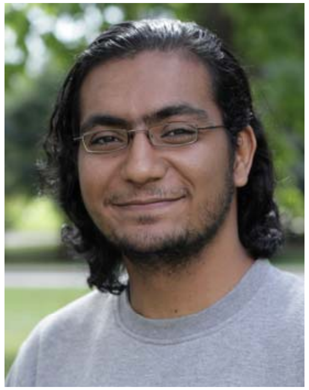
Fast Energy Minimization with Label Costs and Applications in Model Fitting

presented by Anton Osokin

co-authors:

Andrew Delong Yuri Boykov Hossam Isack



some slides taken from Yuri Boykov and Andrew Delong

Popular energy in vision

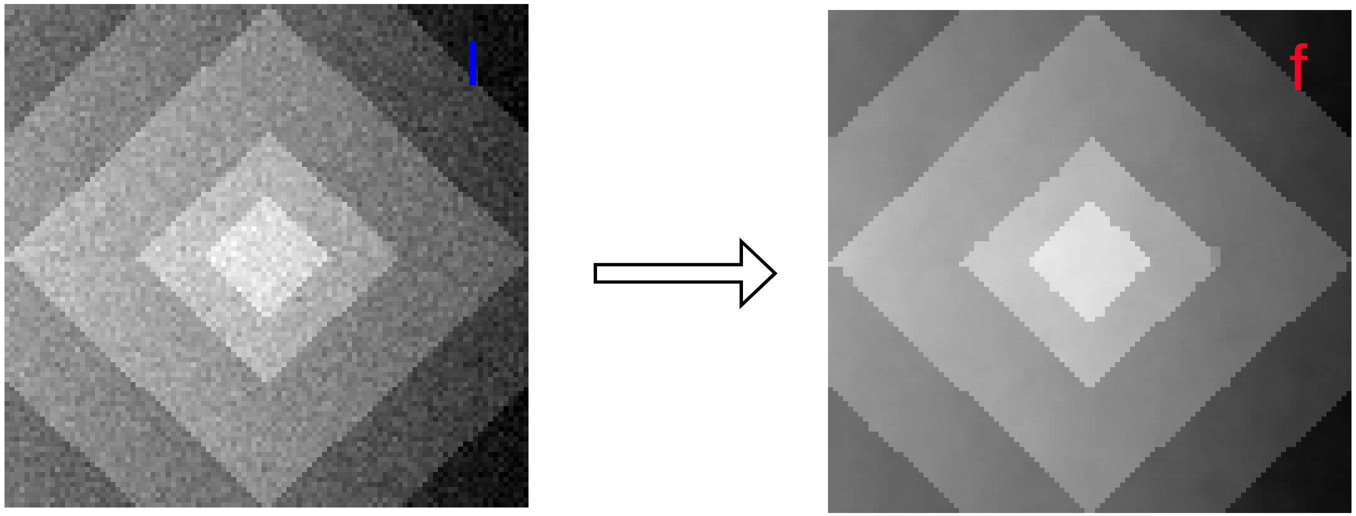
$$f_p \in \{1, \dots, K\}$$

$$E(f) = \sum_p D_p(f_p) + \sum_{(p,q)\in N} V_{pq}(f_p, f_q) \to \min_f$$

$D_p(f_p)$ – unary potentials

$V_{pq}(f_p, f_q)$ – pairwise potentials

Image restoration



observed noisy image I

image labeling f
(restored intensities)

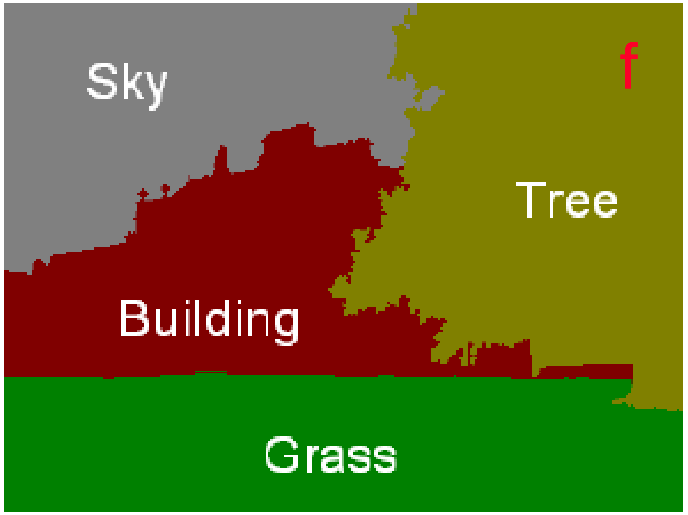
$\mathbf{I} = \{ I_1, I_2, \ldots, I_n \}$

$\boldsymbol{F} = \{ f_1, f_2, \ldots, f_n \}$

data fidelity — *spatial regularization*

$$E(f) = \sum_p (f_p - I_p)^2 + \sum_{(p,q)\in N} (f_p - f_q)^2$$

Image segmentation



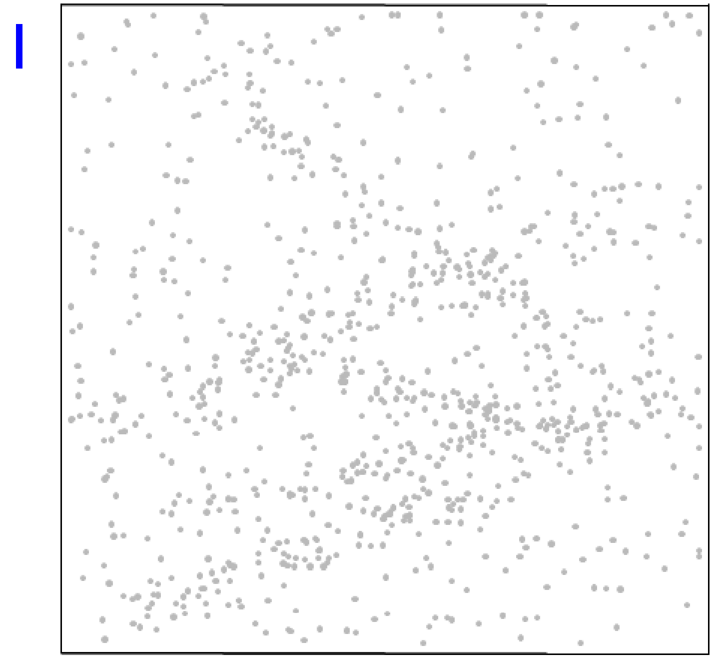
Observed image I

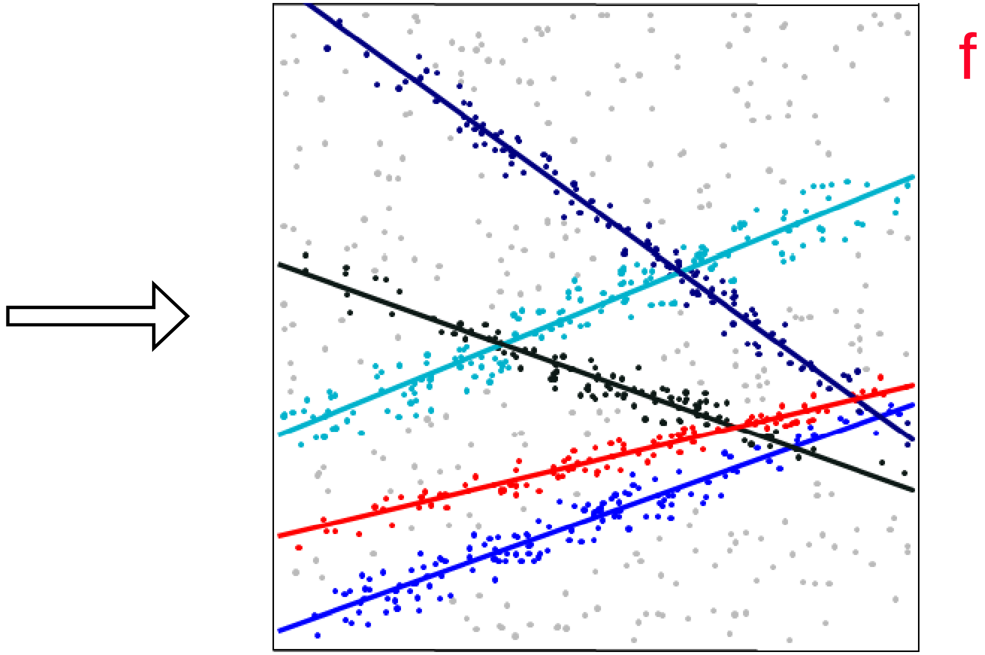
image labeling f (segments)

data fidelity *spatial regularization*

$$E(f) = \sum_{p} -\log P(f_p \mid I_p) + \theta \sum_{(p,q)\in N} [f_p \neq f_q]$$

Geometric model fitting



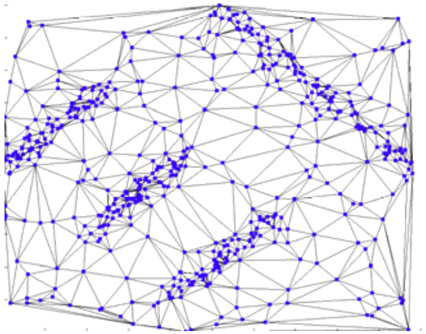


Sampled points I

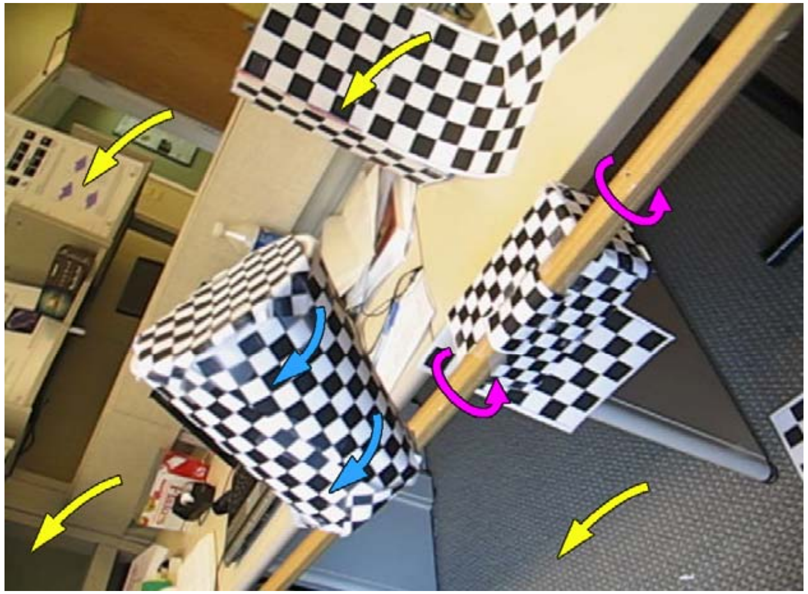
Points clustering f

data fidelity *spatial regularization*

$$E(f) = \sum_p -\log dist(p, f_p) + \theta \sum_{(p,q)\in N} \frac{[f_p \neq f_q]}{dist(p,q)}$$



More geometric model fitting



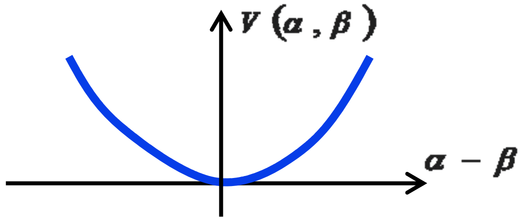
Motion estimation

Plane fitting

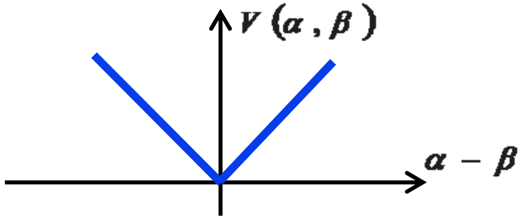
$$E(f)=\sum_{p} D_p(f_p)+\sum_{(p,q)\in N} V_{pq}(f_p,f_q)$$

Optimization

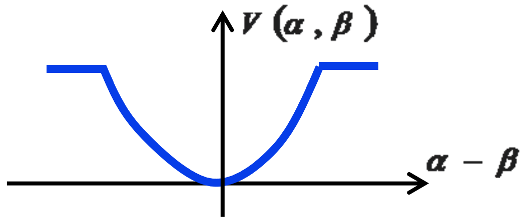
- Convex regularization
  - gradient descent works
  - exact polynomial algorithms



- TV regularization
  - a bit harder (non-differentiable)
  - global minima algorithms (Ishikawa, etc.)

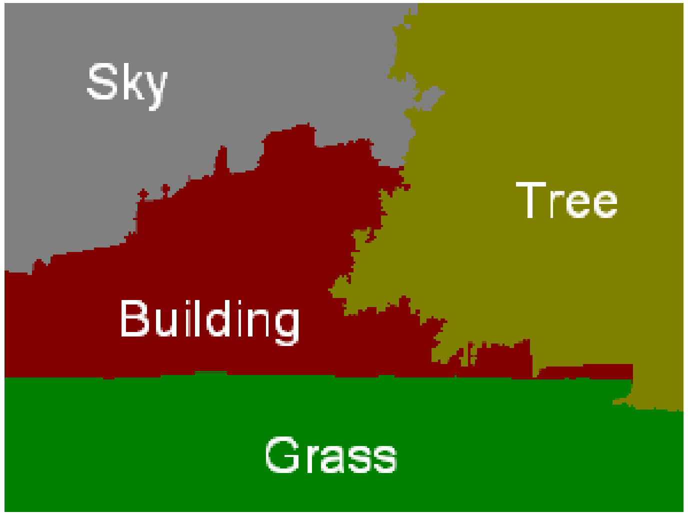


- Robust regularization
  - NP-hard, many local minima
  - good approximations (message passing, a-expansion, a/b-swap)

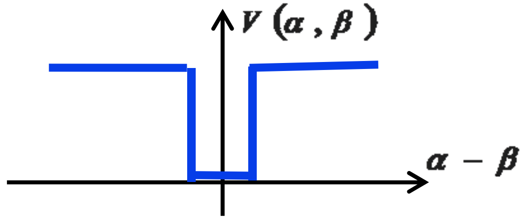


Potts model
(piece-wise constant labeling)

$$E(f) = \sum_p -\log P(f_p \mid I_p) + \theta \sum_{(p,q)\in N} [f_p \neq f_q]$$



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



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

maxflow/mincut
combinatorial algorithms

Adding label costs

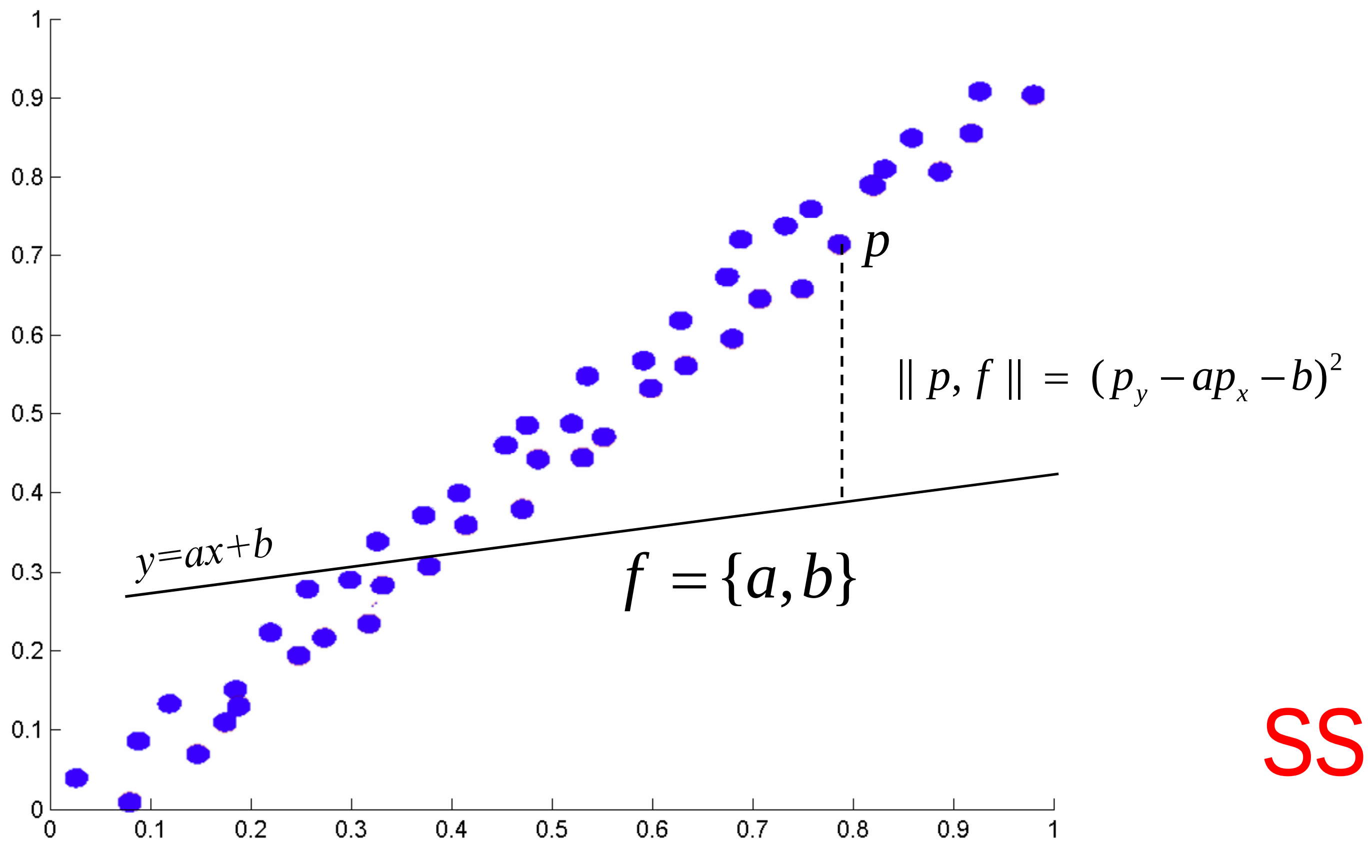
$$E(f) = \sum_{p} D_p(f_p) + \sum_{(p,q)\in N} V(f_p, f_q) \quad + \sum_{f\in\Lambda} h_f \cdot \delta_f(f)$$

Λ - set of labels allowed at each point p

$$\delta_f(f) = \begin{cases} 1, & \exists p : f_p = f \\ 0, & otherwise \end{cases}$$

