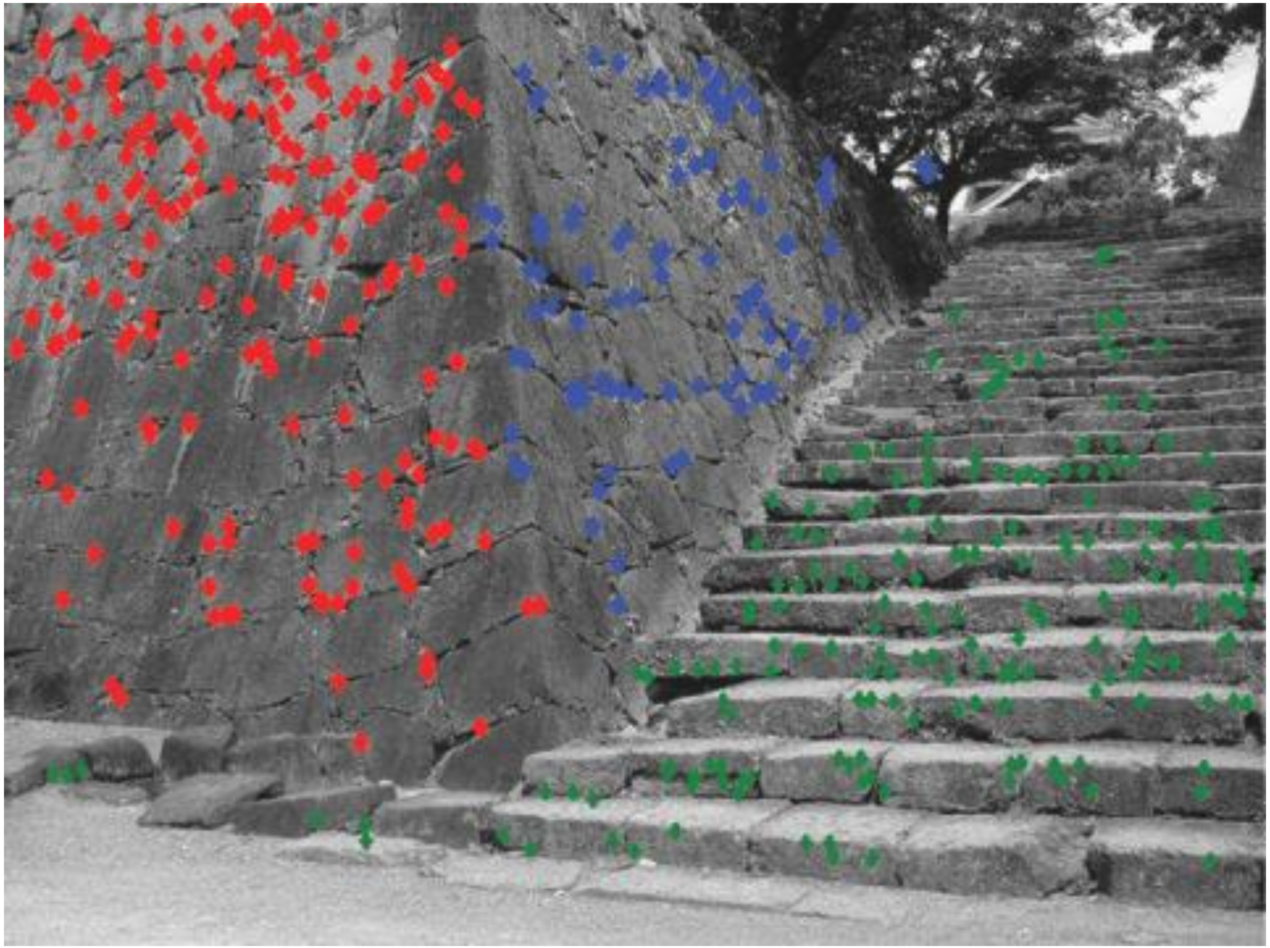
PEARL can integrate both sparsity and spatial regularity
[Delong, Osokin, Isack, Boykov, IJCV 2012]

Q: spatial regularity + EM?

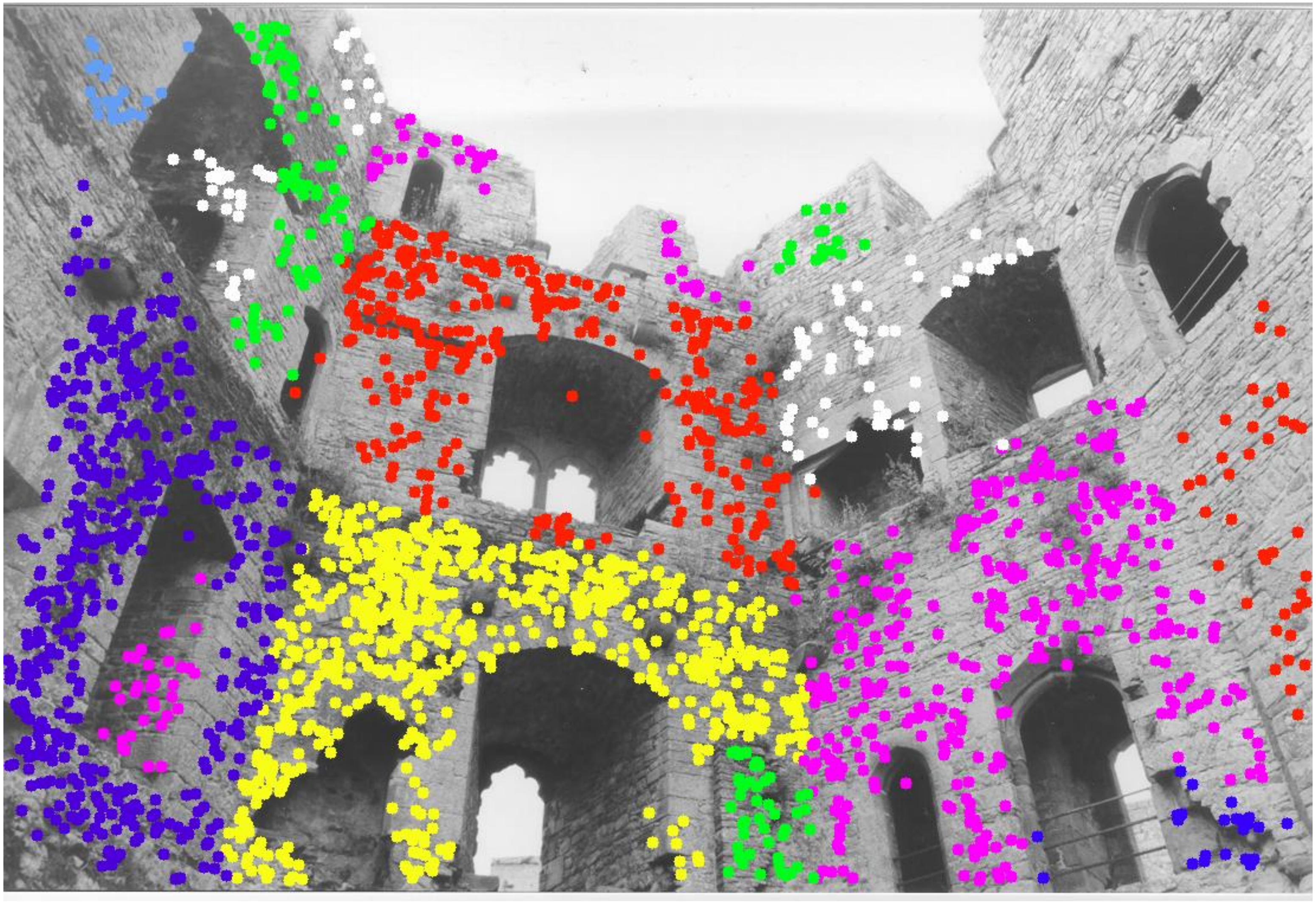
Fitting planes (homographies)

$$E(\mathbf{L}) = \sum_{p} || \, p - L_p \, || + \sum_{(p,q) \in N} w \cdot [L_p \neq L_q] \; + \sum_{L \in \Lambda} h_L \cdot \delta_L(\mathbf{L})$$