Model fitting

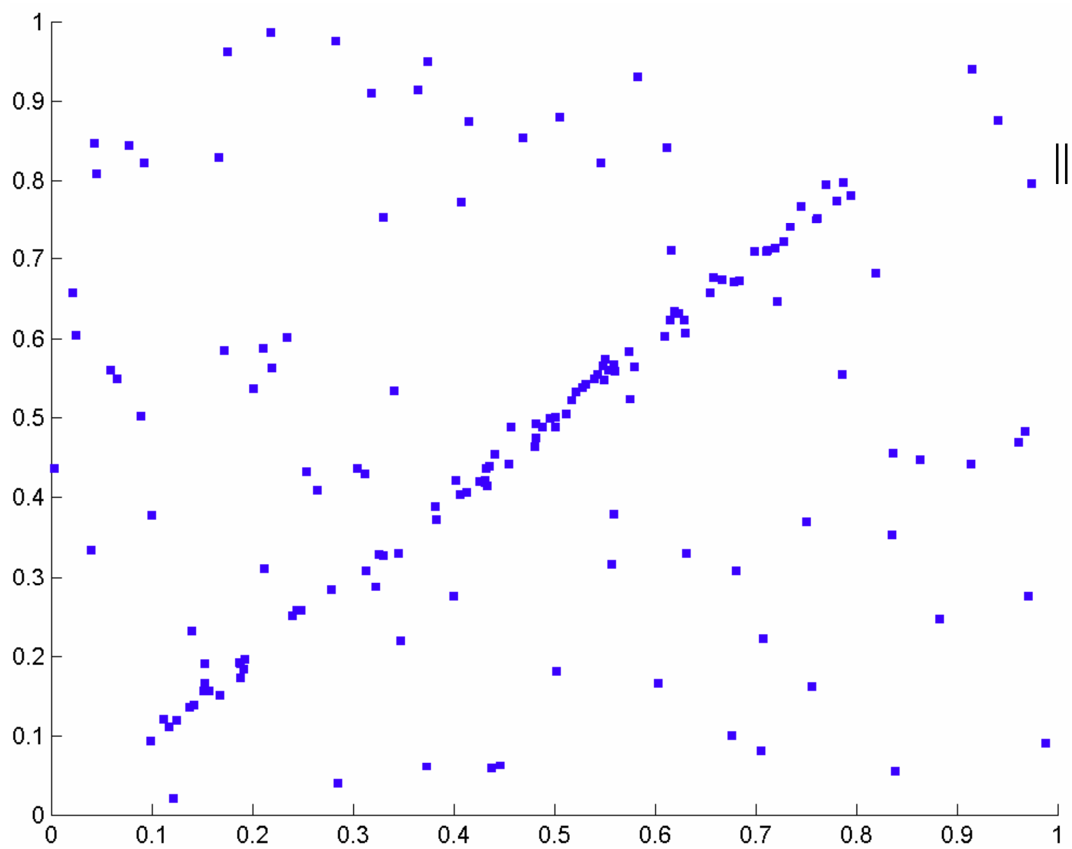
$$f^* = \arg\min_f \sum_p \| p, f \|$$



SSD

Many outliers

quadratic errors fail



use more robust error measures, e.g.

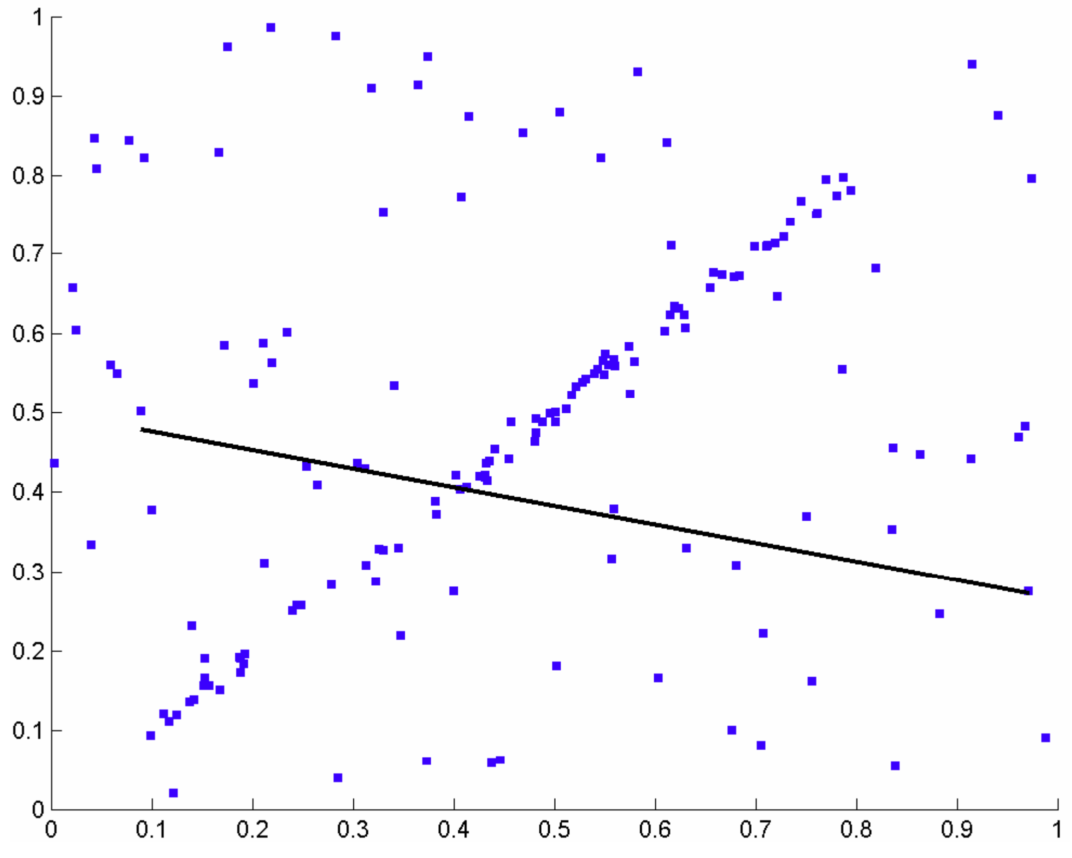
$$\| p, f \| = | p_y - a p_x - b |$$

gives "MEDIAN" line

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

<u>RANSAC</u>

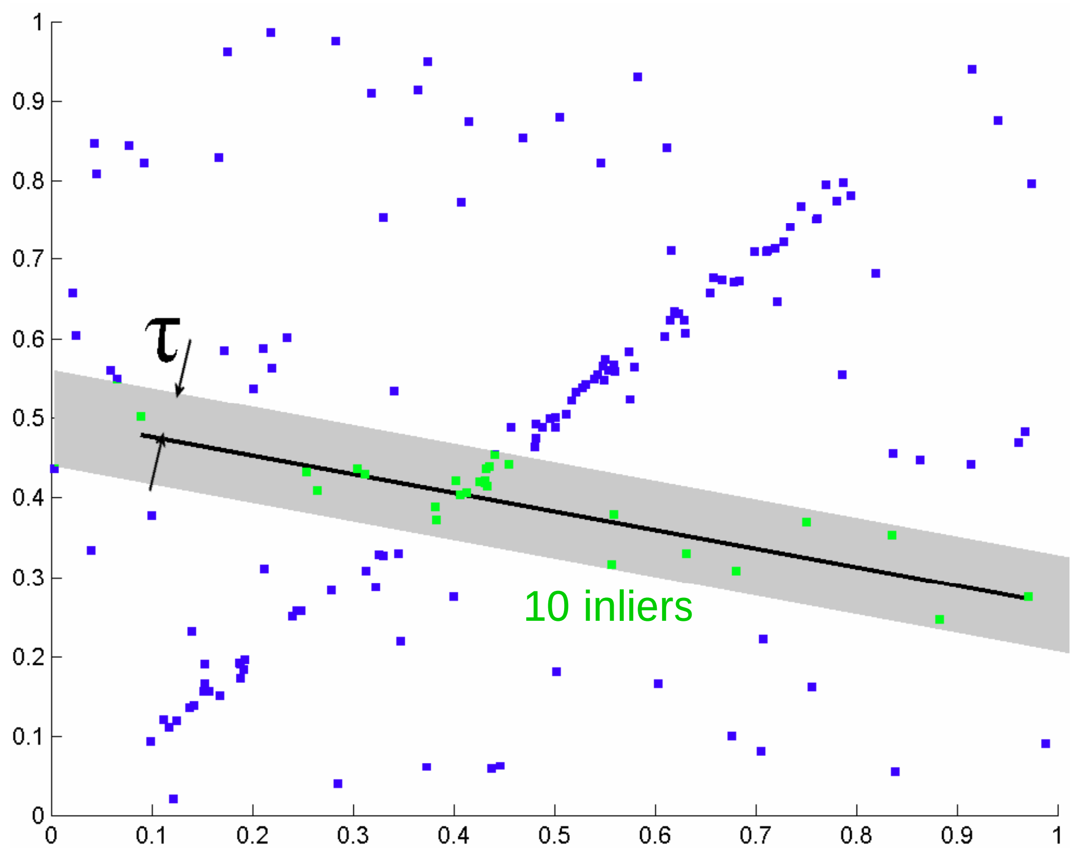
Many outliers



1. sample randomly two points, get a line

RANSAC

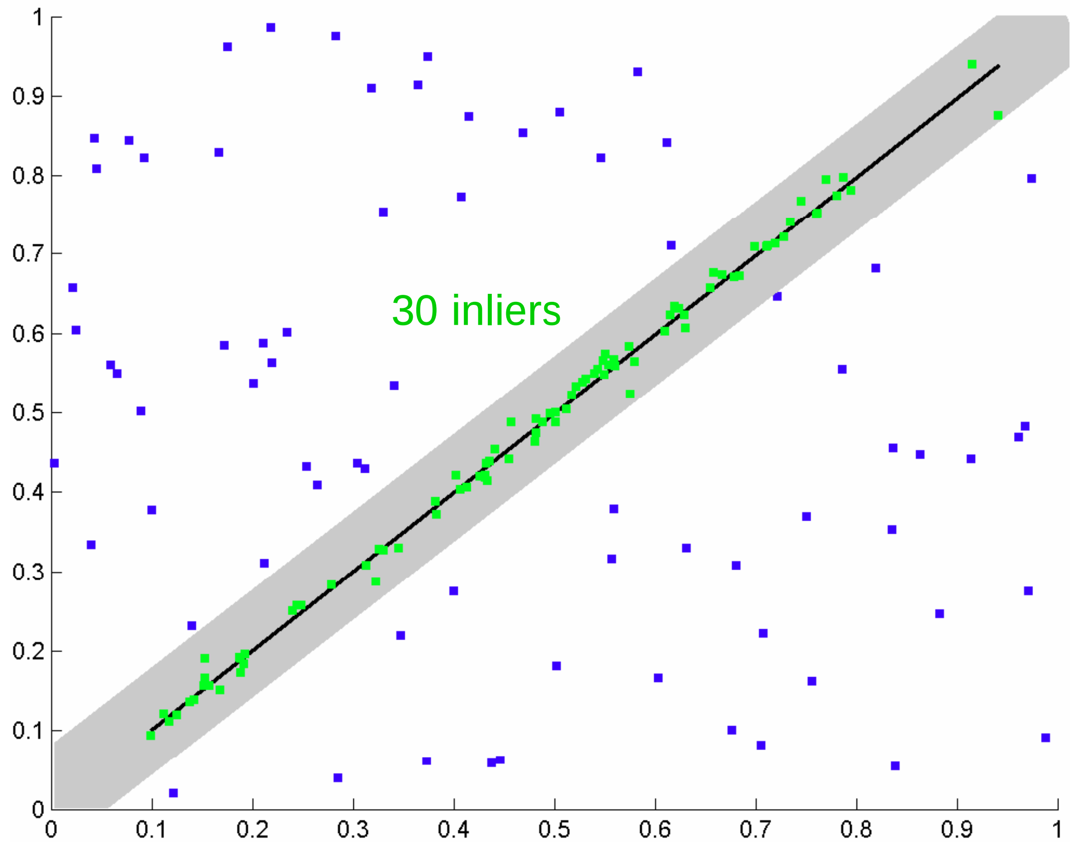
Many outliers



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

RANSAC

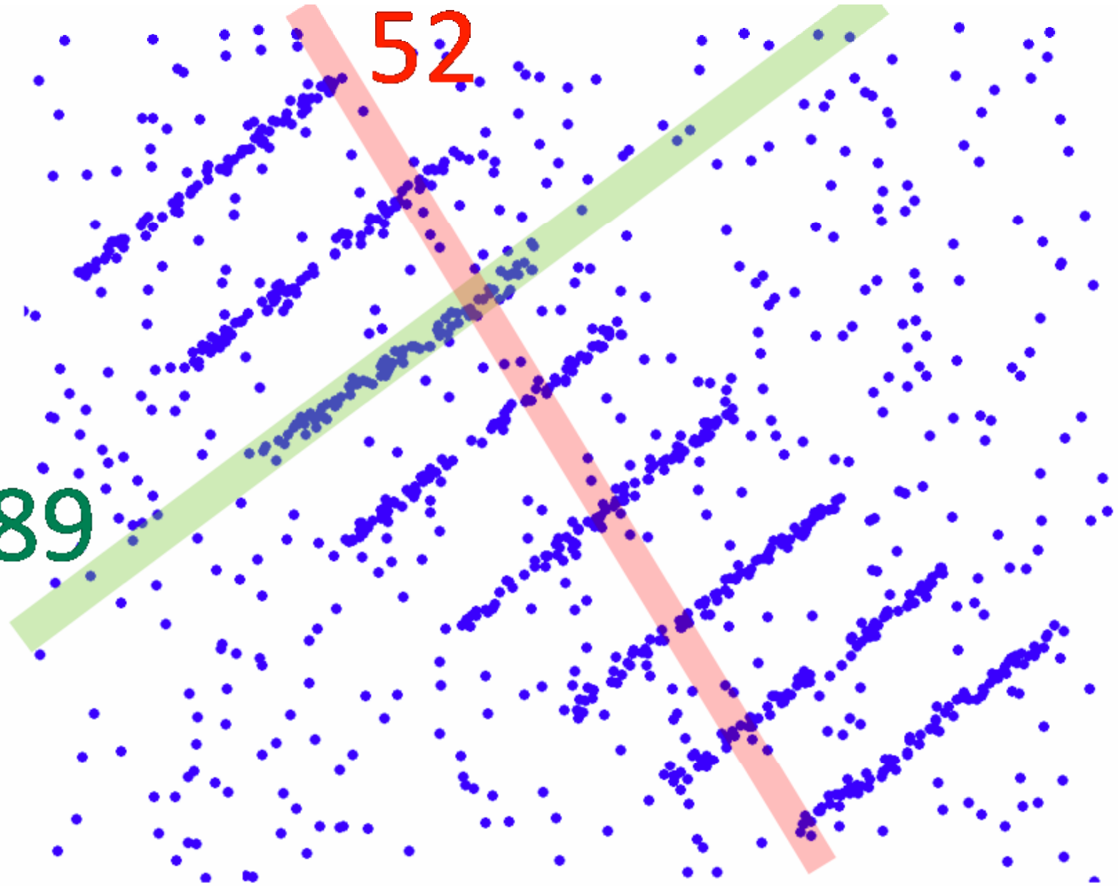
Many outliers



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

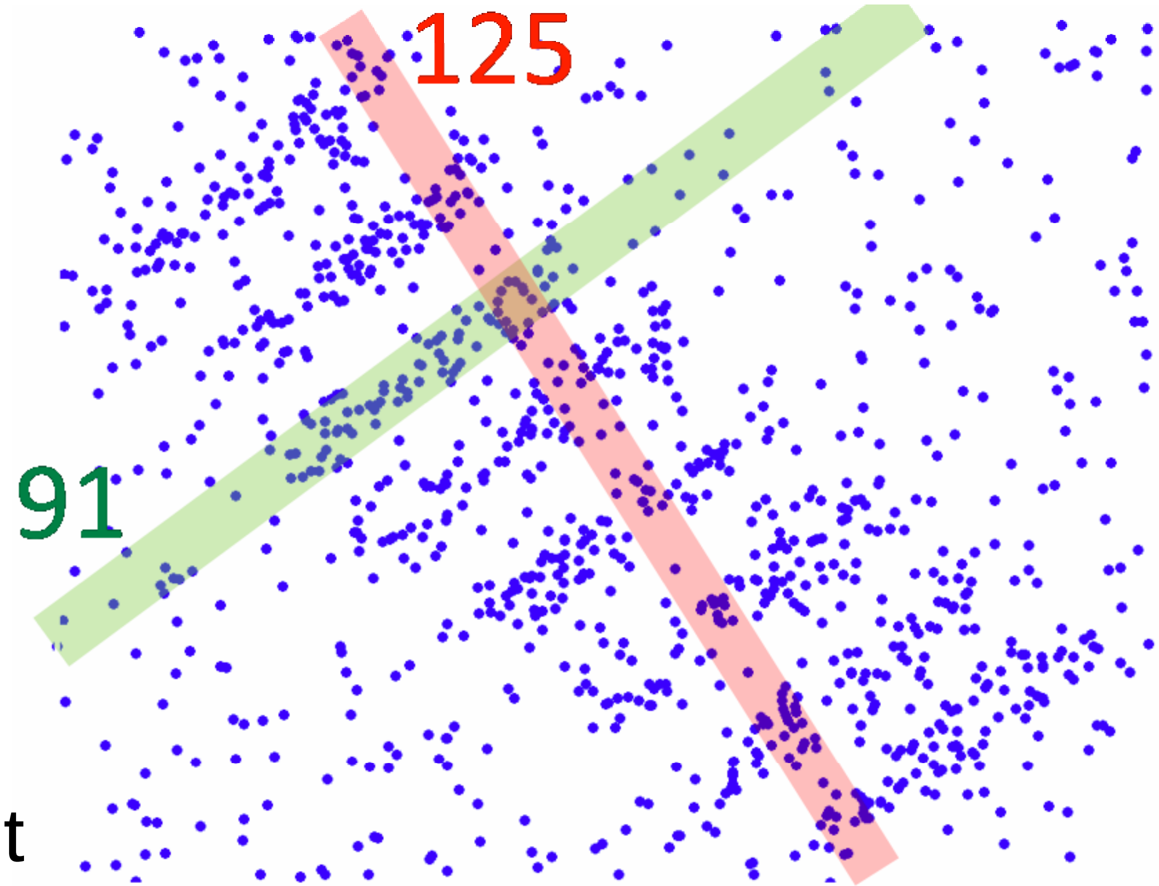
Multiple models and many outliers



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

$$E(f) = \sum_{p} \| p, f \|$$

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

- find optimal **label** f
for one very specific
error measure

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

Energy-based approach

$$E(f) = \sum_{p} \| p, f_p \|$$

If **multiple** models

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

-find optimal **labeling**
$\boldsymbol{f} = \{ f_1, f_2, \ldots, f_n \}$

Need regularization!

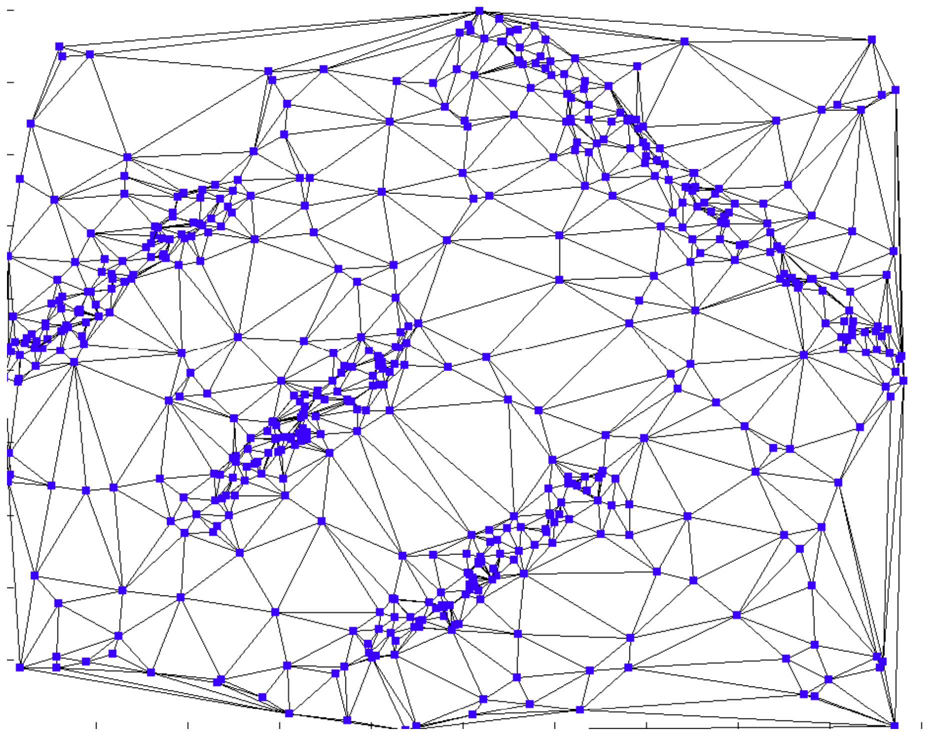
Spatial regularization

$$E(f) = \sum_{p} \| p, f_p \| + \sum_{pq \in N} \theta_{pq} \cdot [f_p \neq f_q]$$

If **multiple** models

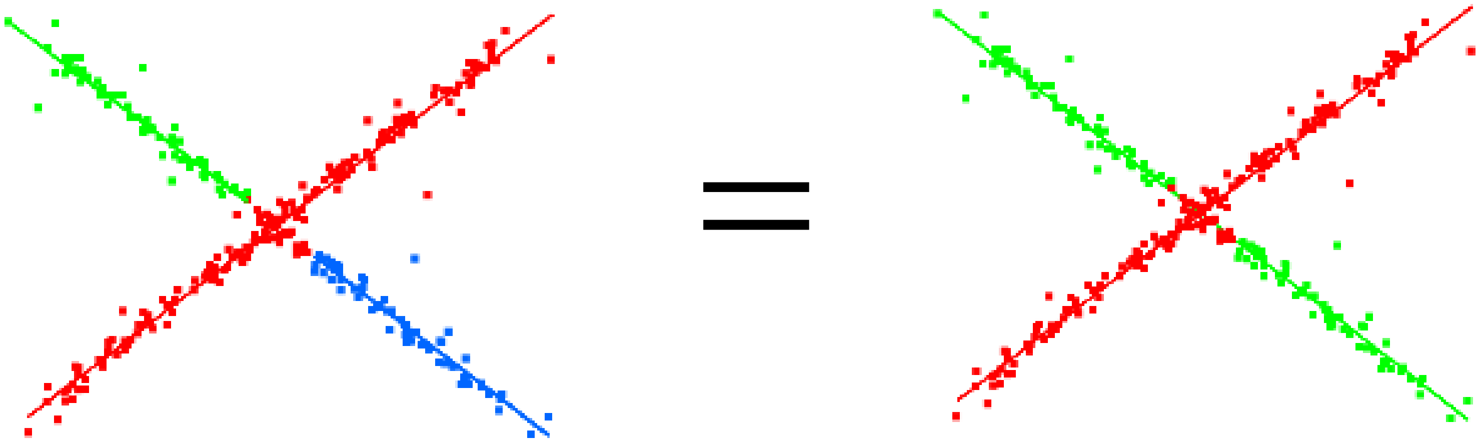
- assign different models (labels f_p) to every point p

-find optimal labeling $\boldsymbol{f} = \{ f_1, f_2, \ldots, f_n \}$



Spatial regularization

$$E(f) = \sum_{p} \| p, f_p \| + \sum_{pq \in N} \theta_{pq} \cdot [f_p \neq f_q]$$



Not enough!!!

Energy-based approach

$$E(f) = \sum_{p} \| p, f_p \| + \sum_{f \in \Lambda} h_f \cdot \delta_f(f)$$

If **multiple** models

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

- find optimal labeling
$\boldsymbol{f} = \{ f_1, f_2, \dots, f_n \}$

Λ - set of labels allowed at each point p

$$\delta_f(f) = \begin{cases} 1, & \exists p : f_p = f \\ 0, & otherwise \end{cases}$$

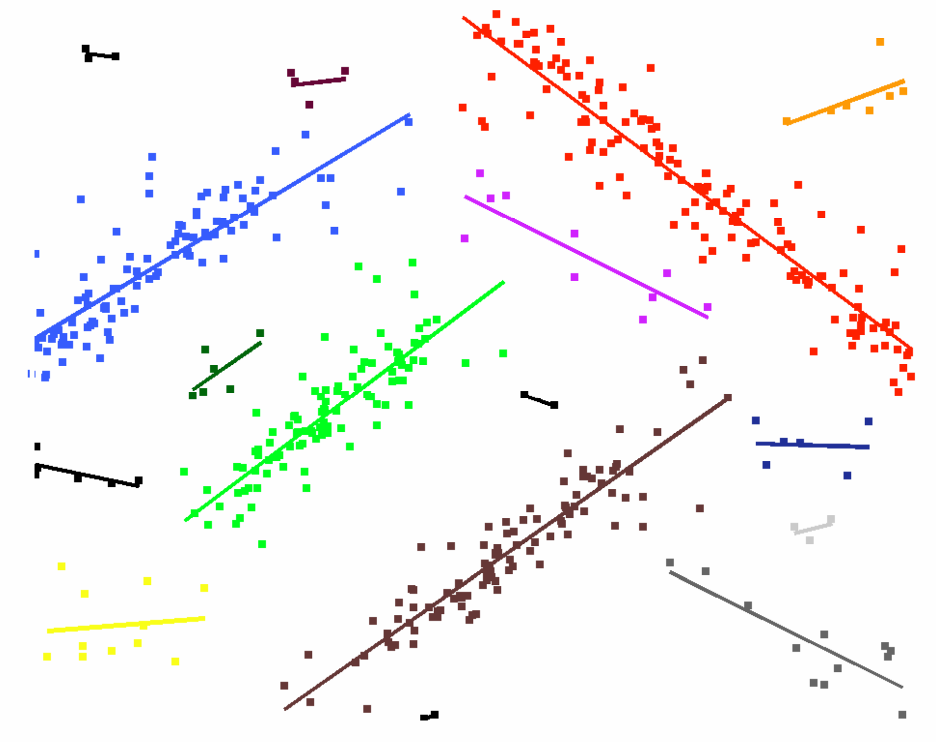
Energy-based approach

$$E(f) = \sum_p \| p, f_p \| + \sum_{(p,q)\in N} \theta_{pq} \cdot [f_p \neq f_q] \; + \sum_{f \in \Lambda} h_f \cdot \delta_f(f)$$

If **multiple** models

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

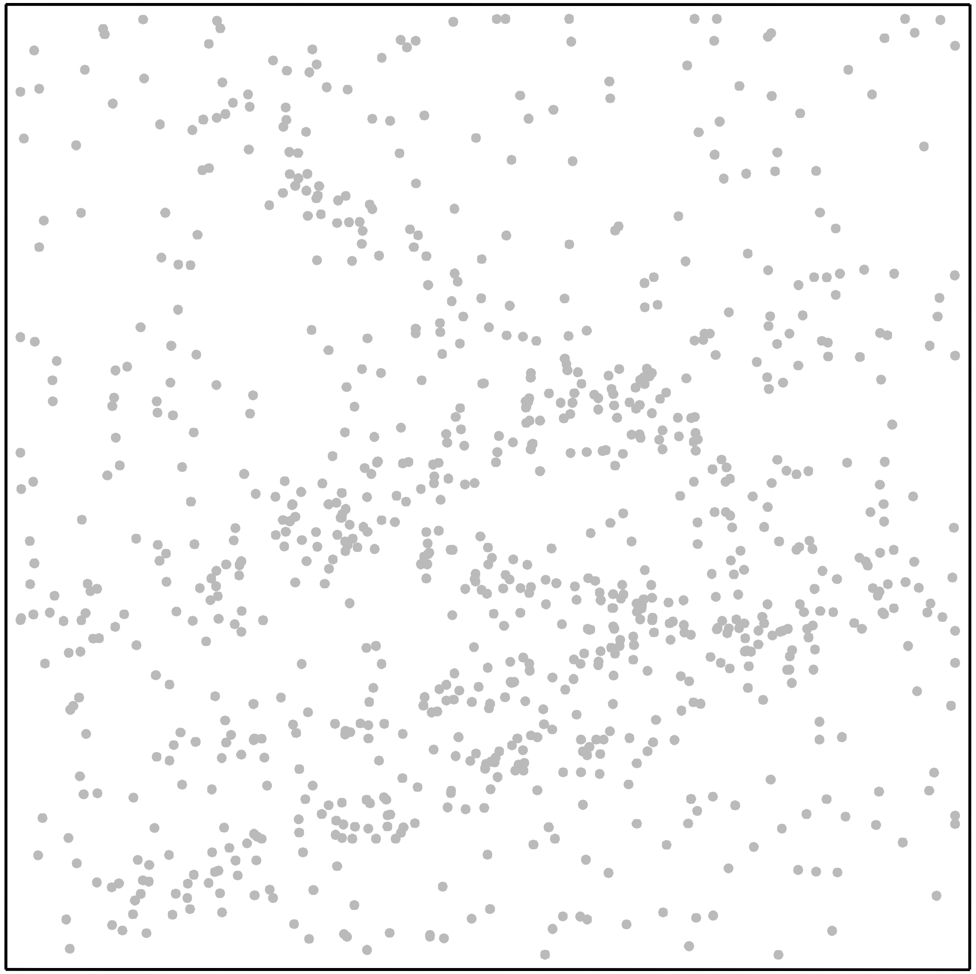
-find optimal labeling
$\boldsymbol{f} = \{ f_1, f_2, \ldots, f_n \}$



Practical problem: number of potential labels (models) is huge, how are we going to use a-expansion?

PEARL

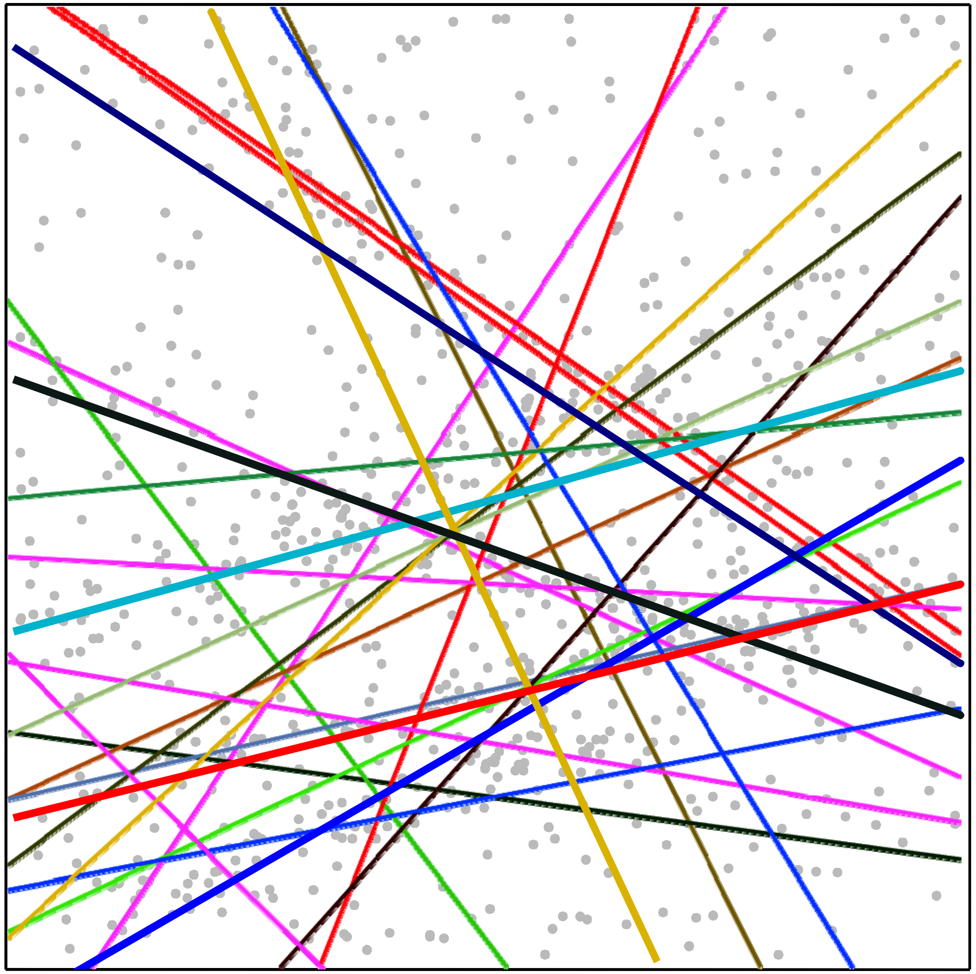
Propose
Expand
And
Reestimate
Labels



data points

PEARL

Propose
Expand
And
Reestimate
Labels



data points + randomly sampled models

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

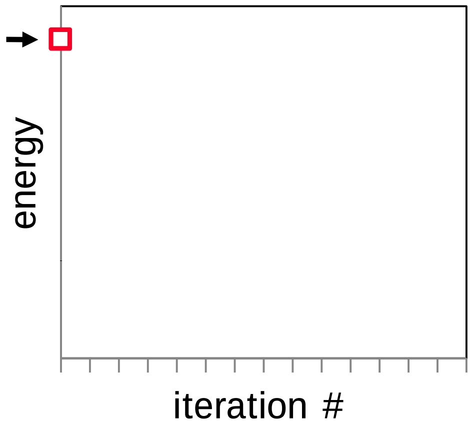
Λ

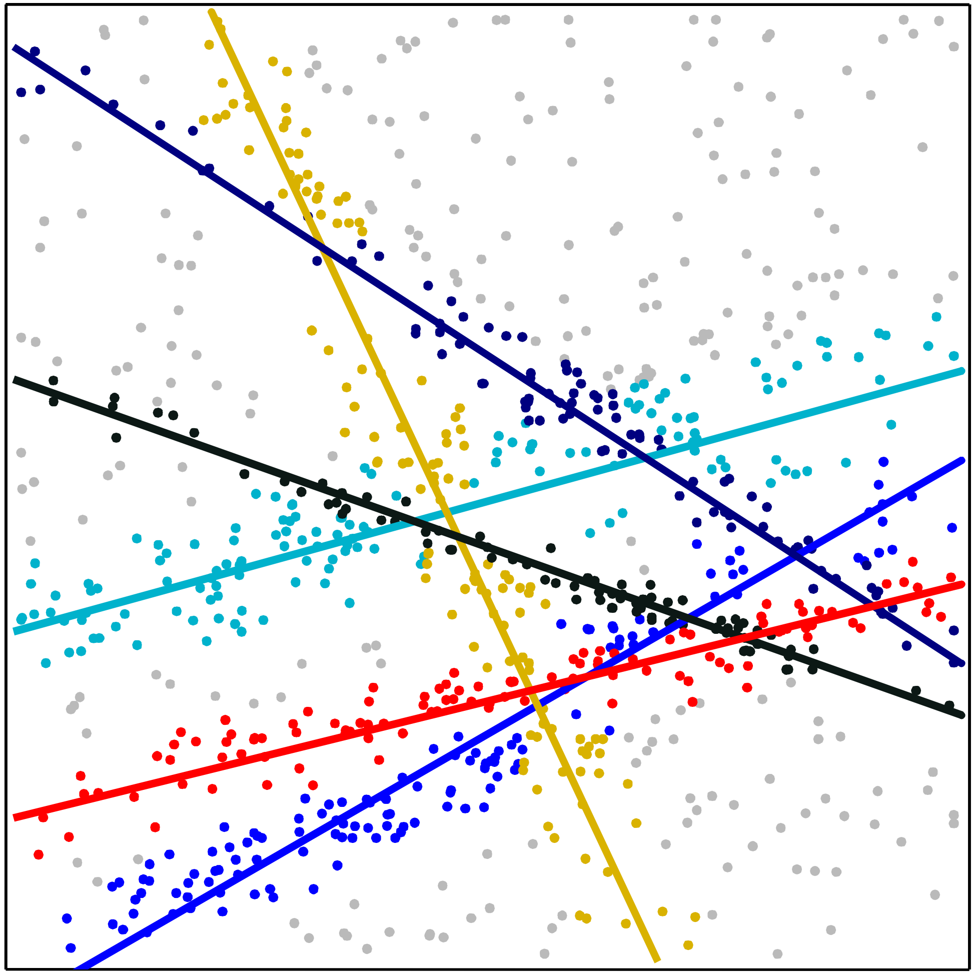
$$E(f) = \sum_{p} \| p, f_p \| + \sum_{pq \in N} \theta_{pq} \cdot [f_p \neq f_q]$$

$$f_p \in \Lambda$$

PEARL

Propose
Expand
And
Reestimate
Labels





iteration 1: optimize labeling f

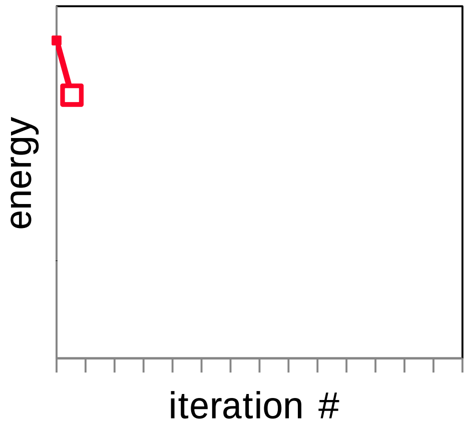
a-expansion:
minimize $E(f)$
over a fixed
set of labels
Λ