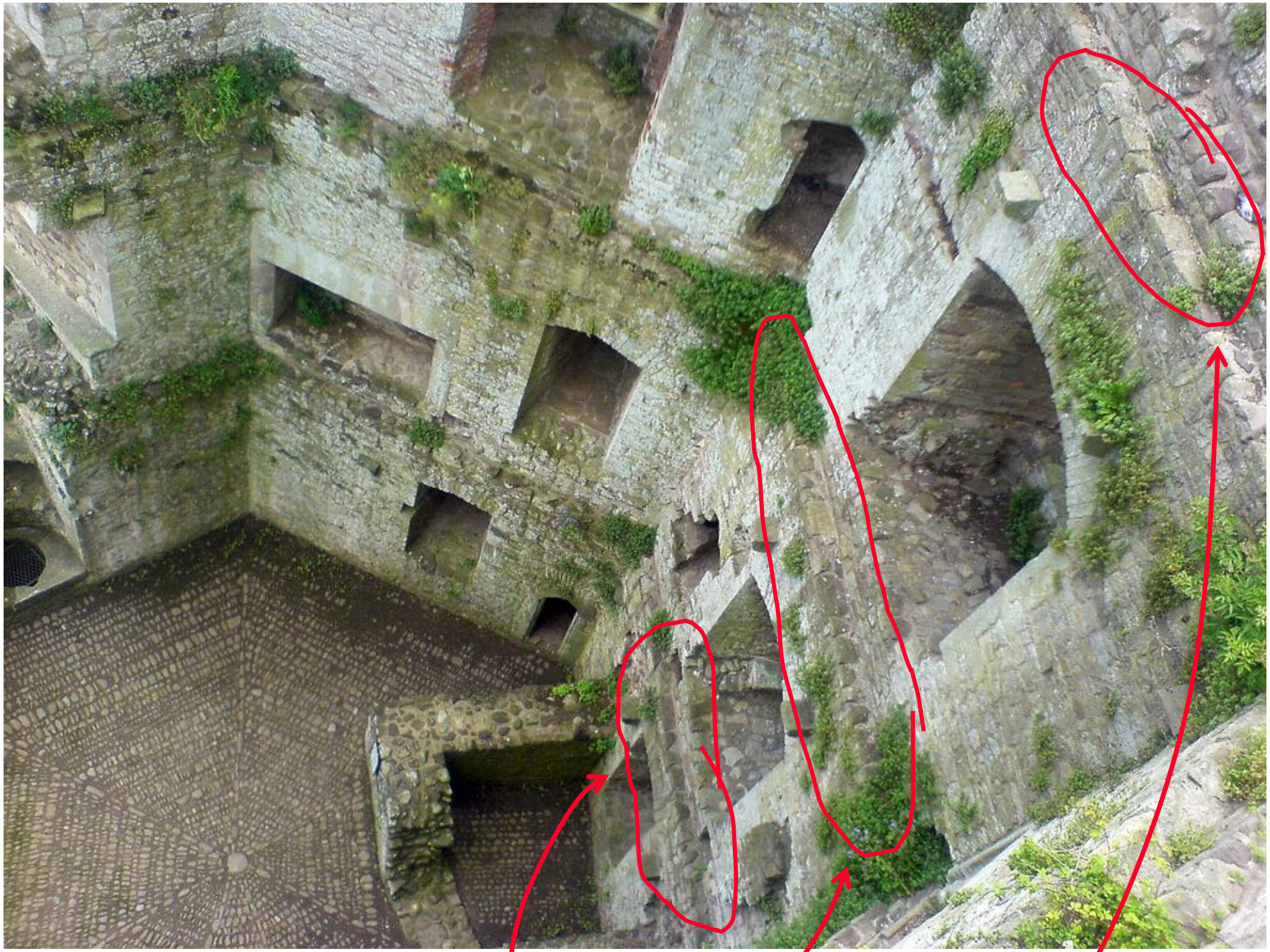
PEARL



Fitting planes (homographies)



same scene
from a different view point…

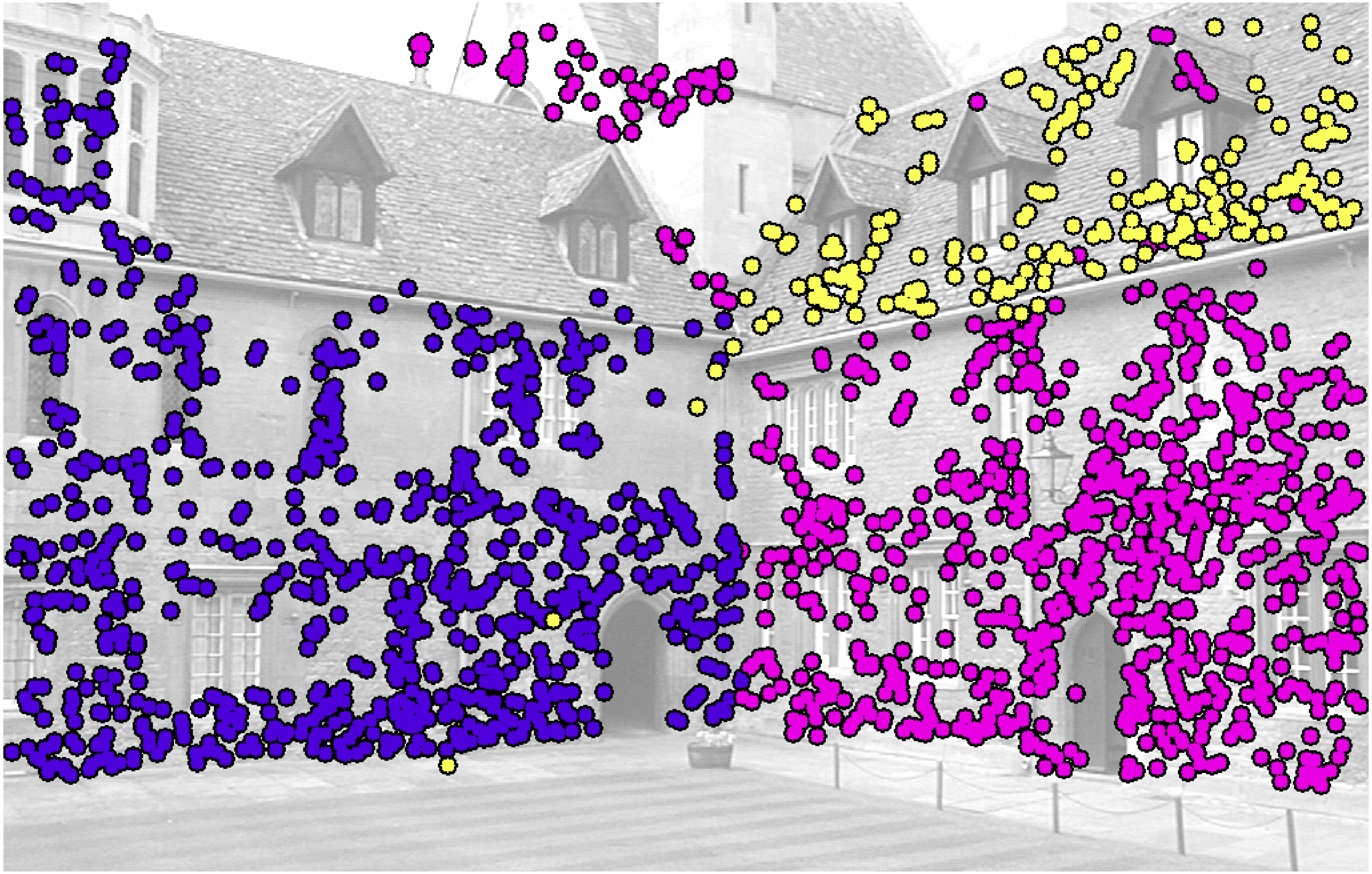


Note **very small steps** between each floor

Fitting planes (homographies)

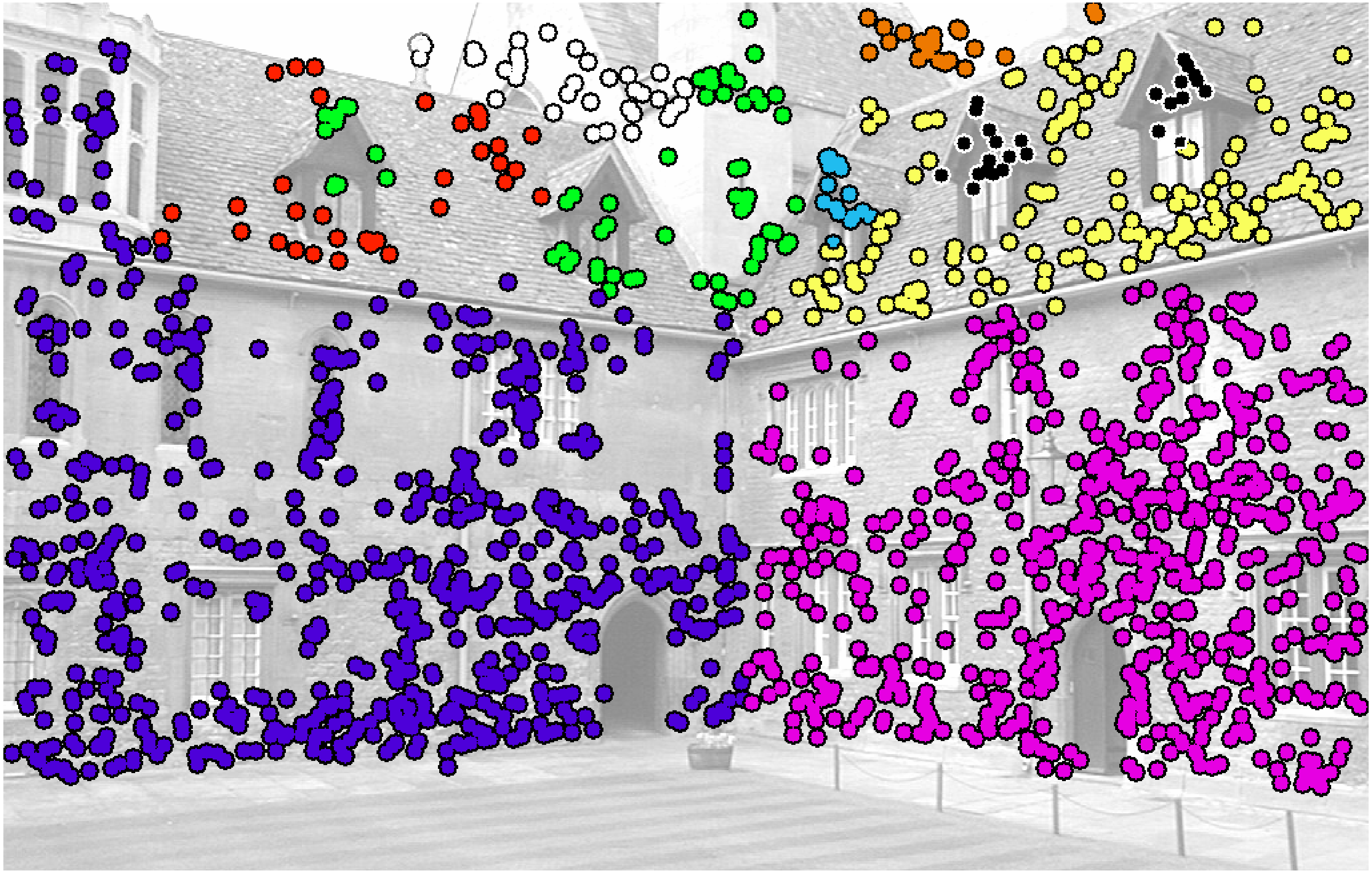
Original image (one of 2 views)

Fitting planes (homographies)



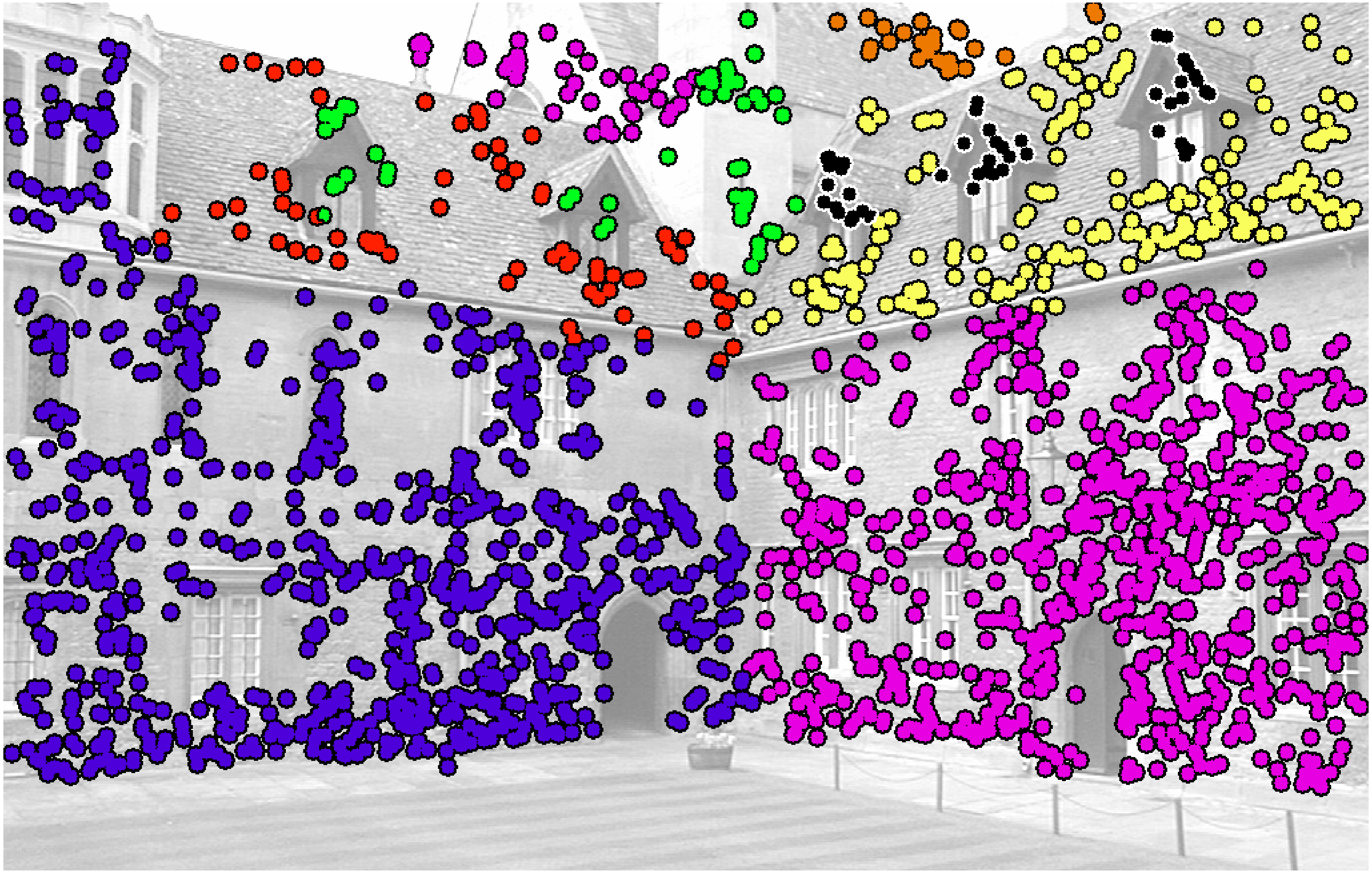
(a) Label costs only

Fitting planes (homographies)



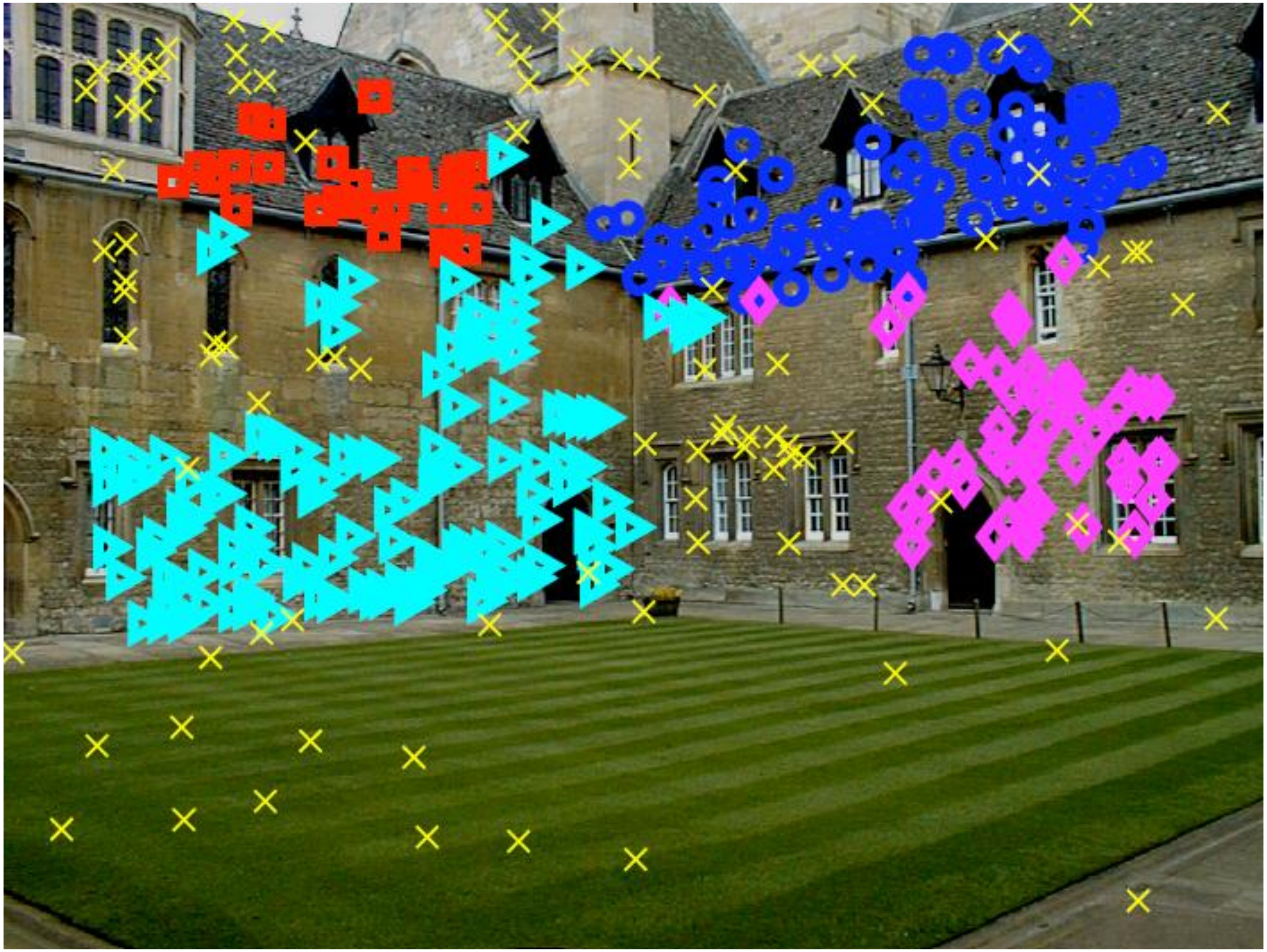
(b) Spatial regularity only

Fitting planes (homographies)