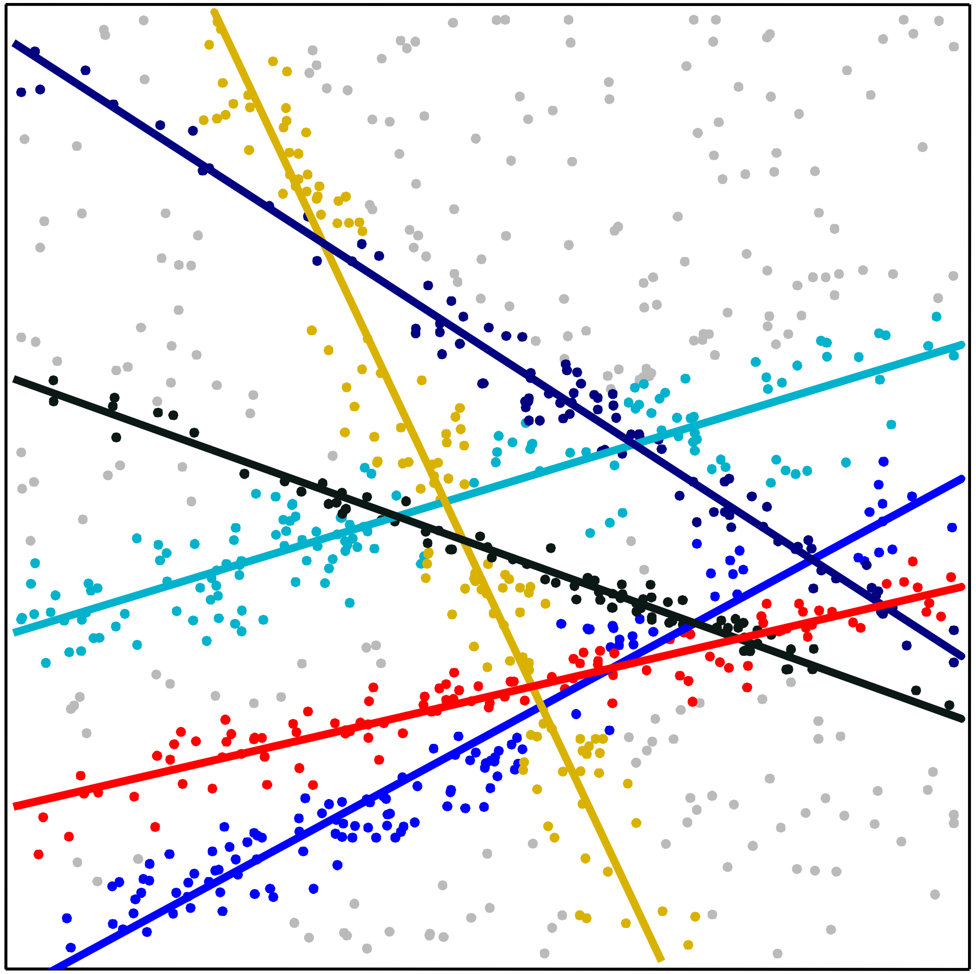
$$E(f) = \sum_p \| p, f_p \| + \sum_{pq \in N} \theta_{pq} \cdot [f_p \neq f_q]$$

fixed

PEARL

Propose
Expand
And
Reestimate
Labels





iteration 1: reestimate models

reestimating labels in Λ for given inliers

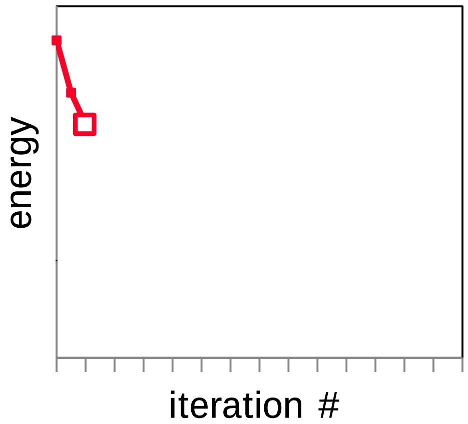
minimizing the first term of energy $E(l)$

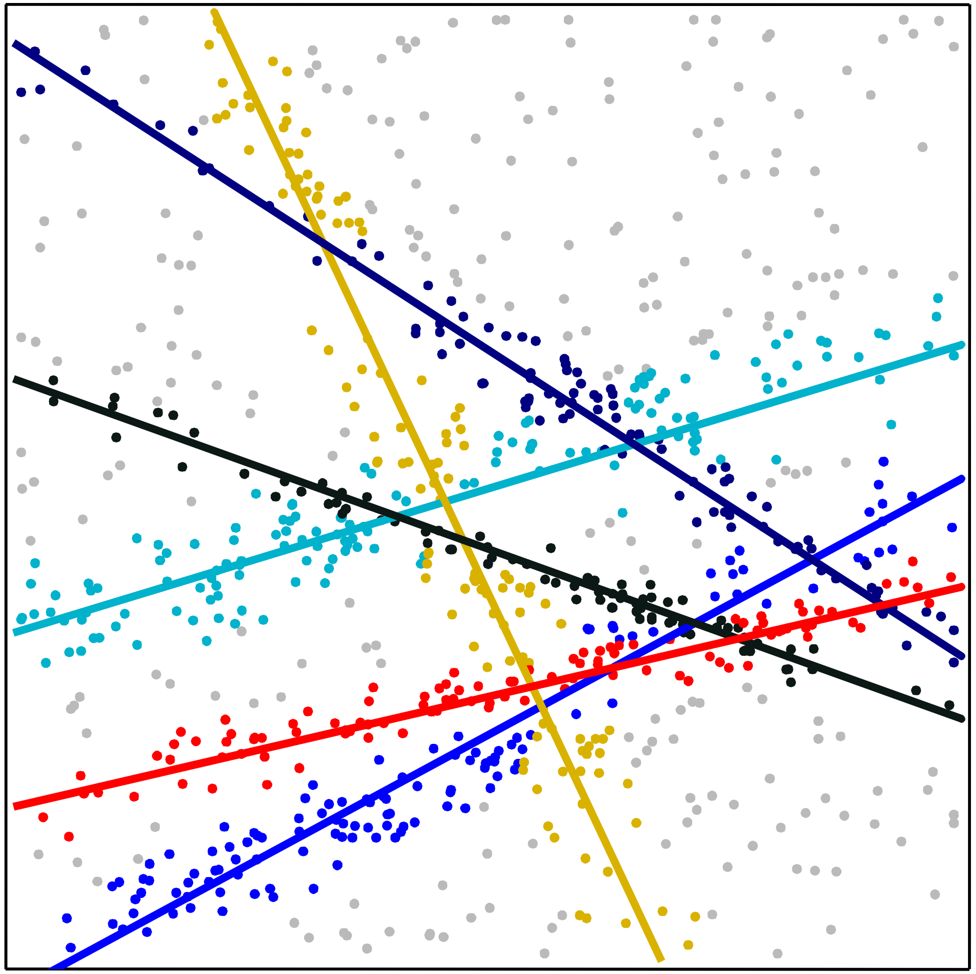
$$E(f) = \sum_{p} \| p, f_p \| + \sum_{pq \in N} \theta_{pq} \cdot [f_p \neq f_q]$$

PEARL

$$f_p \in \Lambda$$

Propose
Expand
And
Reestimate
Labels





iteration 2: optimize labeling f

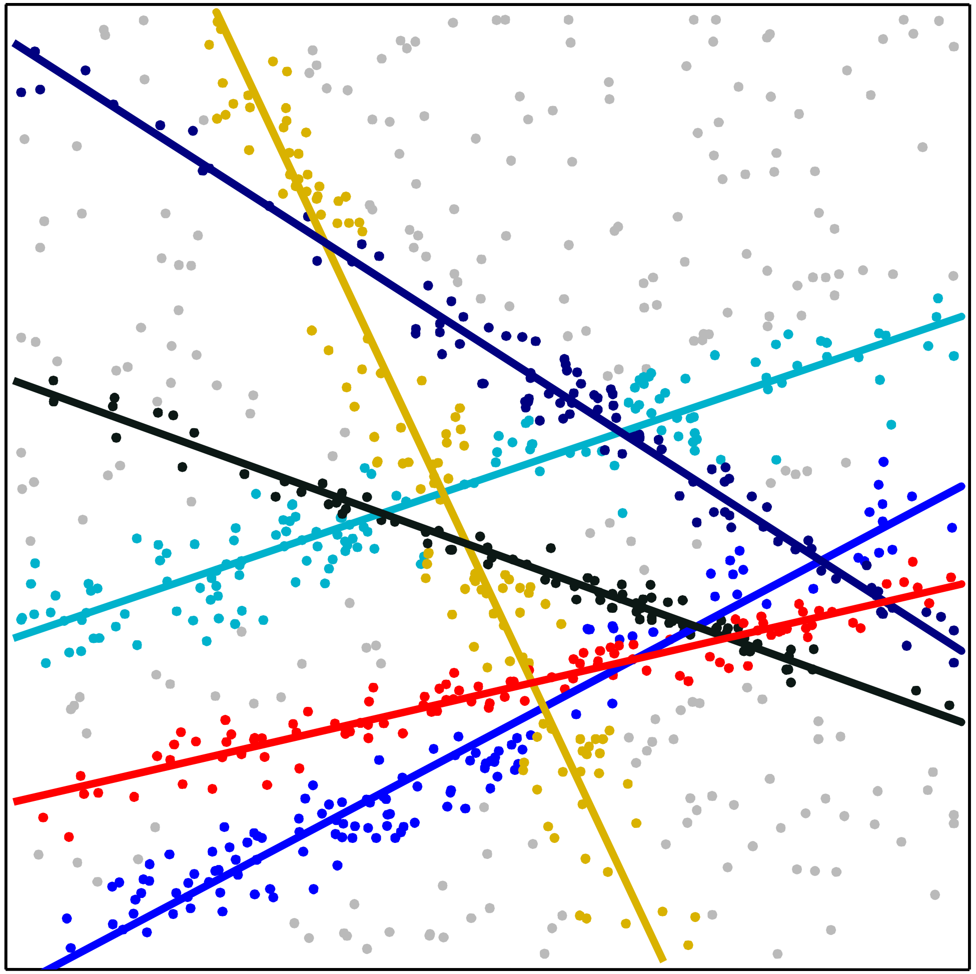
a-expansion:
minimize *E(f)*
over a fixed
set of labels
Λ

$$E(f) = \sum_p \| p, f_p \| + \sum_{pq \in N} \theta_{pq} \cdot [f_p \neq f_q]$$

fixed

PEARL

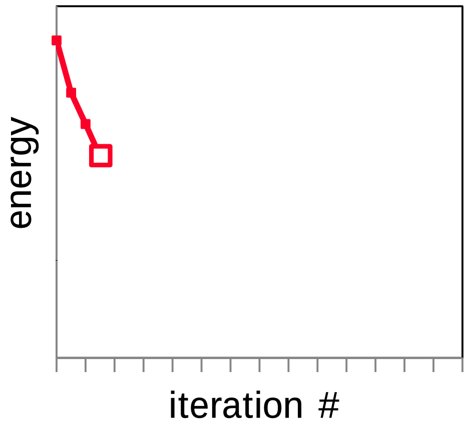
Propose
Expand
And
Reestimate
Labels



iteration 2: reestimate models

reestimating labels in for given inliers

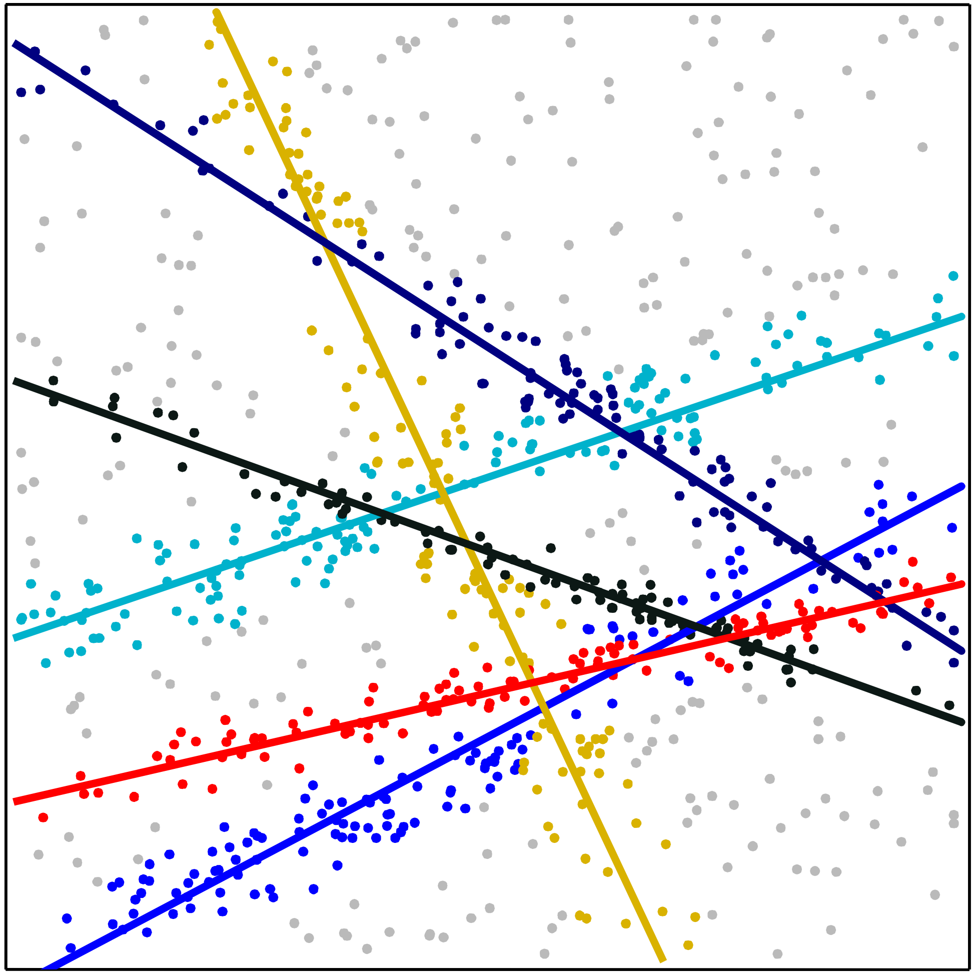
minimizing the first term of energy $E(f)$



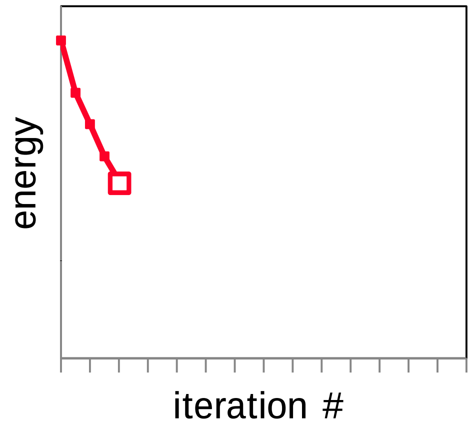
$$E(f) = \sum_{p} \| p, f_p \| + \sum_{pq \in N} \theta_{pq} \cdot [f_p \neq f_q]$$

PEARL

Propose
Expand
And
Reestimate
Labels



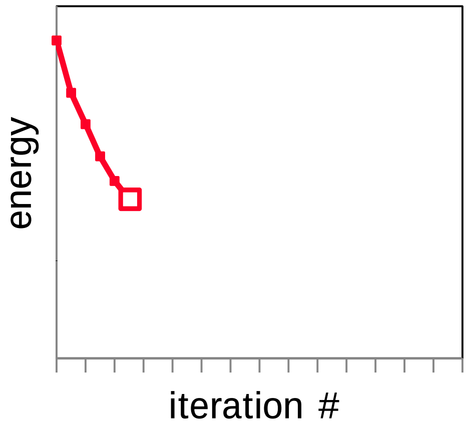
iteration 3: optimize labeling f

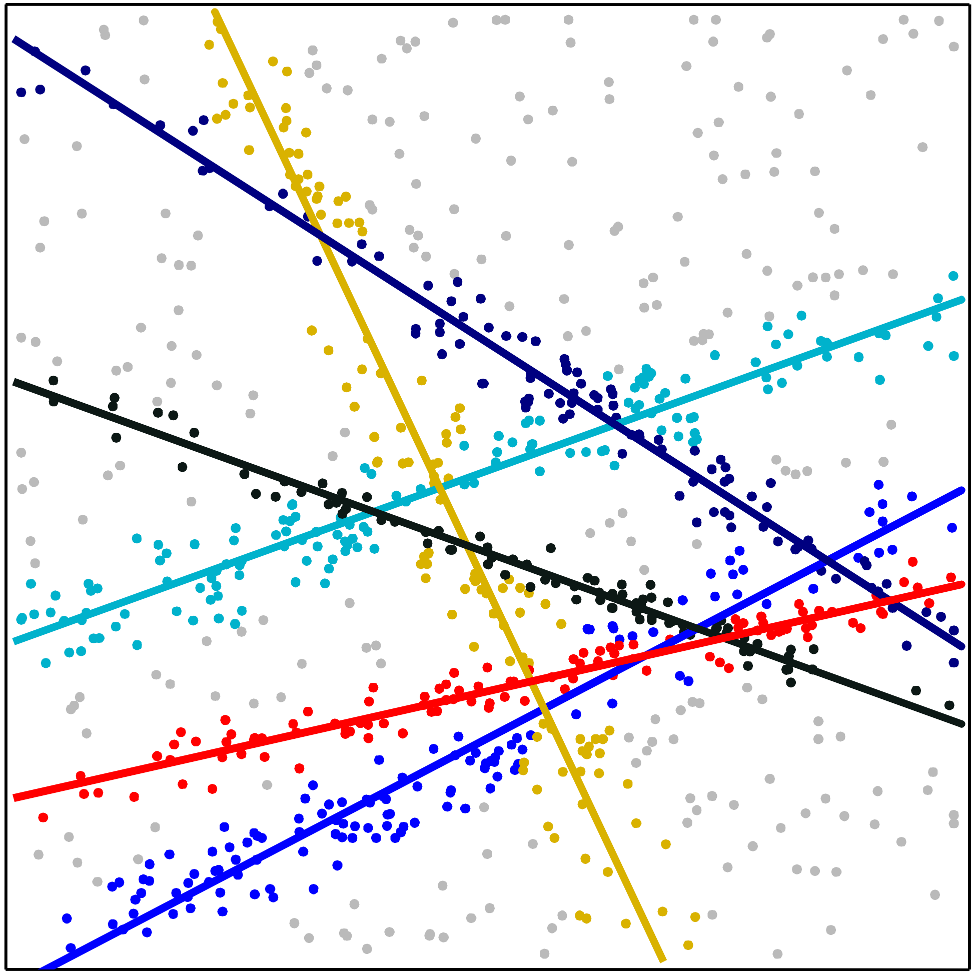


$$E(f) = \sum_{p} \| p, f_p \| + \sum_{pq \in N} \theta_{pq} \cdot [f_p \neq f_q]$$