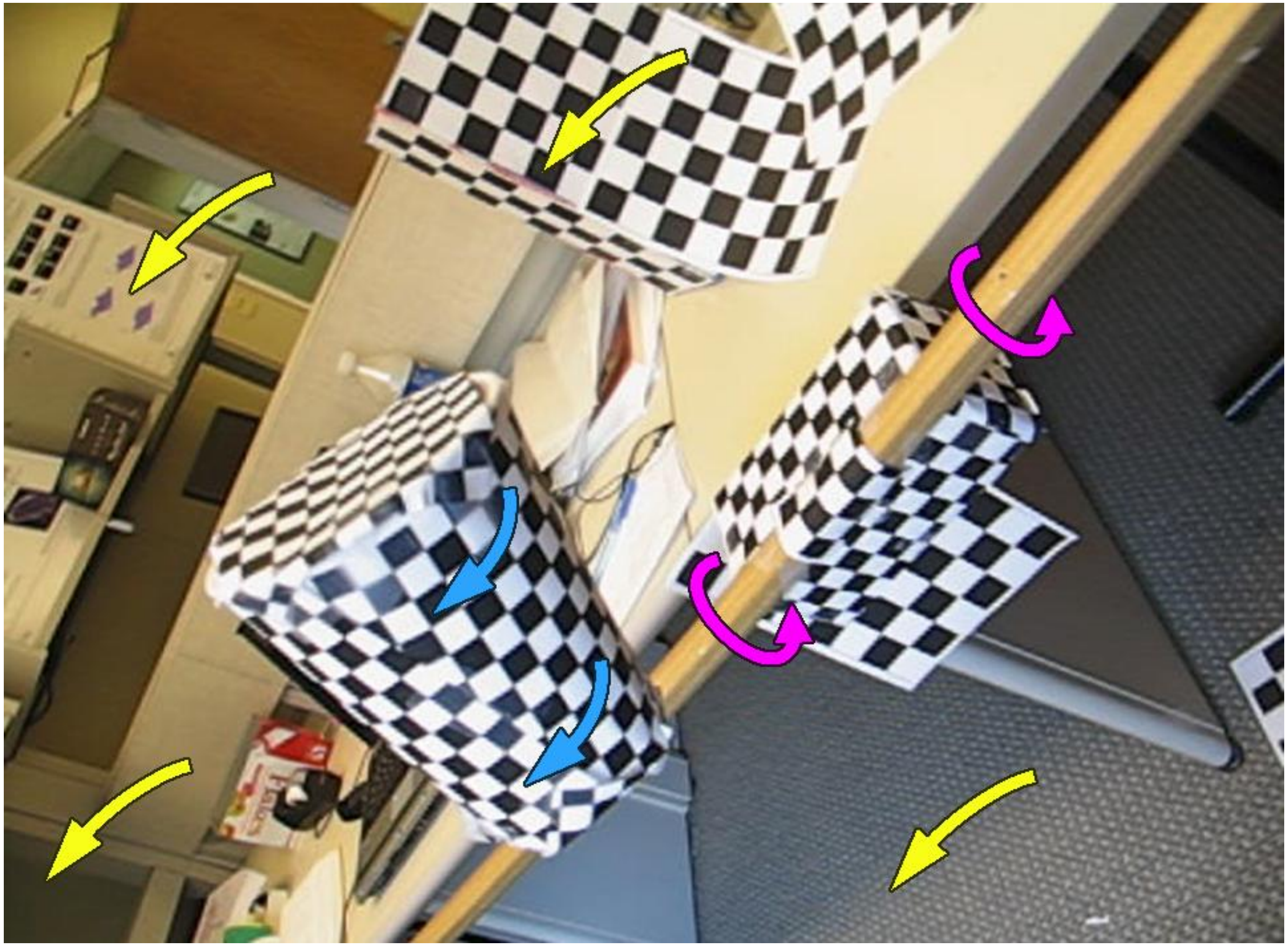
(c) Spatial regularity + label costs

Comparison



based on spectral clustering - Chin, Wang, Sutter ICCV 2009

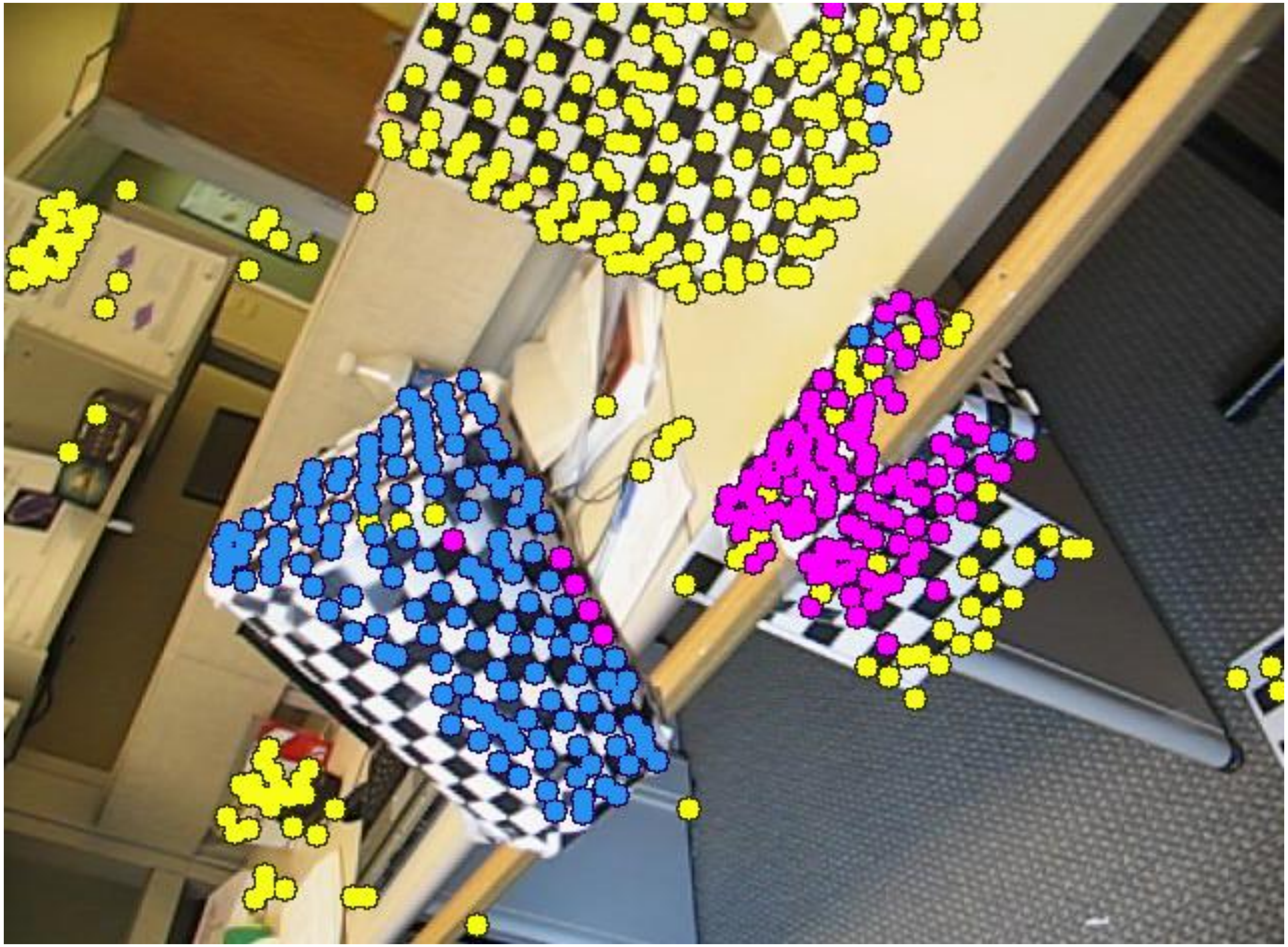
Fitting Rigid Motions (fundamental matrices)



3 motions

Original image [Rene Vidal]