PEARL

Propose
Expand
And
Reestimate
Labels

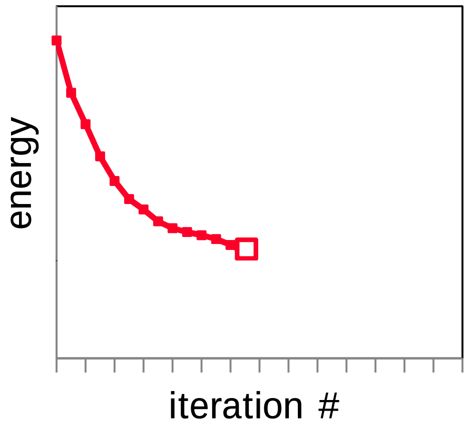


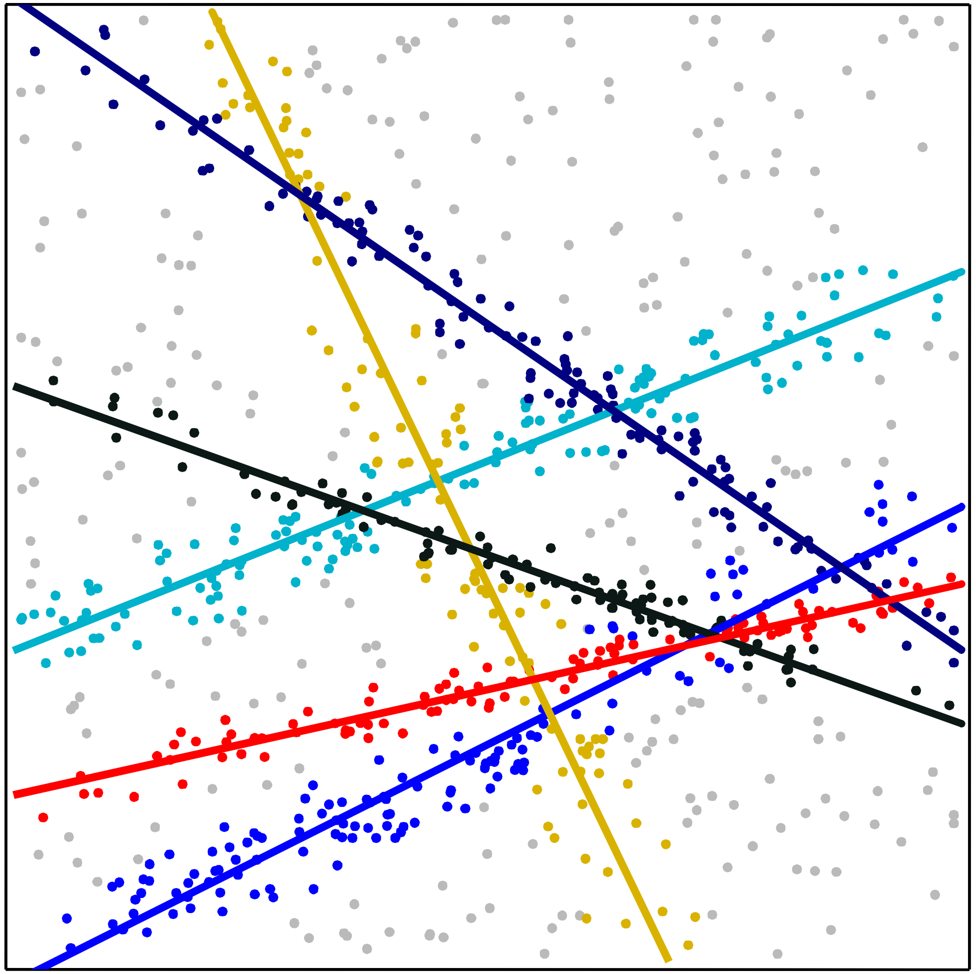


iteration 3: reestimate models

$$E(f) = \sum_{p} \| p, f_p \| + \sum_{pq \in N} \theta_{pq} \cdot [f_p \neq f_q]$$

PEARL

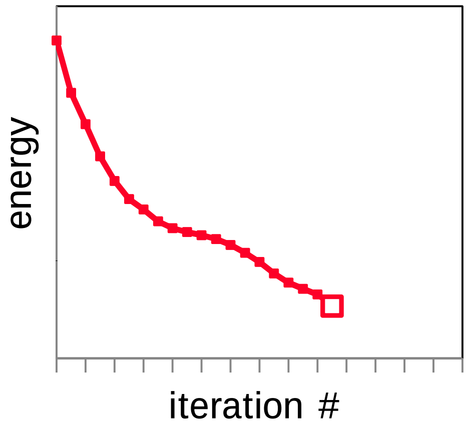


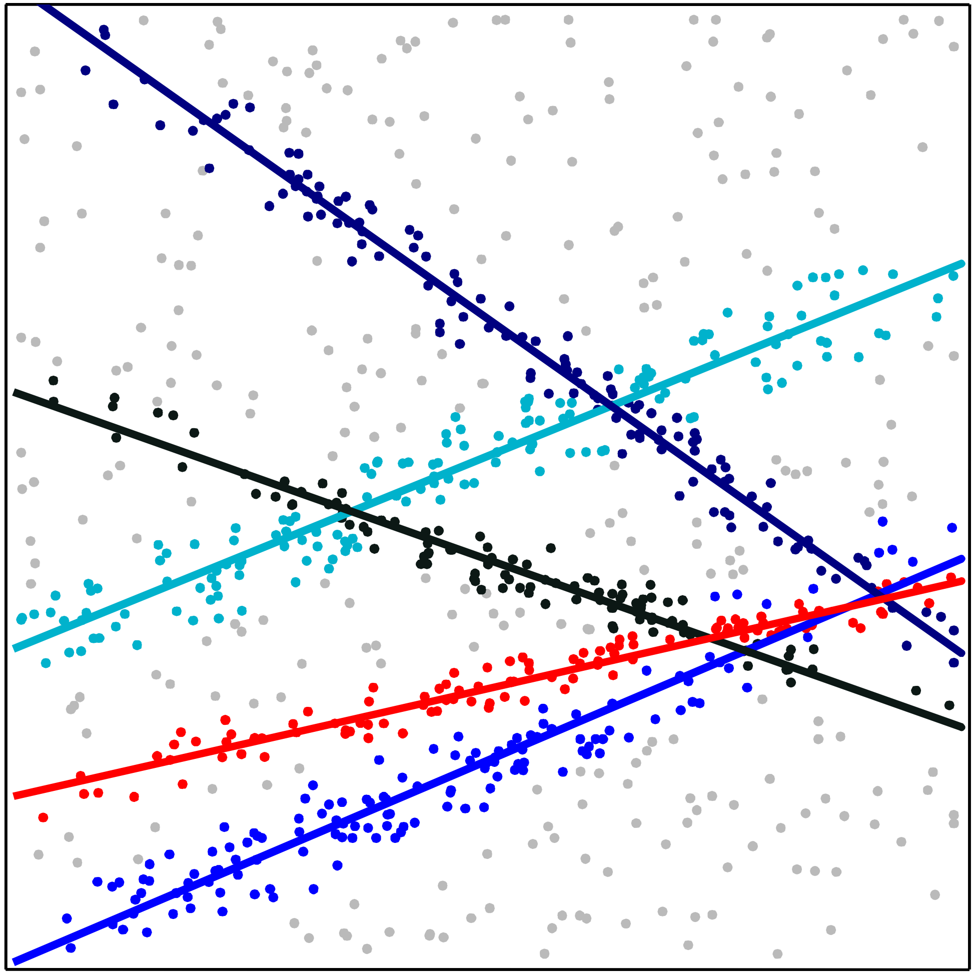


iteration 7...

$$E(f) = \sum_{p} \| p, f_p \| + \sum_{pq \in N} \theta_{pq} \cdot [f_p \neq f_q]$$

PEARL

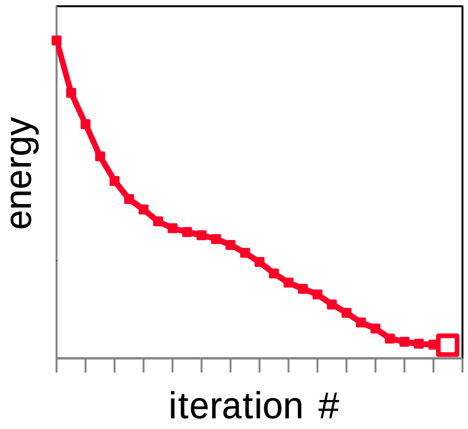


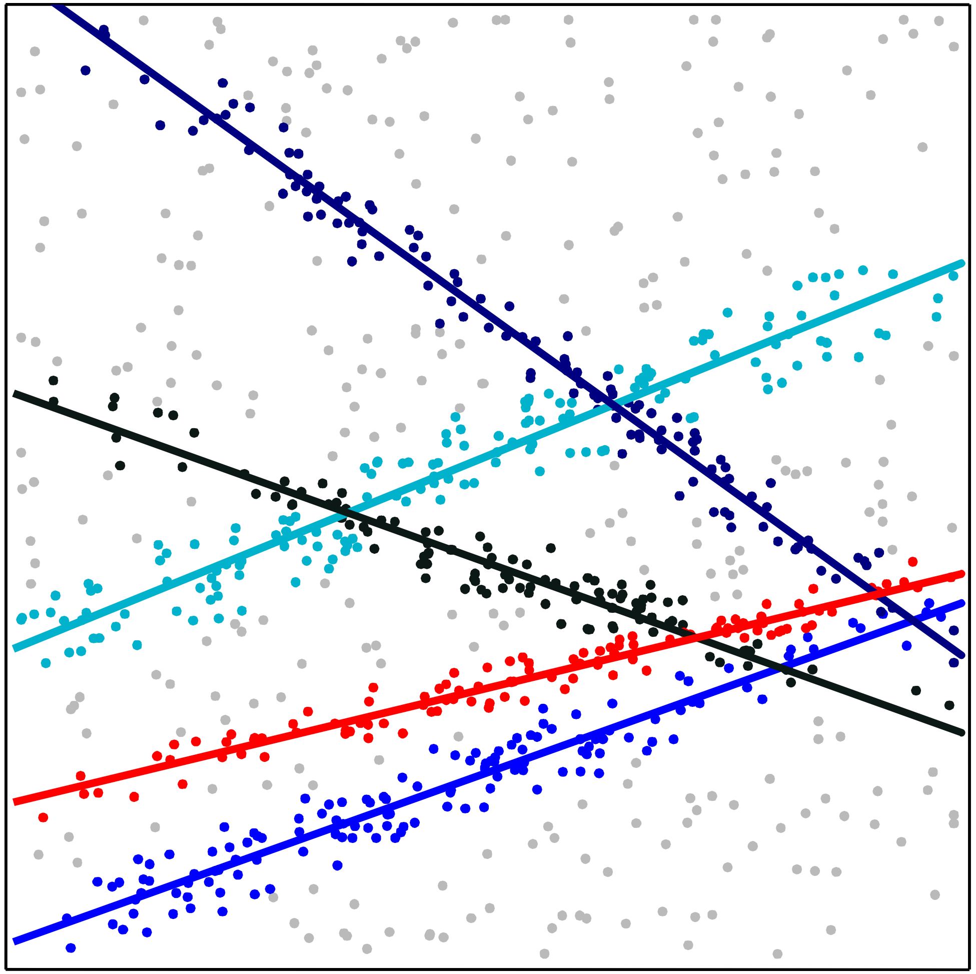


iteration 10...

$$E(f) = \sum_{p} \| p, f_p \| + \sum_{pq \in N} \theta_{pq} \cdot [f_p \neq f_q]$$

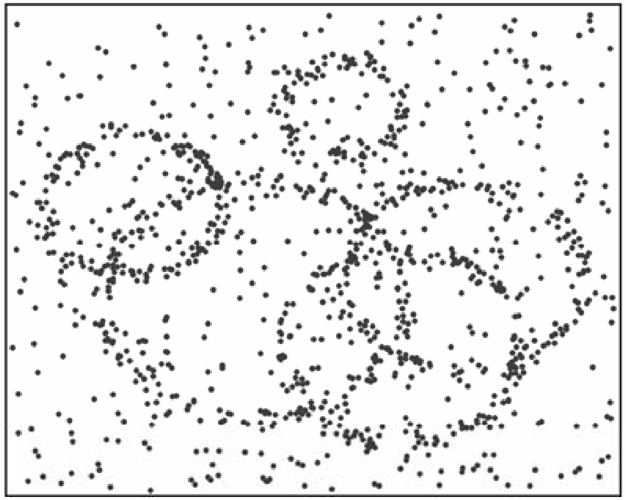
PEARL

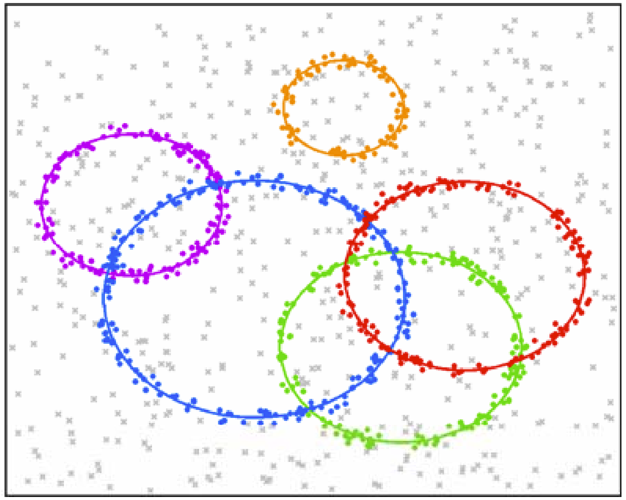




iteration 15... converged.

Fitting circles





regularization with label costs only

Here spatial regularization does not work well

(unsupervised image segmentation)

Fitting color models

(unsupervised image segmentation)

Fitting color models

(c) Spatial regularity + label costs

Zhu and Yuille 96
used continuous
variational formulation
(gradient discent)

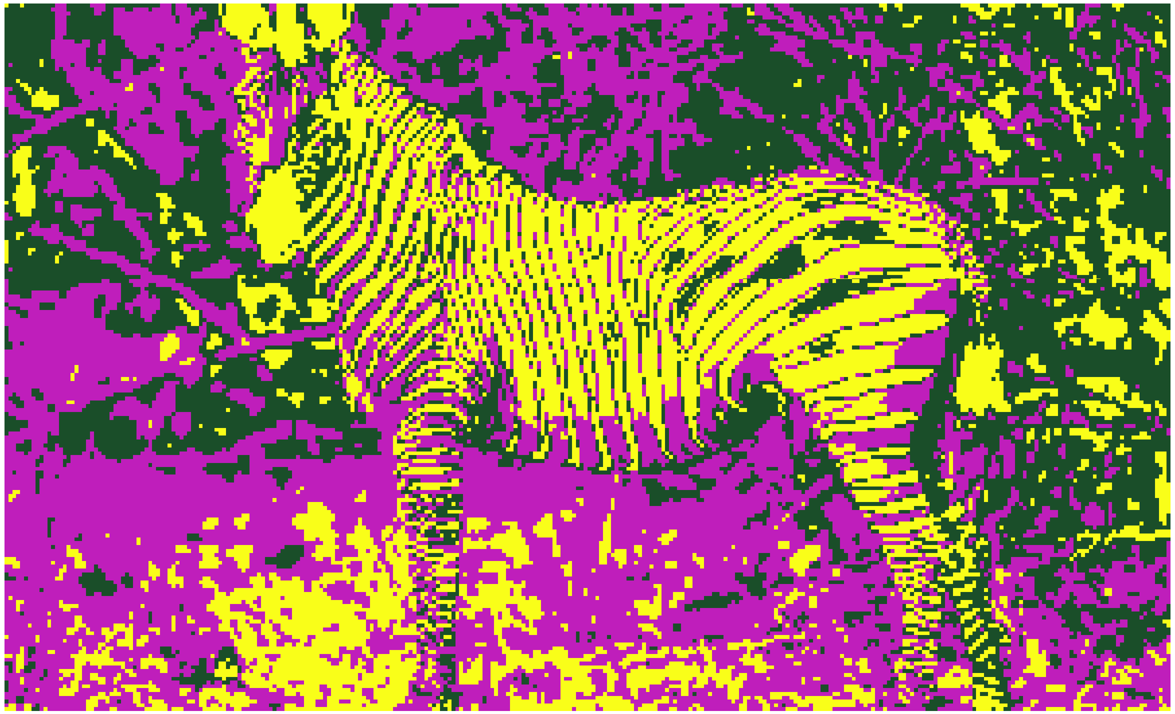
(unsupervised image segmentation)

Fitting color models

(b) Spatial regularity only [Zabih&Kolmogorov CVPR 04]

(unsupervised image segmentation)

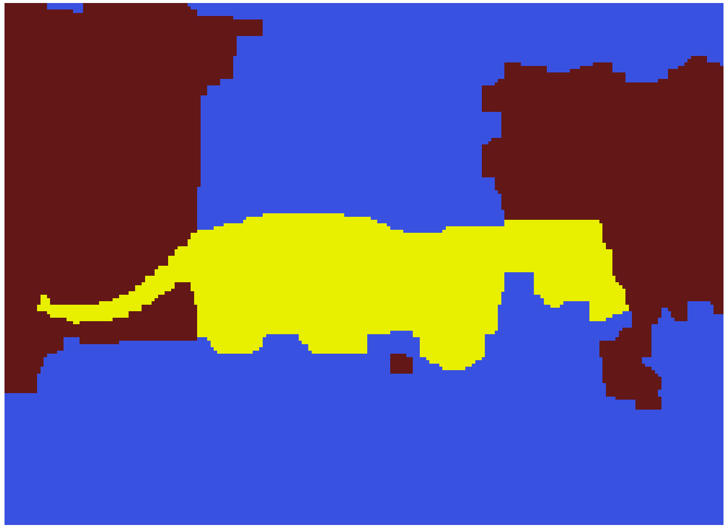
Fitting color models



(a) Label costs only [Li, CVPR 2007]

(unsupervised image segmentation)