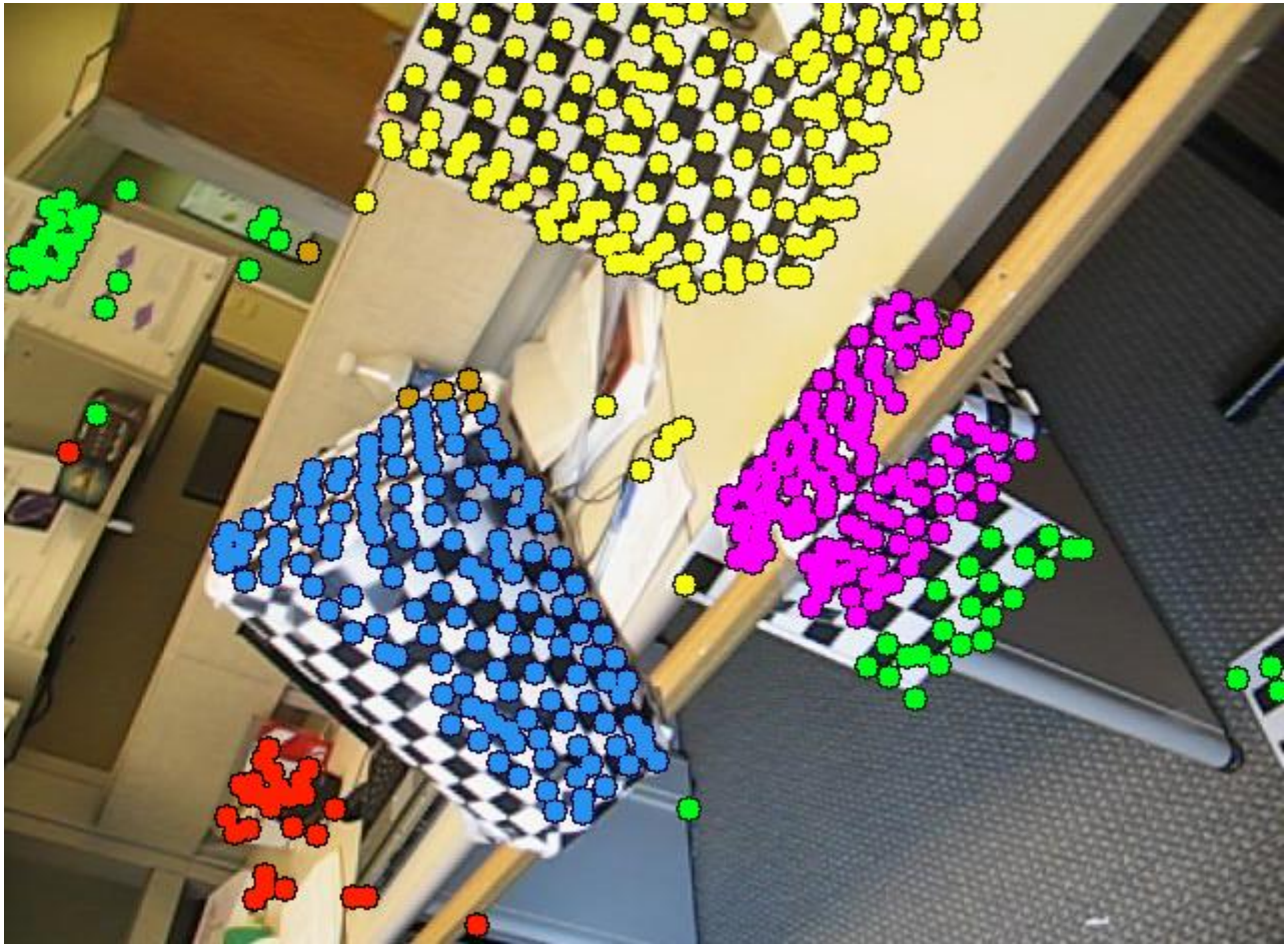
Fitting Rigid Motions (fundamental matrices)



3 motions

(a) Label costs only

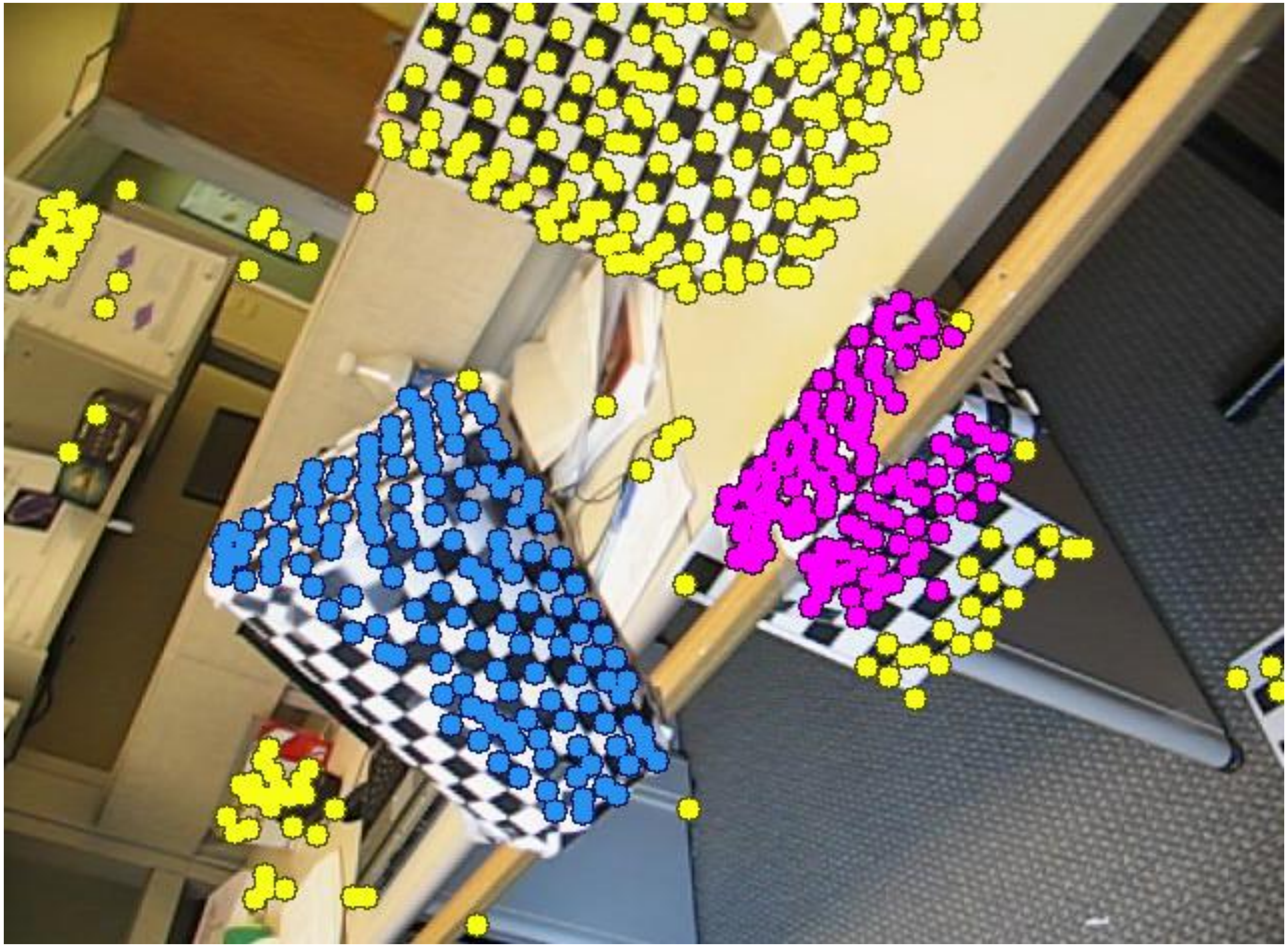
Fitting Rigid Motions (fundamental matrices)



(b) Spatial regularity only

7 motions

Fitting Rigid Motions (fundamental matrices)



3 motions

(c) Spatial regularity + label costs

Fitting Rigid Motions (fundamental matrices)

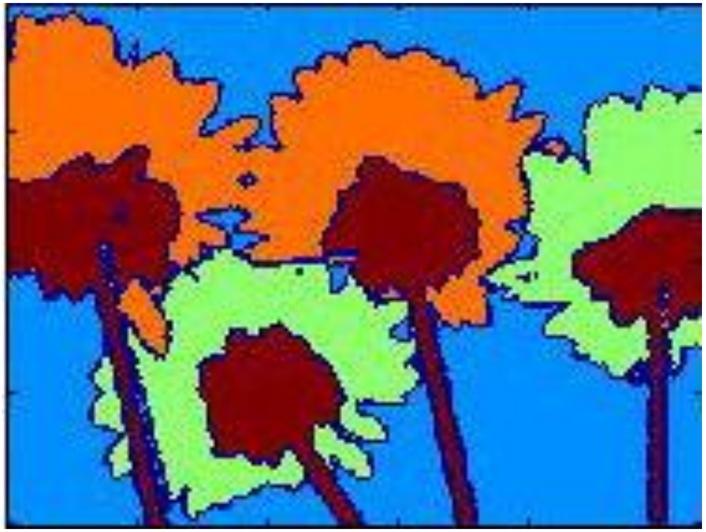
Fitting Rigid Motions (fundamental matrices)

Fitting Rigid Motions (fundamental matrices)

(unsupervised image segmentation)

Fitting color models

label L represents parameters (e.g. mean) of a Gaussian $N(I|L)$

$$-\ln N(I_p \,|\, L_p)$$

$$E_I(\mathbf{L}) = \sum_p (I_p - L_p)^2 + \sum_{(p,q)\in N} w \cdot [L_p \neq L_q]$$

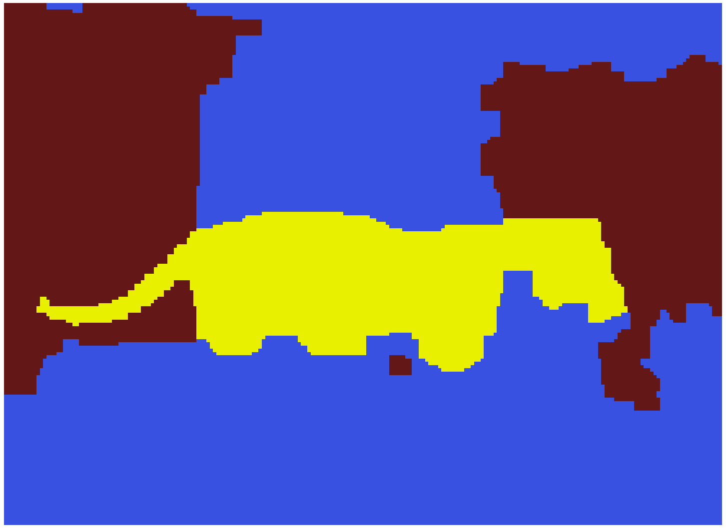
color consistency model (**Chan-Vese**)