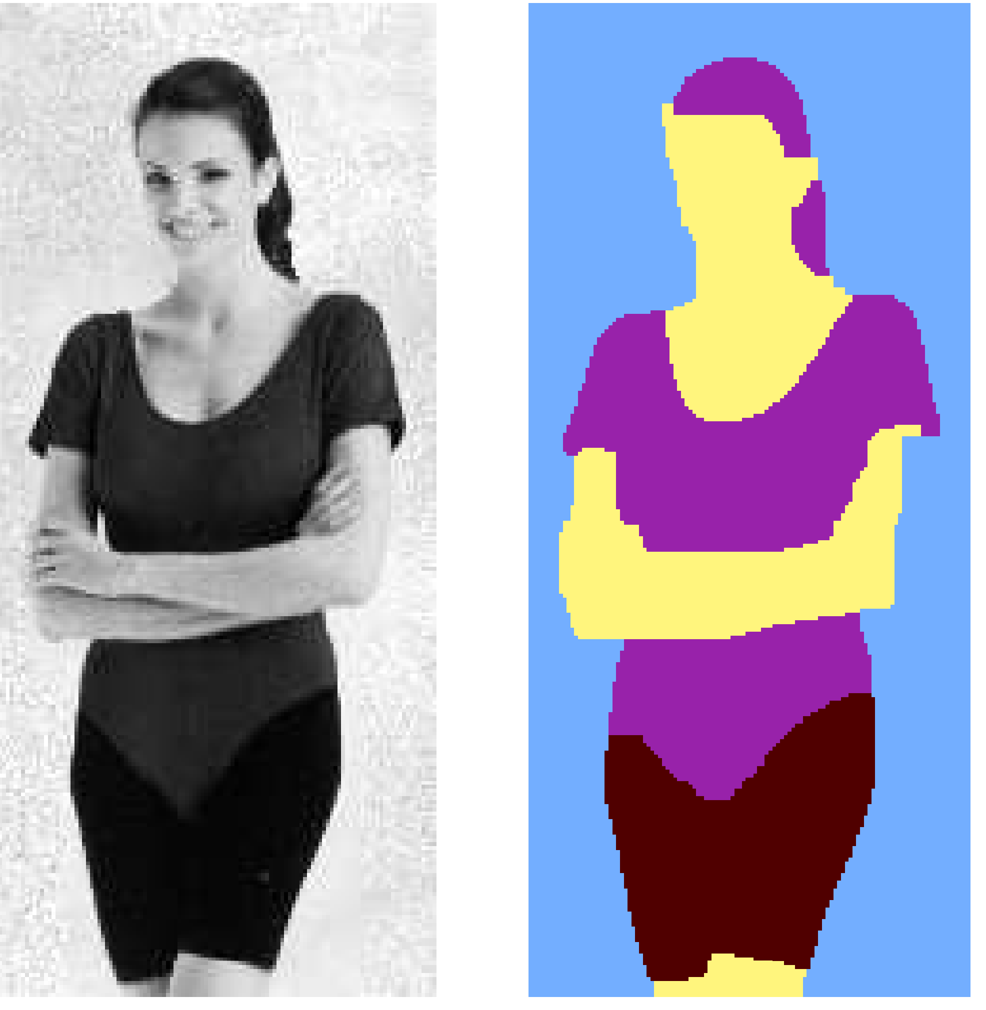
Fitting color models



Spatial regularity + label costs

(unsupervised image segmentation)

Fitting color models



Spatial regularity + label costs

(unsupervised image segmentation)
Fitting color models

Spatial regularity + label costs

(unsupervised image segmentation)

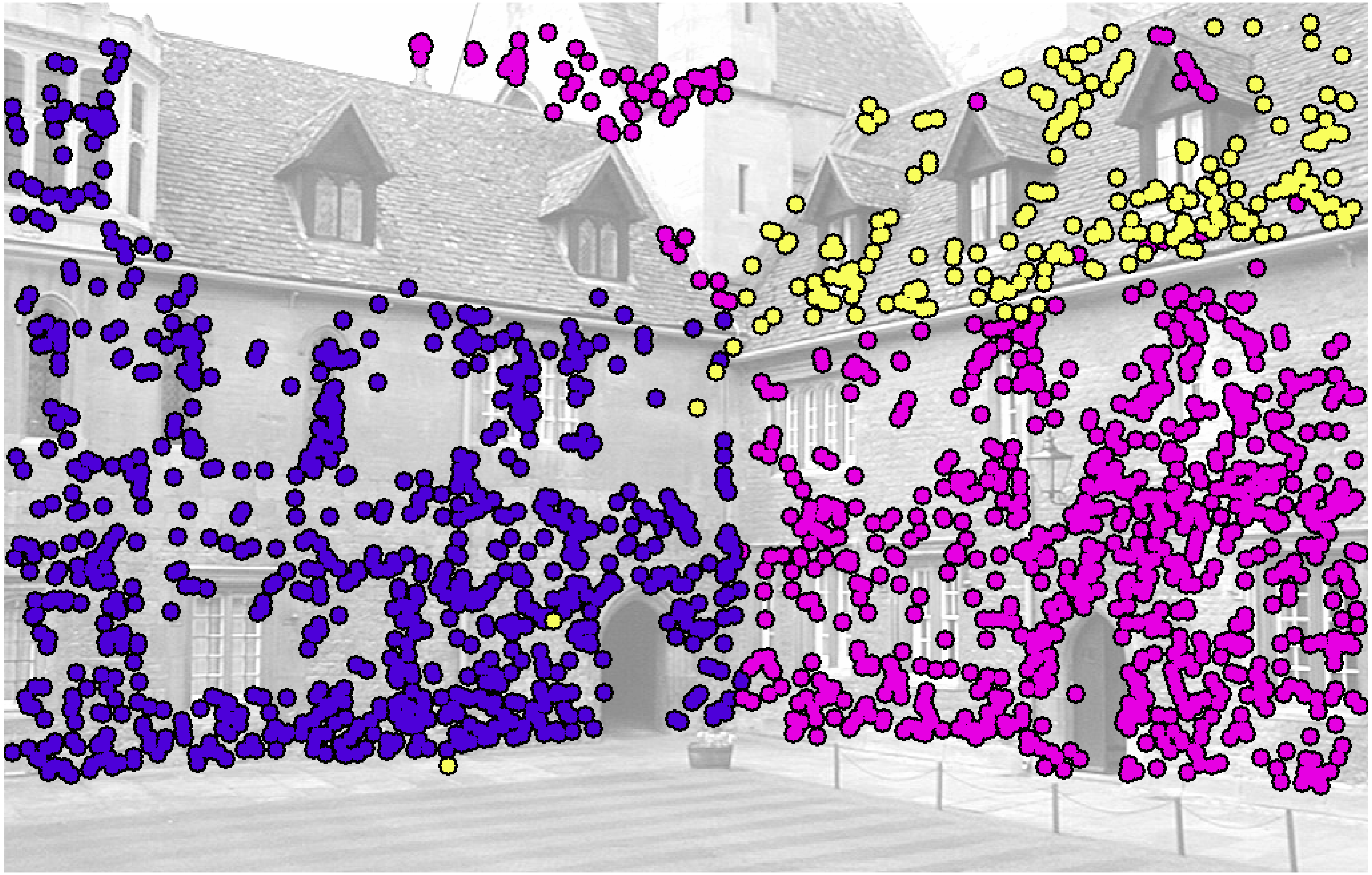
Fitting color models

Spatial regularity + label costs

Fitting planes (homographies)

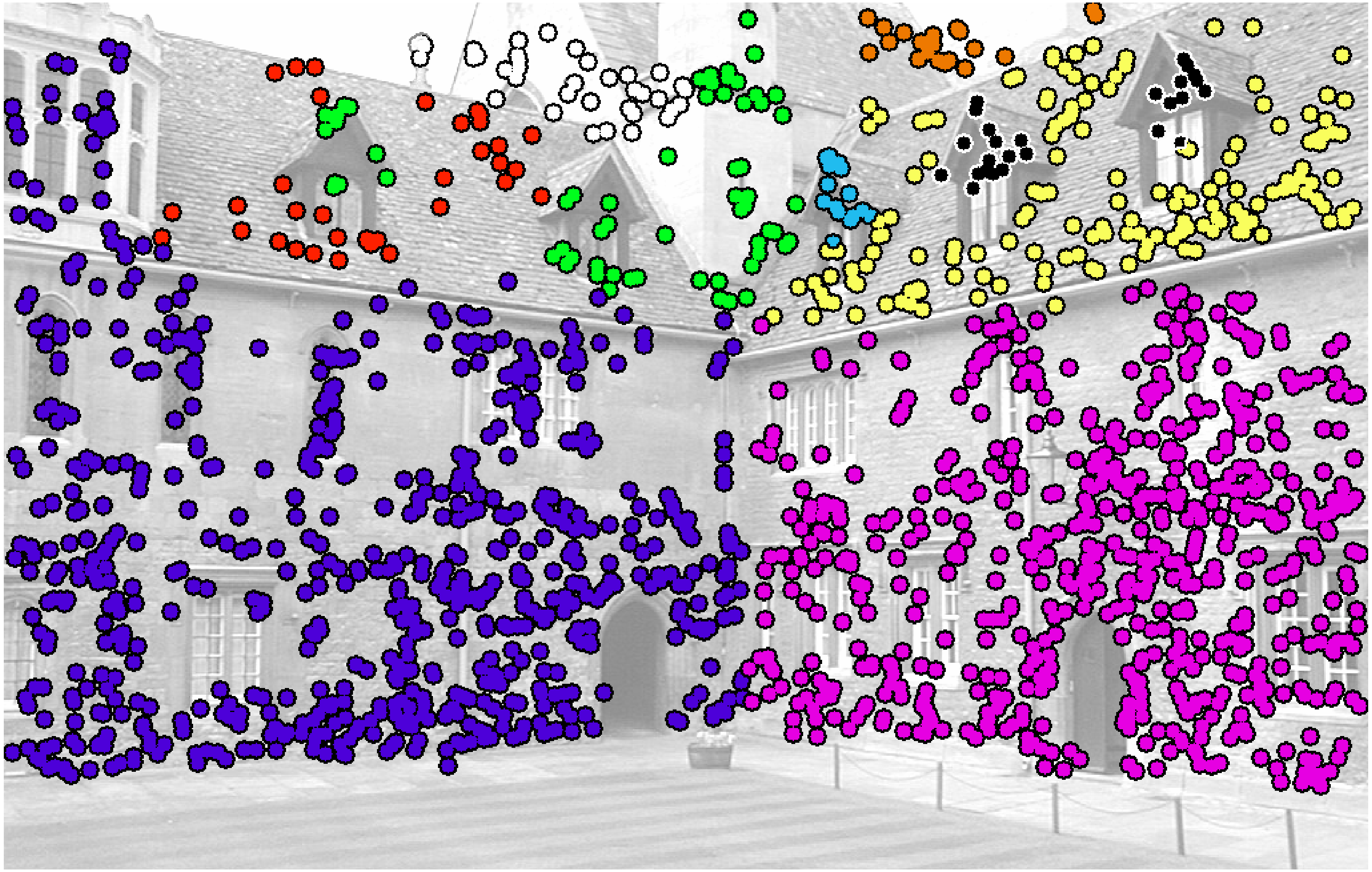
Original image (one of 2 views)

Fitting planes (homographies)



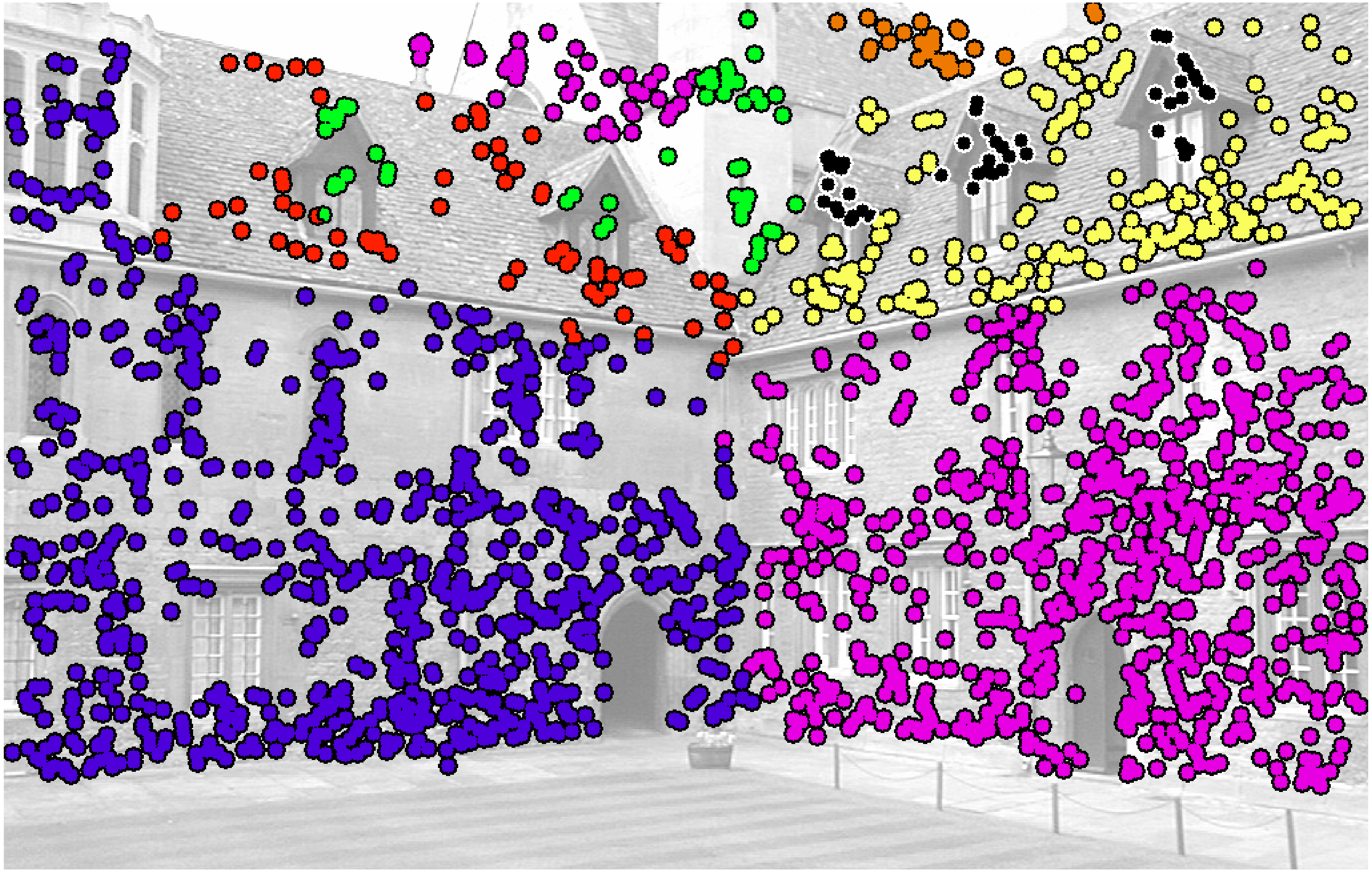
(a) Label costs only

Fitting planes (homographies)



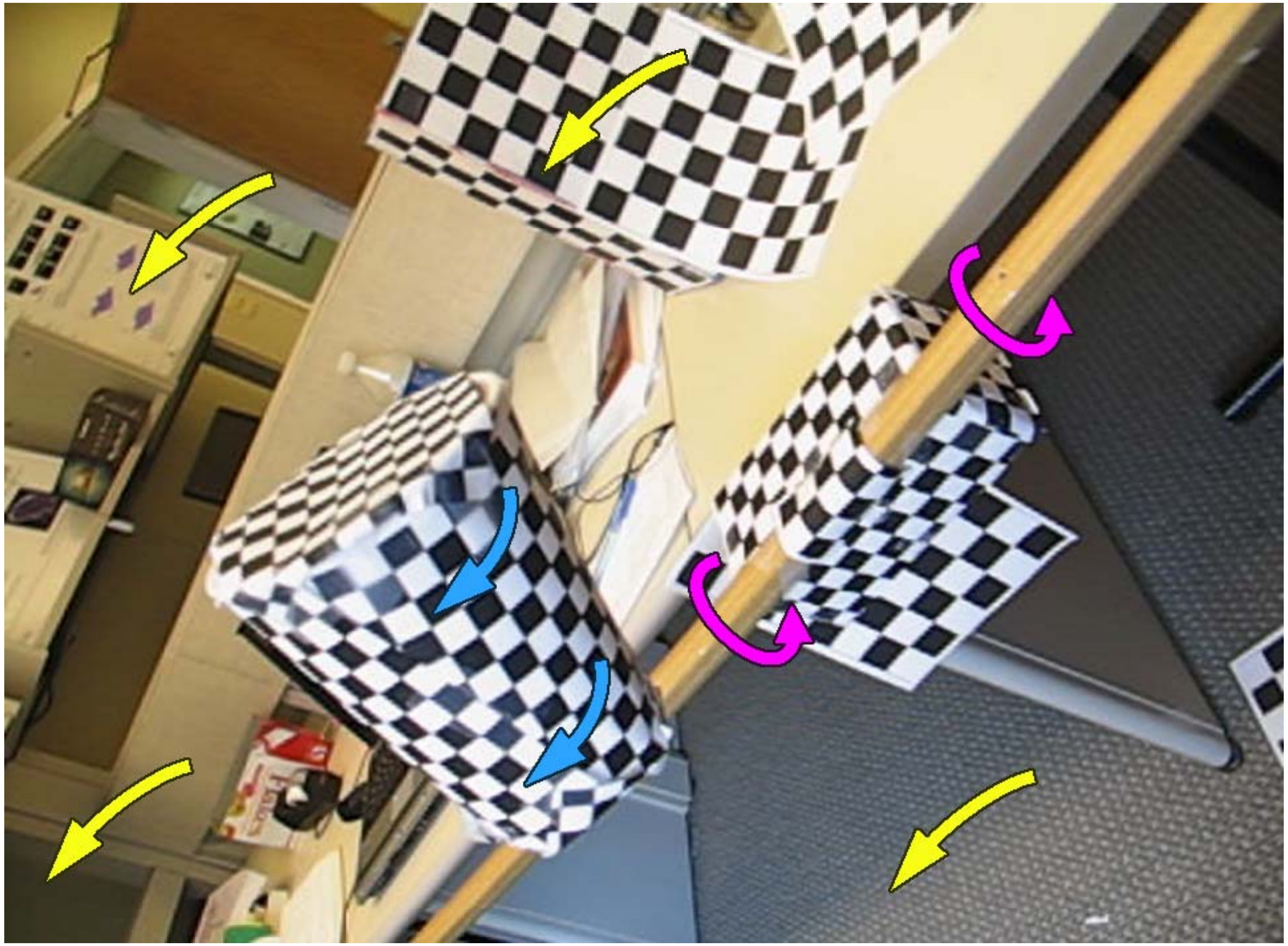
(b) Spatial regularity only

Fitting planes (homographies)