(unsupervised image segmentation)

Fitting color models

more generally...

label L represents parameters of an arbitrary distribution $Pr(I|L)$



$-\ln Pr(I_p \mid L_p)$

$$E_I(\mathbf{L}) = \sum_p || p - L_p || + \sum_{(p,q)\in N} w \cdot [L_p \neq L_q] + \sum_{L\in\Lambda} h_L \cdot \delta_L(\mathbf{L})$$

information theory (MDL) interpretation:

= number of bits to compress image I **losslessly**

(unsupervised image segmentation)
Fitting color models

Spatial smoothness + label costs

Zhu & Yuille, PAMI 1996
used continuous formulation (gradient descent)

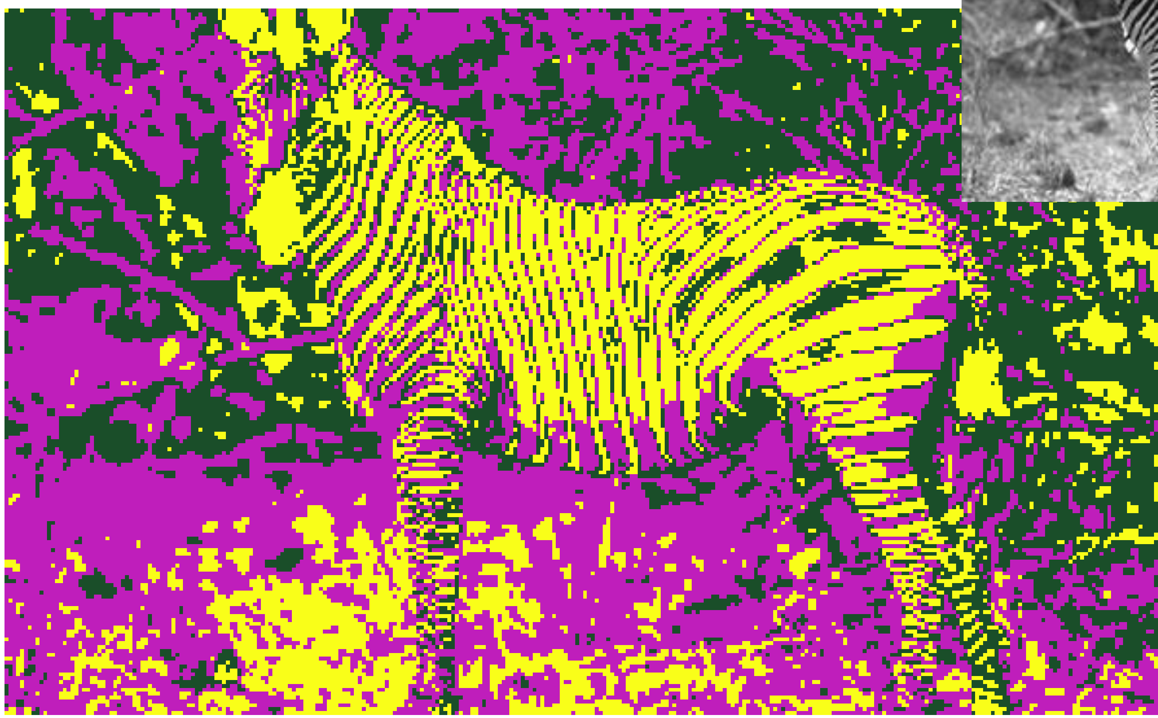
(unsupervised image segmentation)

Fitting color models

Spatial smoothness only [Zabih & Kolmogorov, CVPR 04]

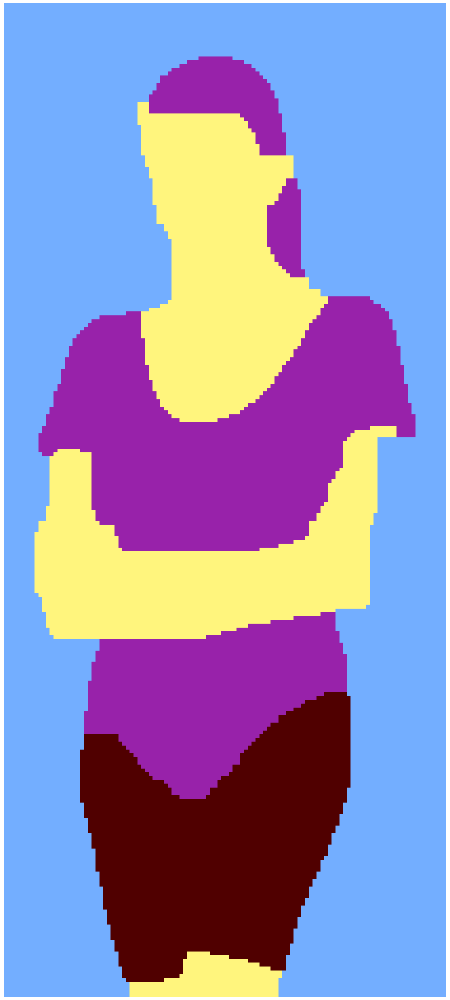
(unsupervised image segmentation)

Fitting color models



Label costs only

(unsupervised image segmentation)
Fitting color models



Spatial smoothness + label costs

Lossy image compression

$\bar{I}$

$$E(\bar{I},L) = E_{\bar{I}}(L) + \lambda \cdot \sum_p \| \bar{I}_p - I_p \|$$

color model fitting
(optimal bits for $\bar{I}$)

distortion of I

Lossy image compression

$\bar{I}$

$$E(\bar{I}, L) = E_{\bar{I}}(L) + \lambda \cdot \sum_p \| \bar{I}_p - I_p \|$$

color model fitting
(optimal bits for $\bar{I}$)

distortion of I

Lossy image compression

$\bar{I}$

$$E(\bar{I},L) \; = \; E_{\bar{I}}(L) \quad + \quad \lambda \cdot \sum_p \| \bar{I}_p - I_p \|$$

color model fitting
(optimal bits for $\bar{I}$)

distortion of I

Lossy image compression

$\bar{I}$

$$E(\bar{I},L) = E_{\bar{I}}(L) \quad + \quad \lambda \cdot \sum_p \| \bar{I}_p - I_p \|$$

color model fitting
(optimal bits for $\bar{I}$)

distortion of I

Lossy image compression

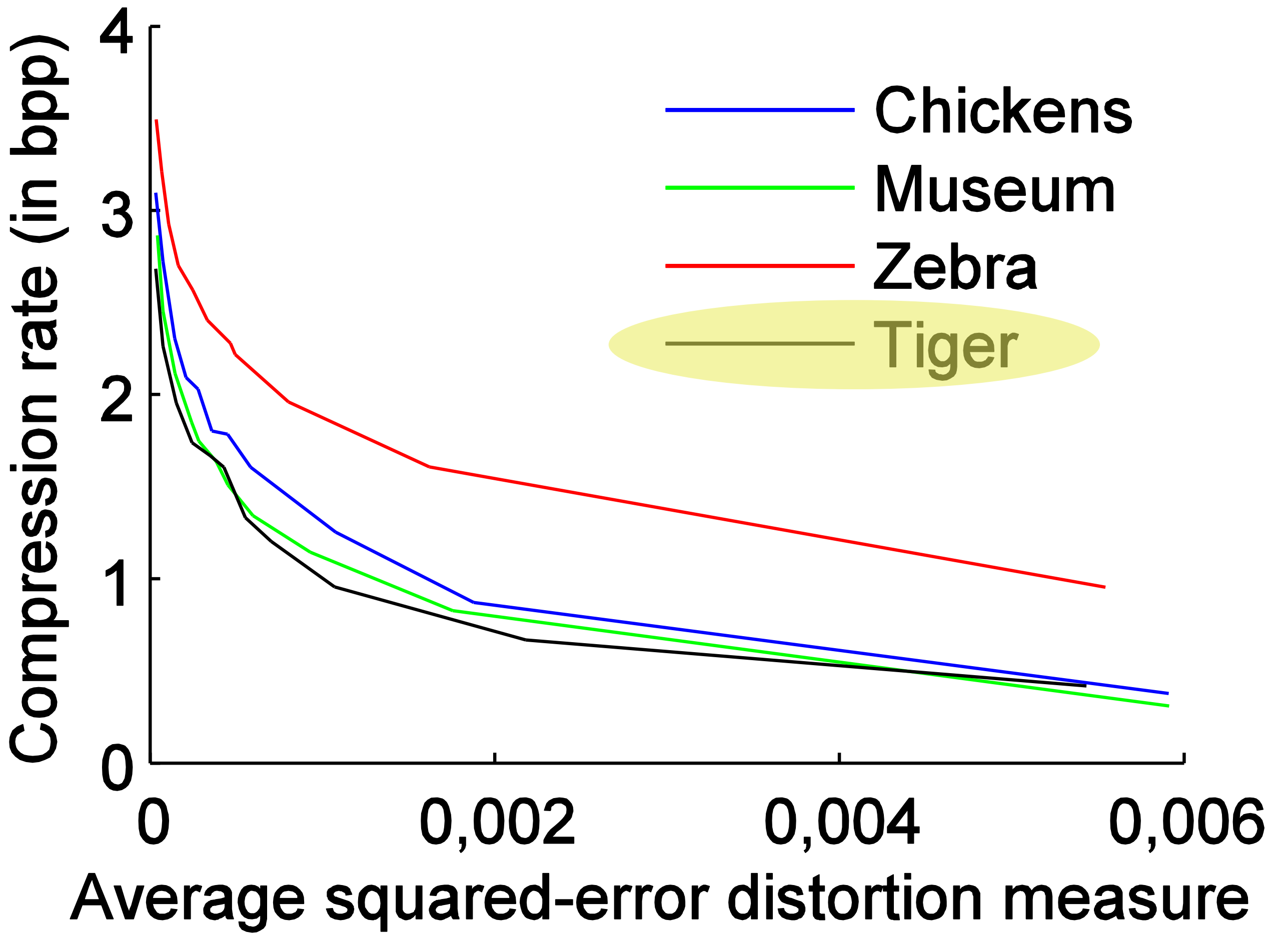
$\bar{I}$

$$E(\bar{I}, L) = E_{\bar{I}}(L) + \lambda \cdot \sum_p \| \bar{I}_p - I_p \|$$

color model fitting (optimal bits for $\bar{I}$)

distortion of I

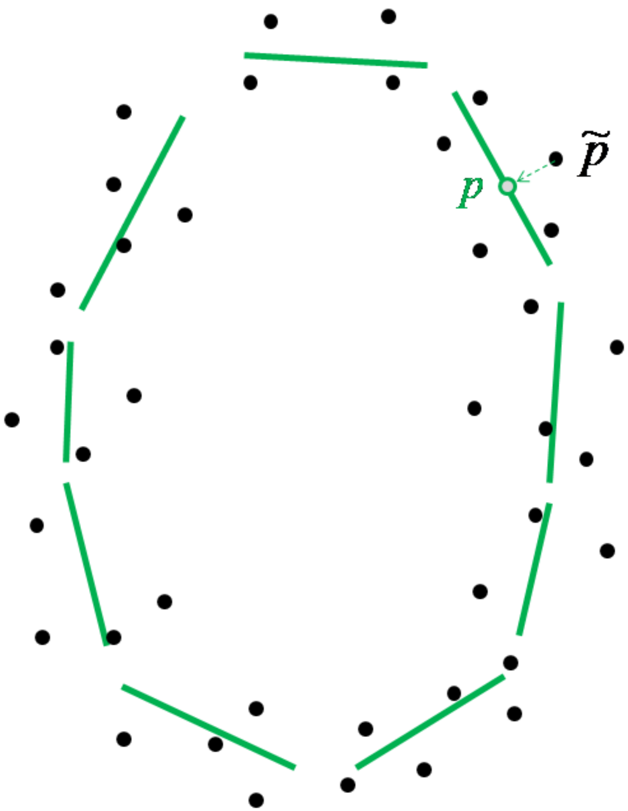
Rate-Distortion Plot

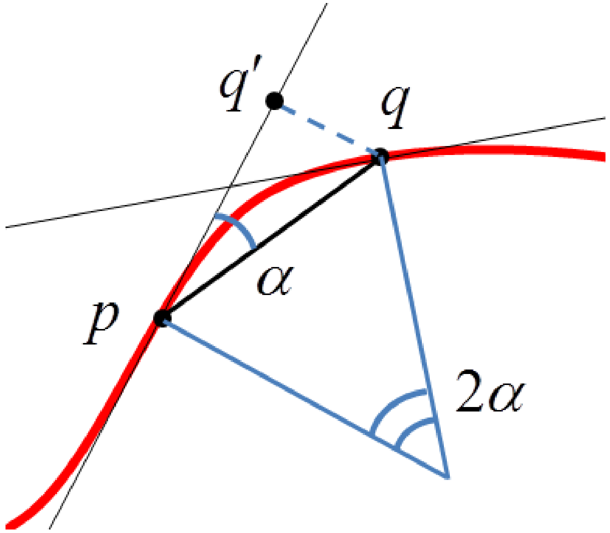


Conclusions

- Energy-based multi-model fitting
- Algorithms for minimizing label-costs energies with global optimality guarantees
  - **extended *a-expansion***, standard ***UFL* heuristics**
- Exploring a continuum of labels, ***PEARL***

Extensions
Piece-wise smooth model fitting

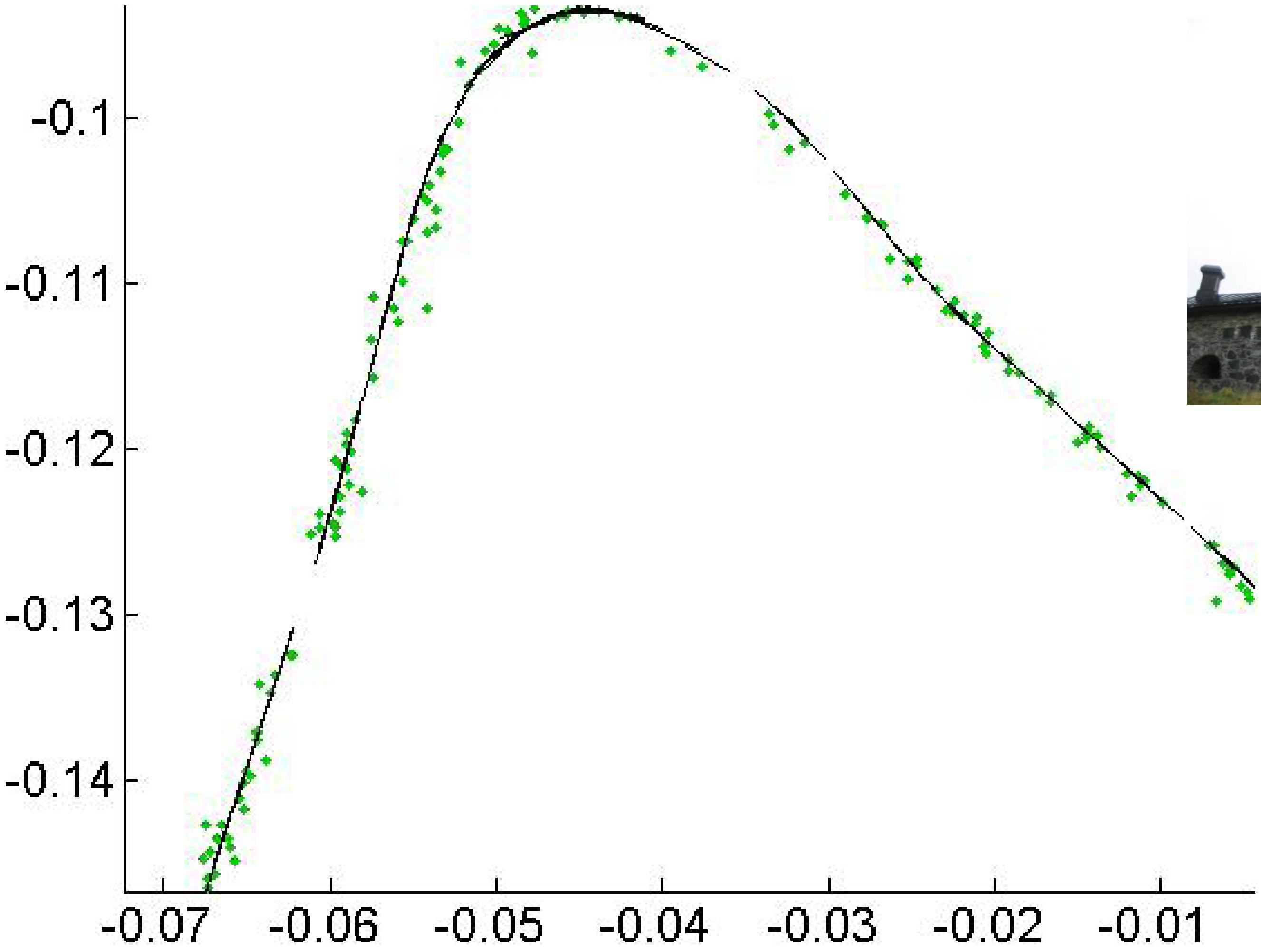




$$\kappa = \frac{|q - q'|}{2|p - q|^2}$$

[Olsson, Boykov CVPR12]

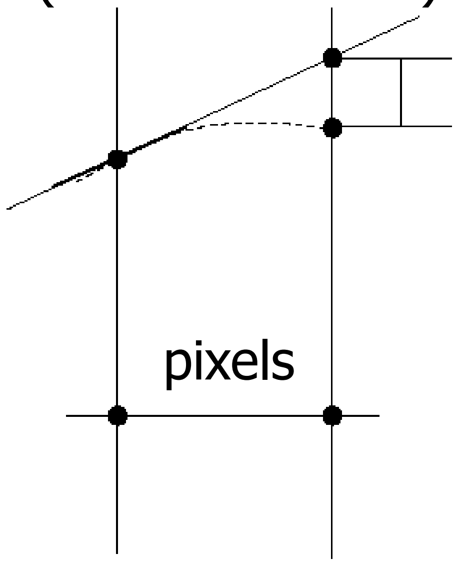
Extensions
Piece-wise smooth model fitting



[Olsson, Boykov CVPR12]

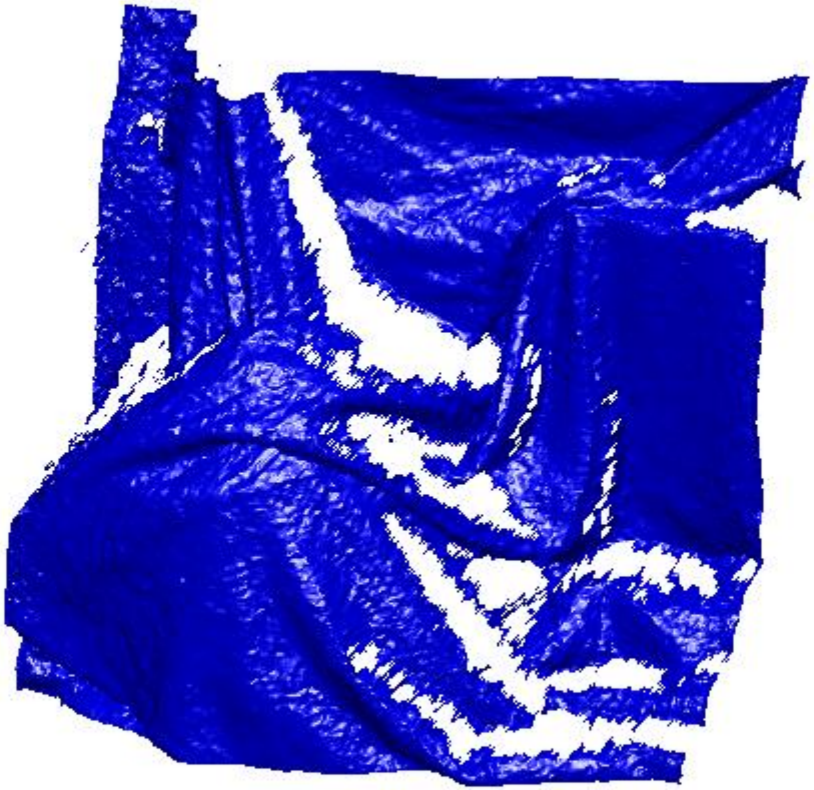
Extensions
Piece-wise smooth stereo

Labels are tangents
(incl. orientation)

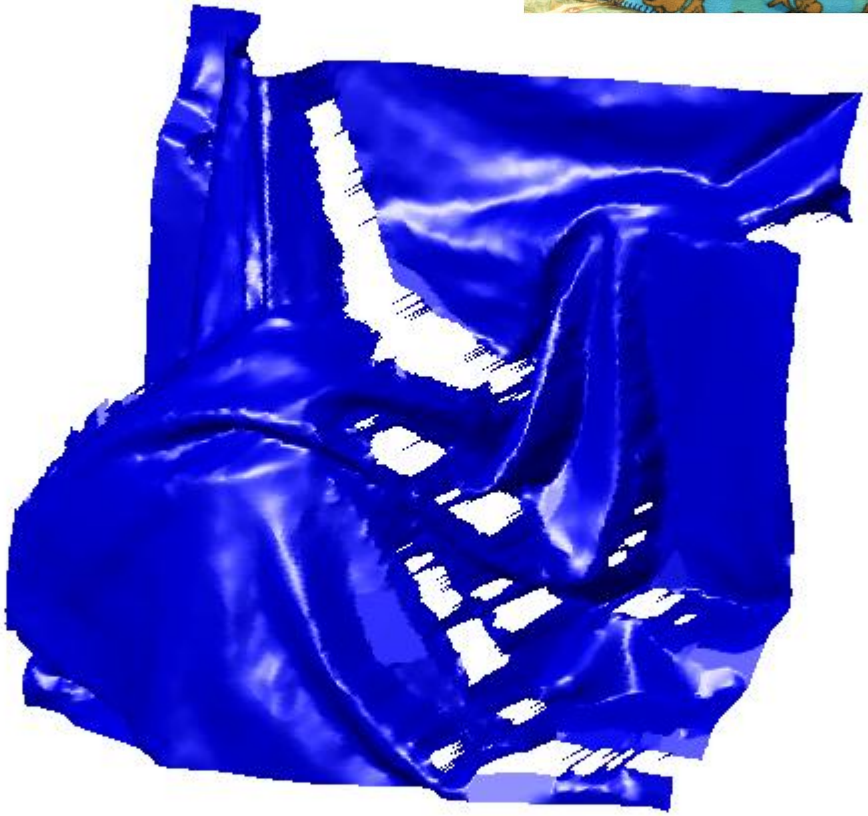


[Olsson, Ulen, Boykov CVPR13]

Extensions
Piece-wise smooth stereo



First-order smoothness

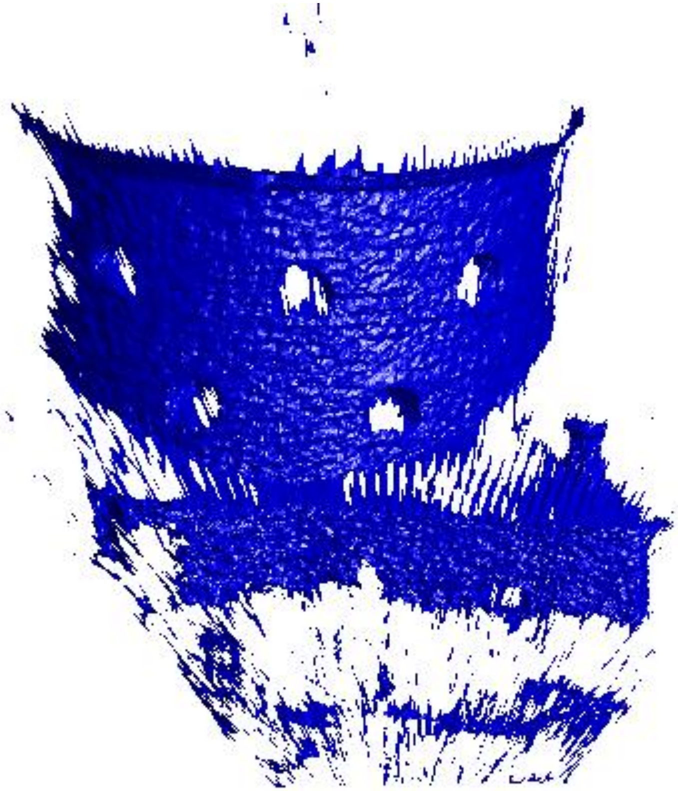


Second-order smoothness

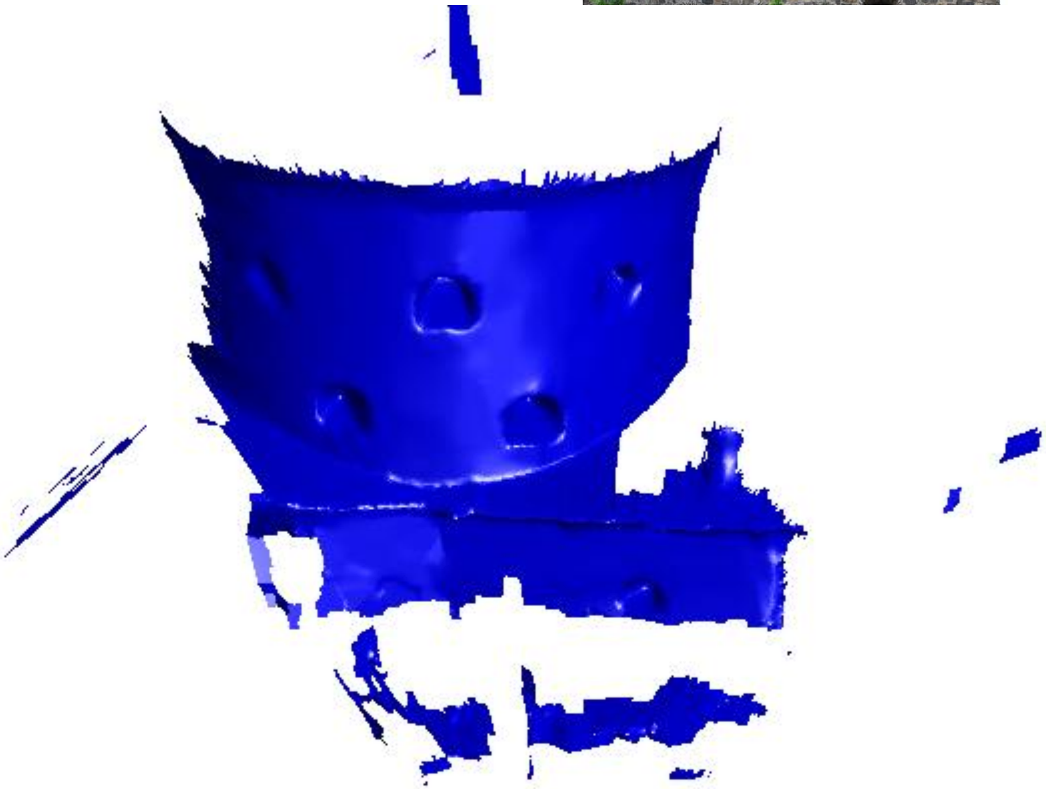
Unlike Woodford et al.'08, first-order interactions

[Olsson, Ulen, Boykov CVPR13]

Extensions
Piece-wise smooth stereo



First-order smoothness



Second-order smoothness

Unlike Woodford et al.'08, first-order interactions

[Olsson, Ulen, Boykov CVPR13]

://www.youtube.com/watch?v=2HAFSwFRoR8&list=UUVS7P9dioyjoN7j9mHStQ_Q&feature=player_detailpage&t=7