(c) Spatial regularity + label costs

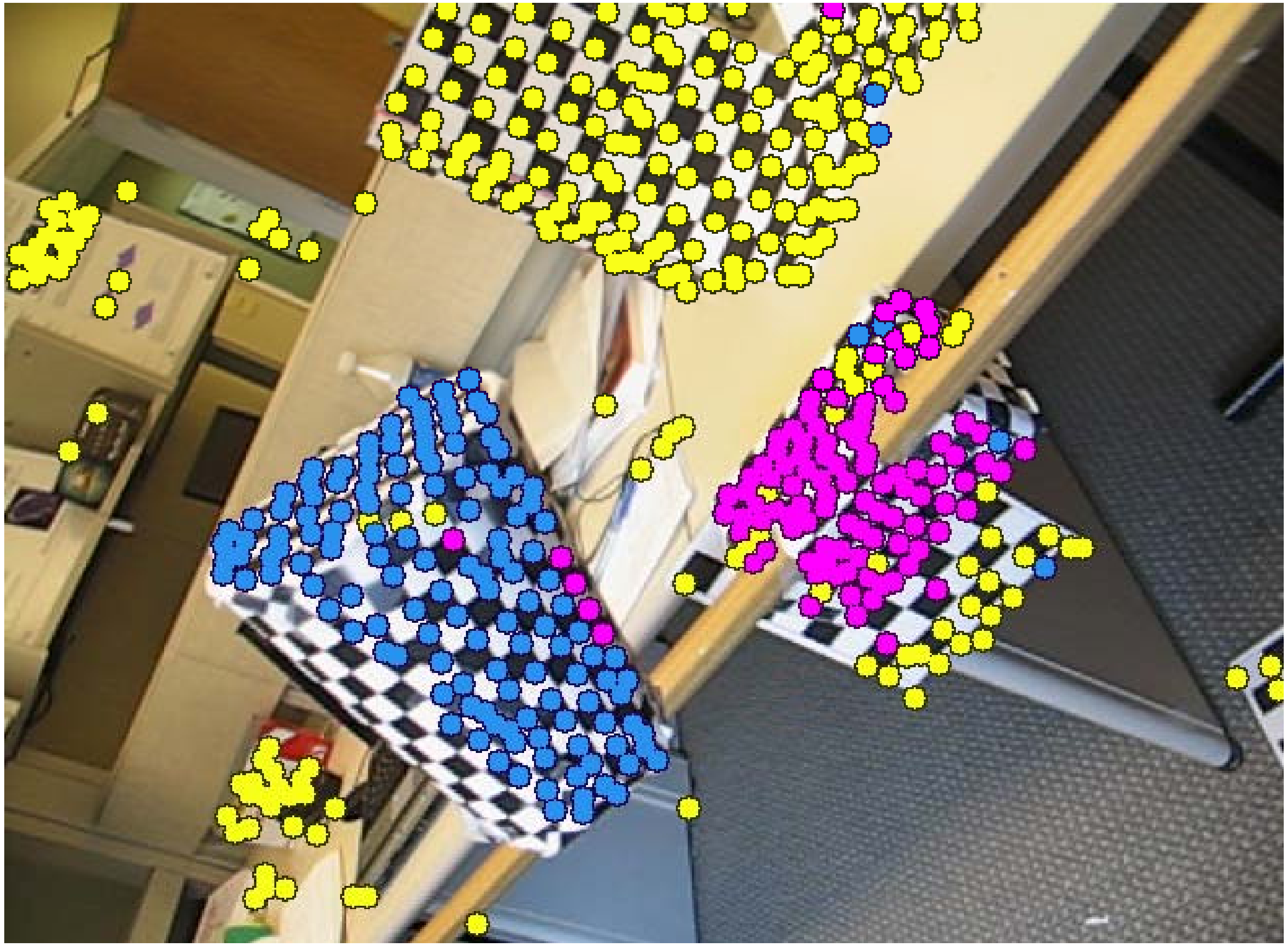
(rigid) Motion Estimation



3 motions

Original image [Rene Vidal]

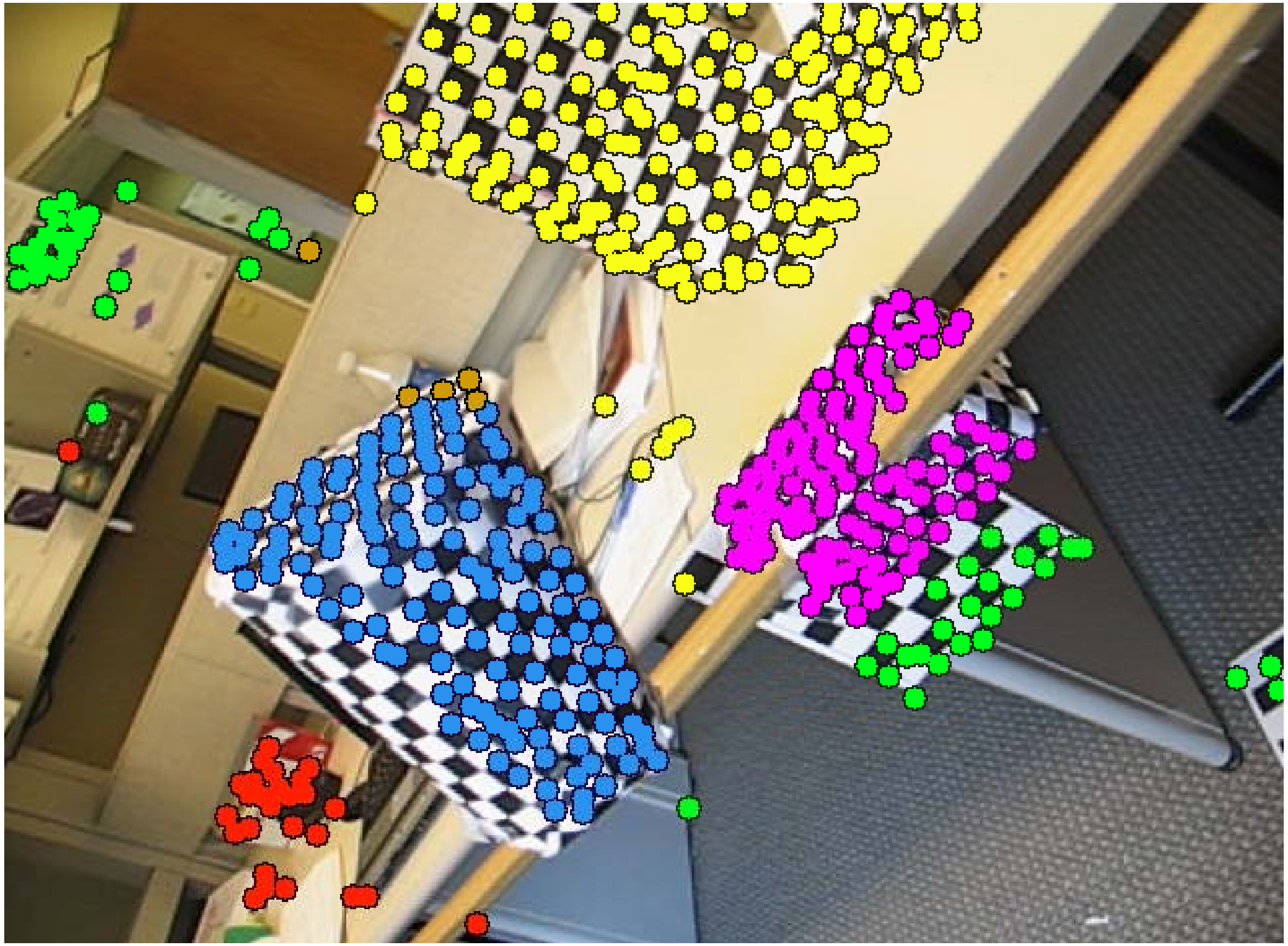
(rigid)
Motion Estimation



(a) Label costs only

3 motions

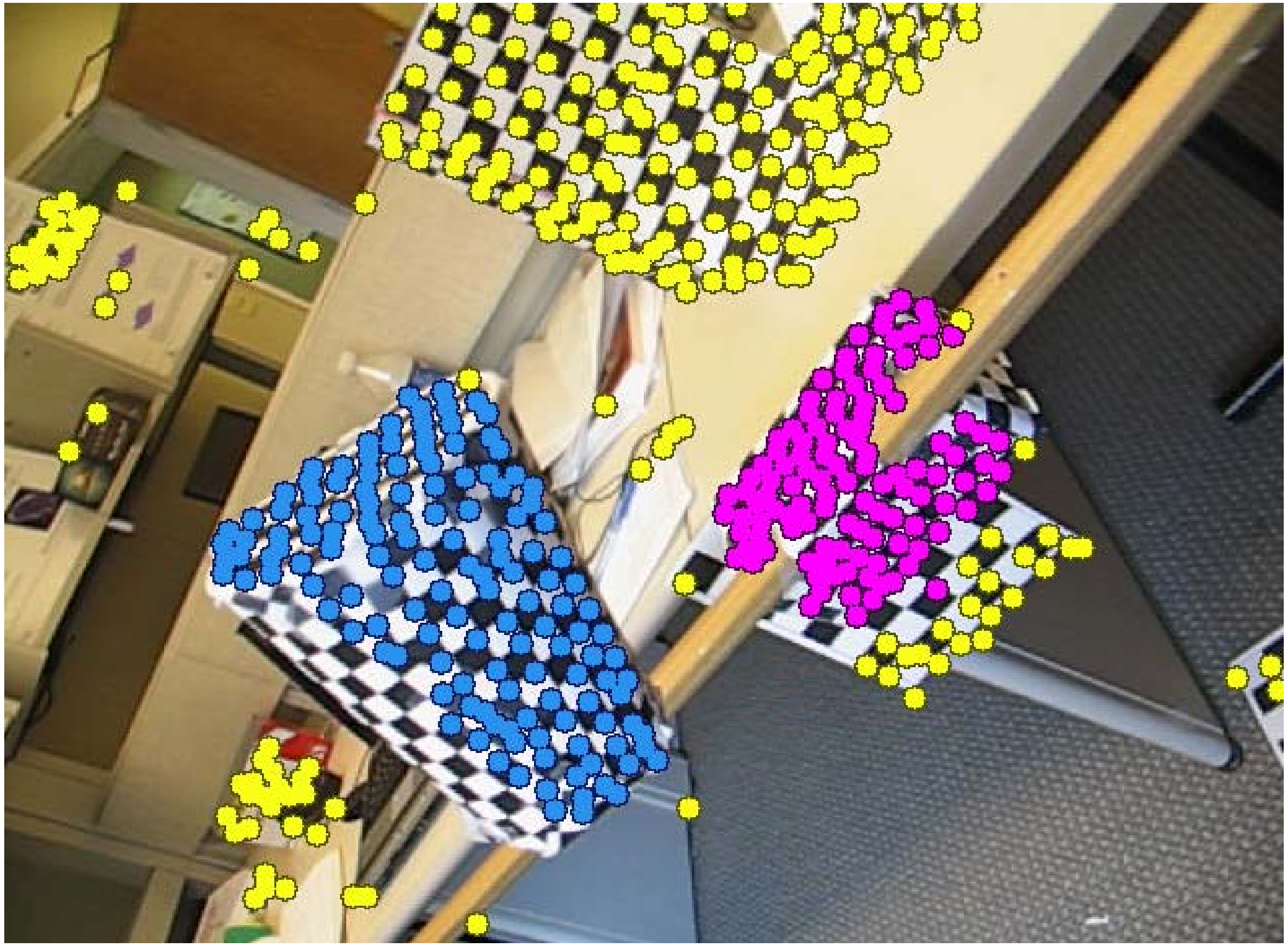
(rigid)
Motion Estimation



7 motions

(b) Spatial regularity only

(rigid)
Motion Estimation



(c) Spatial regularity + label costs

3 motions

(rigid)
Motion Estimation

Multi-Label Energy Formulation

Input: Set of data points P ← pixels, features, matches,...

Set of candidate labels Λ ← objects, motions, homographies,...

Goal: Labeling f that minimizes energy E

$$\min_{labeling} \{ \text{ data costs } + \text{ smooth costs } + \text{ label costs } \}$$

$$E(f) = \sum_{p} D_p(f_p) + \sum_{(p,q)} V_{pq}(f_p, f_q) + \sum_{l} h_l \delta(\exists p : f_p = l)$$

α-expansion

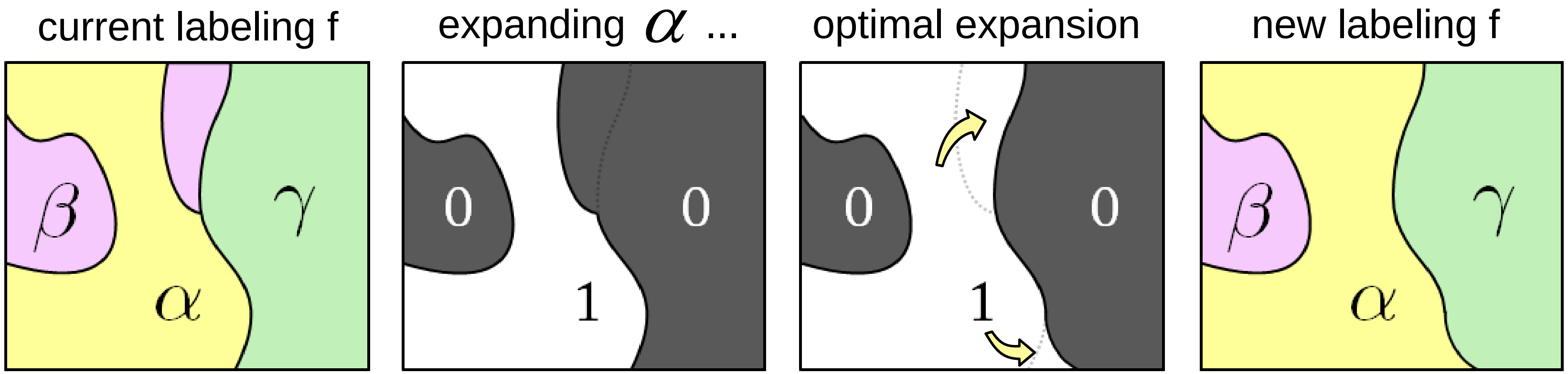
$$E(f) = \sum_p D_p(f_p) + \sum_{(p,q)} V_{pq}(f_p, f_q)$$

- NP-hard in general when $|\Lambda| \geq 3$
- α-expansion is standard algorithm
  - finds local optimum w.r.t. “expansions”
  - optimality guarantees
  - fast & effective in practice

Boykov, Veksler, Zabih. *PAMI* 2001

α -expansion

- α -expansion main idea:
  - convert multi-label problem into sequence of *binary* problems
  - choose label α , and only let it “expand”



Deriving the Graph Construction

Let f be current labeling

$\mathbf{x}$ be labeling of binary subproblem

f^{α} be the labeling induced by $\mathbf{x}$

$$x_p = 0 \iff f_p^{\alpha} = f_p$$

$$x_p = 1 \iff f_p^{\alpha} = \alpha$$

keep current label

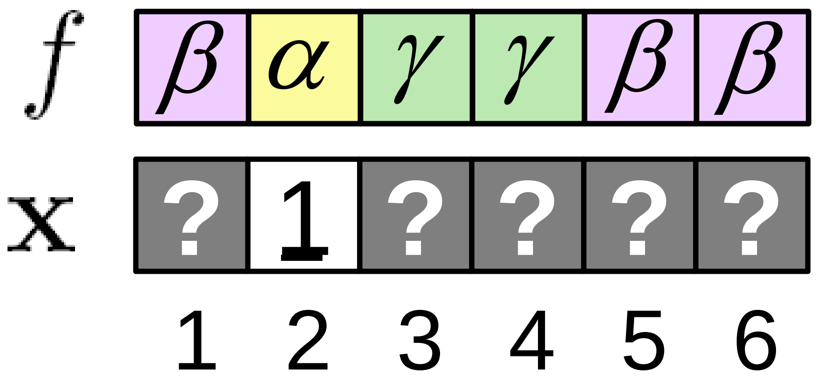
switch to α

Deriving the Graph Construction

- How to add cost h_β to binary problem?

$$E_h^\alpha(x) = E^\alpha(x) + h_\beta \underbrace{(1 - x_1 x_5 x_6)}$$

indicator function $\delta_\beta(f^\alpha) = \begin{cases} 1 & \exists p : f_p^\alpha = \beta \\ 0 & \text{otherwise.} \end{cases}$



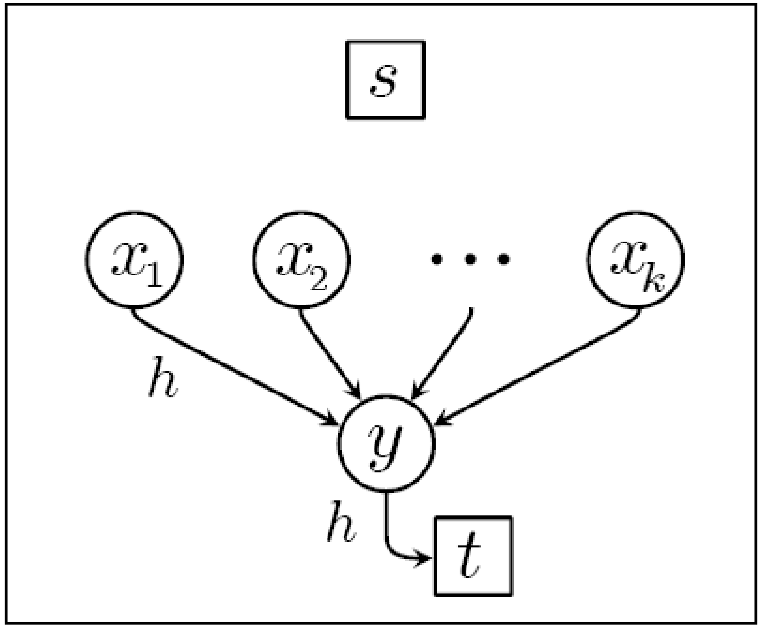
Modified energy $E_h^\alpha(\mathbf{x})$ pays h_β iff f^α contains label β

Deriving the Graph Construction

Add one auxiliary variable:

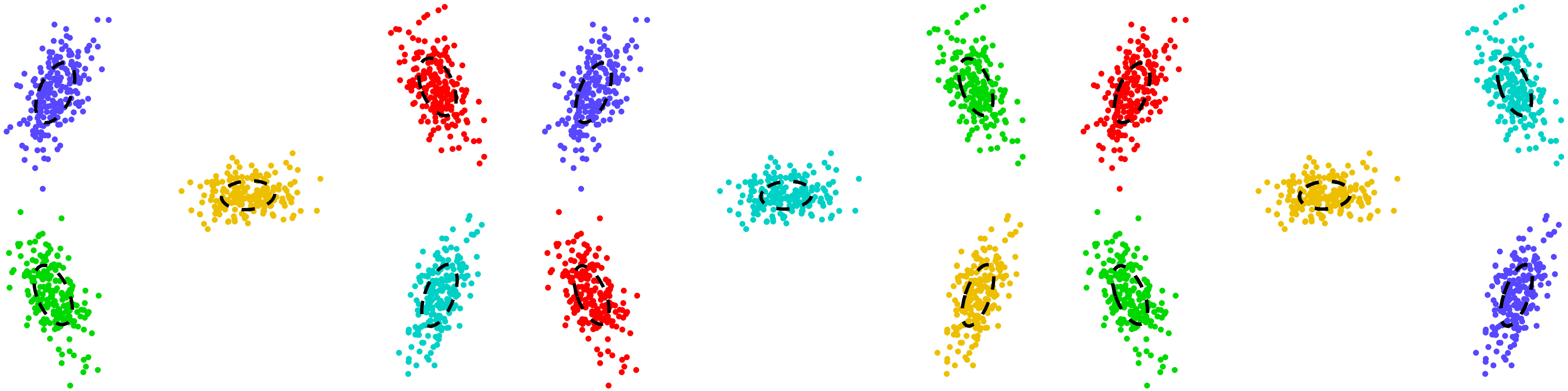
$$h(1 - x_1 x_5 x_6) = h \min_{y \in \{0,1\}} \left[\bar{y} + \bar{x}_1 y + \bar{x}_5 y + \bar{x}_6 y \right]$$

Same in terms of graph:



FMM

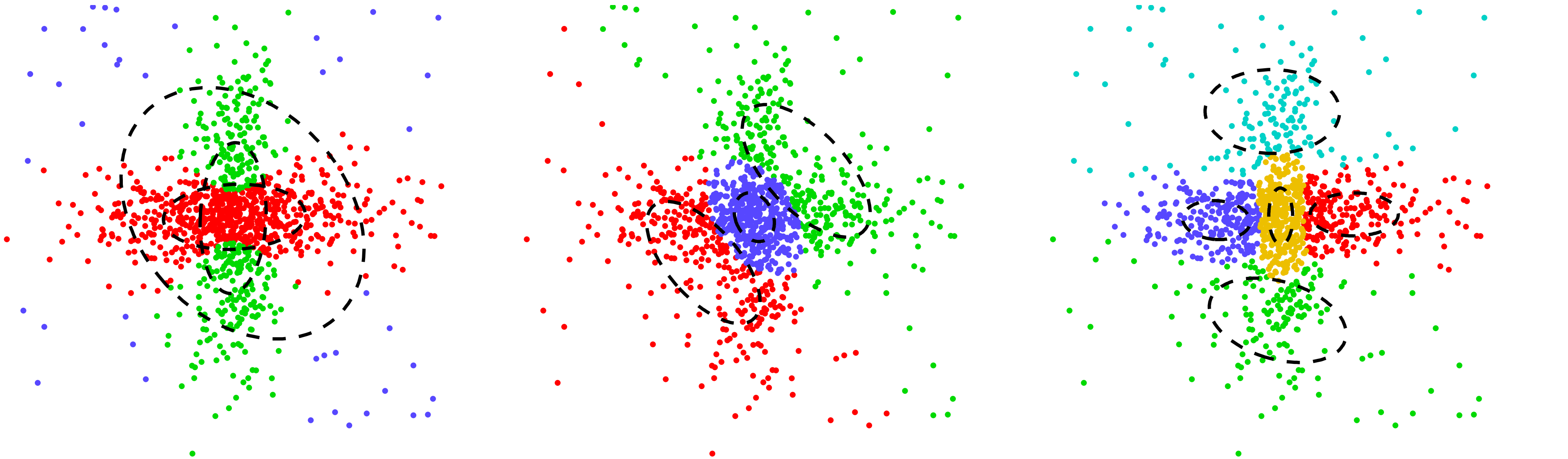
EM | elliptical K-means | Our approach



5 initial models | 5 initial models | 15 initial models

FMM

EM | elliptical K-means | Our approach



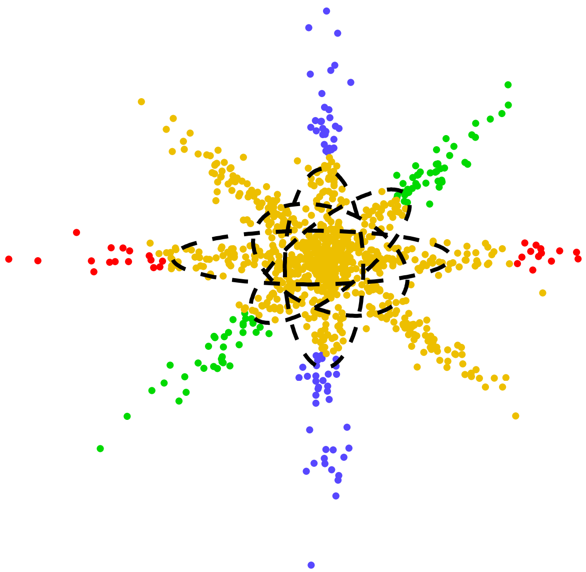
3 initial models | 3 initial models | 15 initial models

works well for
overlapping models
due to ***soft assignments***

hard assignments *fail if models overlap spatially*

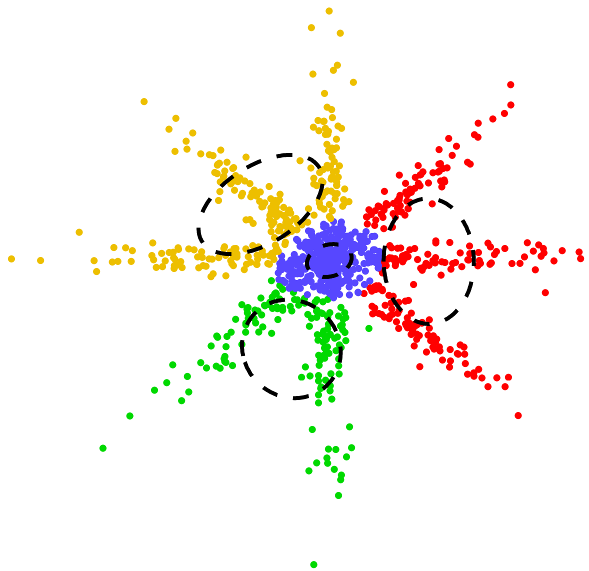
FMM

EM



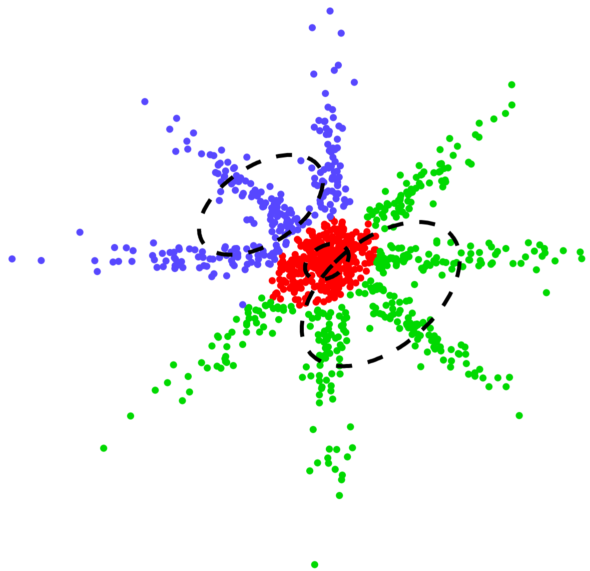
4 initial models

elliptical K-means



4 initial models

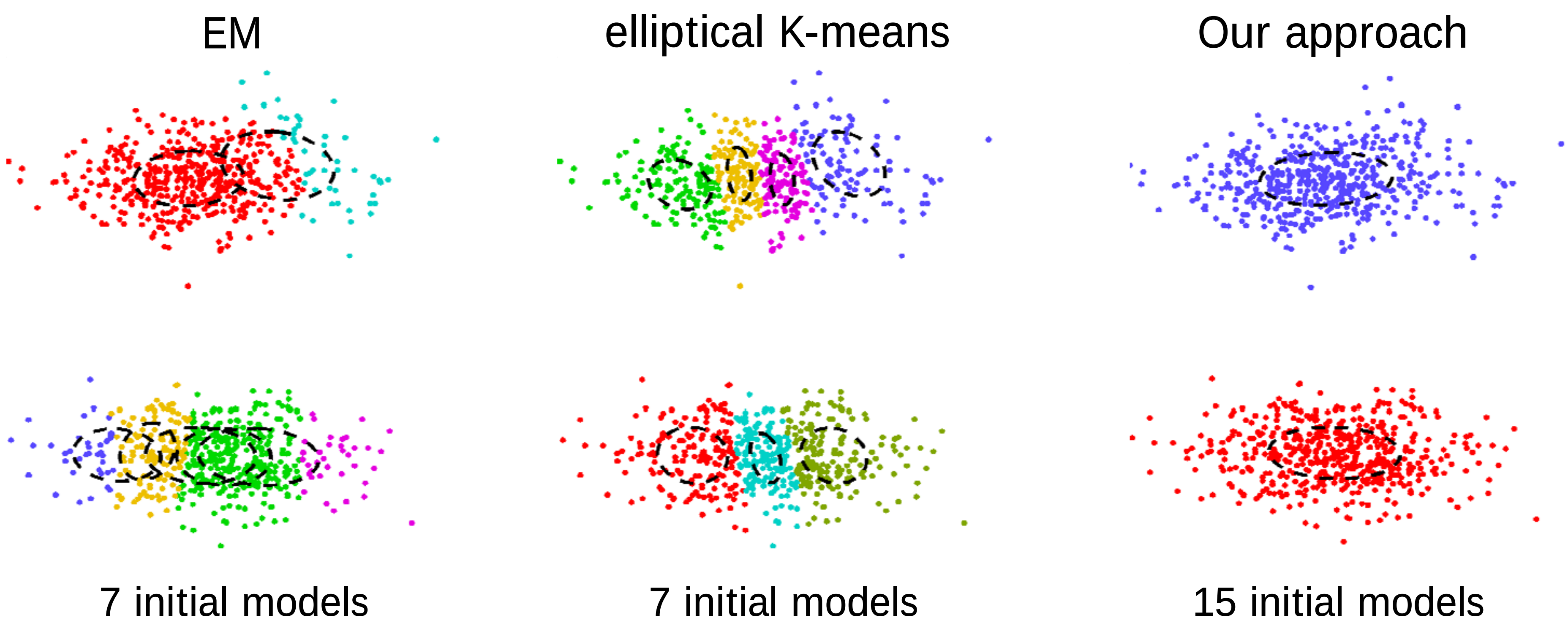
Our approach



15 initial models

soft assignments
is not a panacea

FMM



standard techniques must know the exact number of models *observation:*

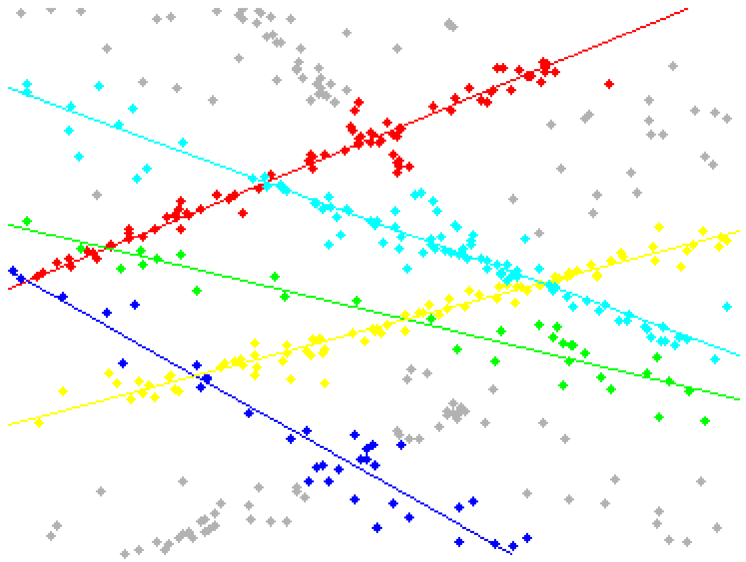
our labeling approach makes hard assignments which may cause problems if models have spatial overlap

Does not happen in vision

K-means vs. PEARL

$$E(f) = \sum_p \| p, f_p \| + \text{hard constraint on number of models}$$

K-means



5 random initial lines + outlier model

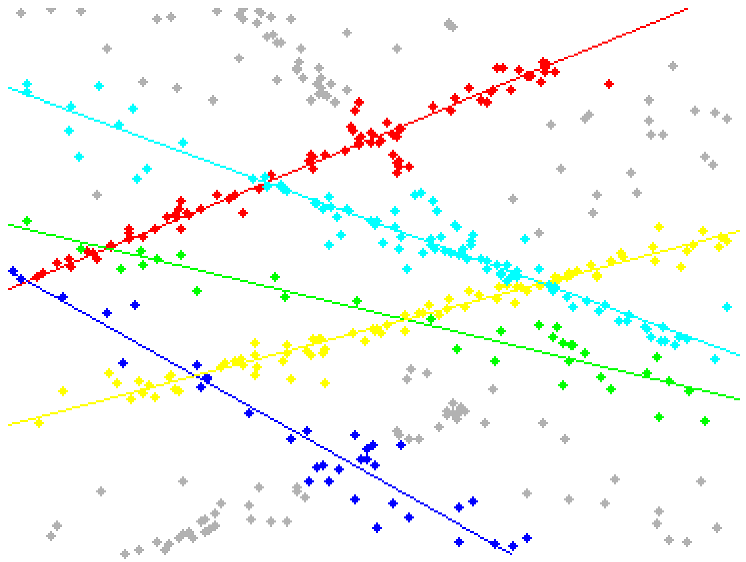
gets stuck in local minima

K-means vs. PEARL

$$E(f) = \sum_{p} \| p, f_p \| + \sum_{f \in \Lambda} h_f \cdot \delta_f(f)$$

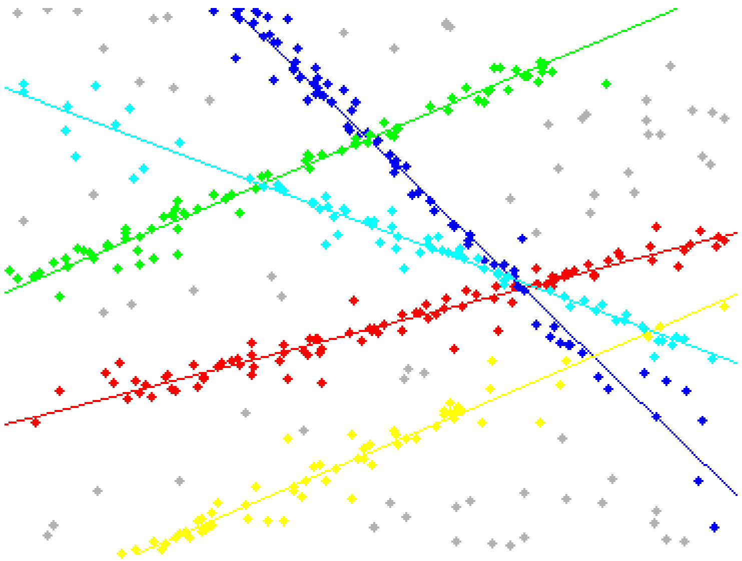
K-means

Our approach $h_f = 1000$



5 random initial lines + outlier model

gets stuck in local minima



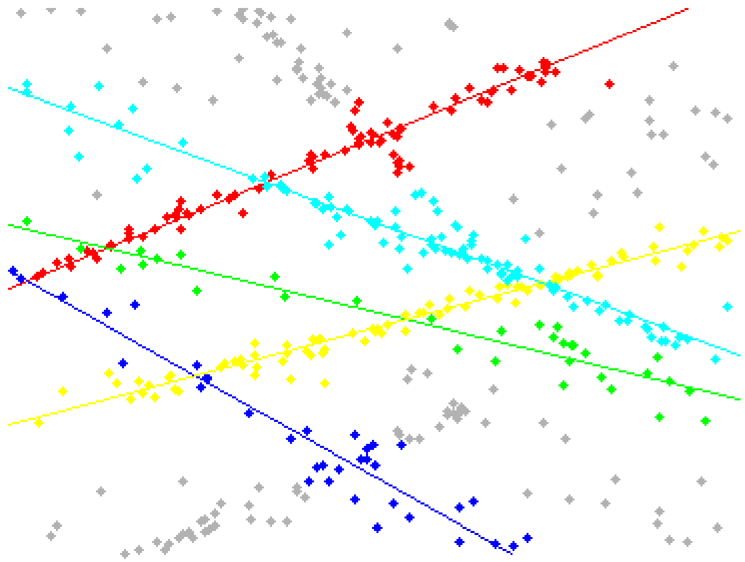
1000 initial lines + outlier model

better explores label space

K-means vs. PEARL

$$E(f) = \sum_{p} \| p, f_p \| + \sum_{f \in \Lambda} h_f \cdot \delta_f(f)$$

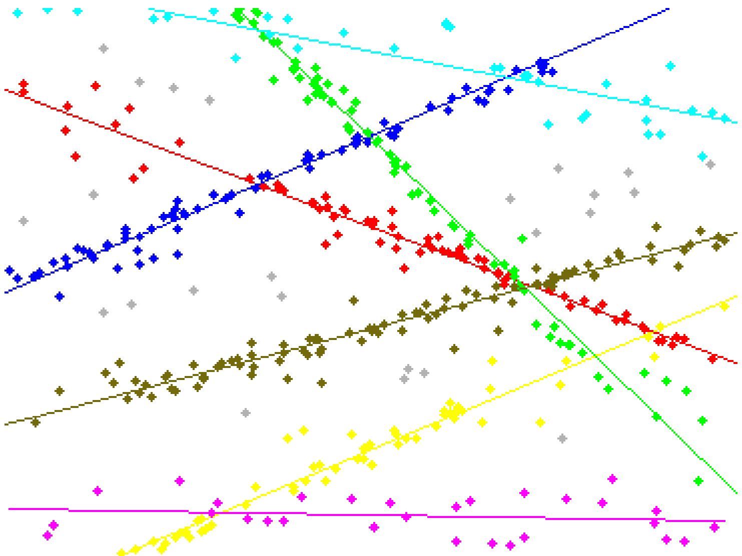
K-means



5 random initial lines + outlier model

gets stuck in local minima

Our approach $h_f = 500$



1000 initial lines + outlier model

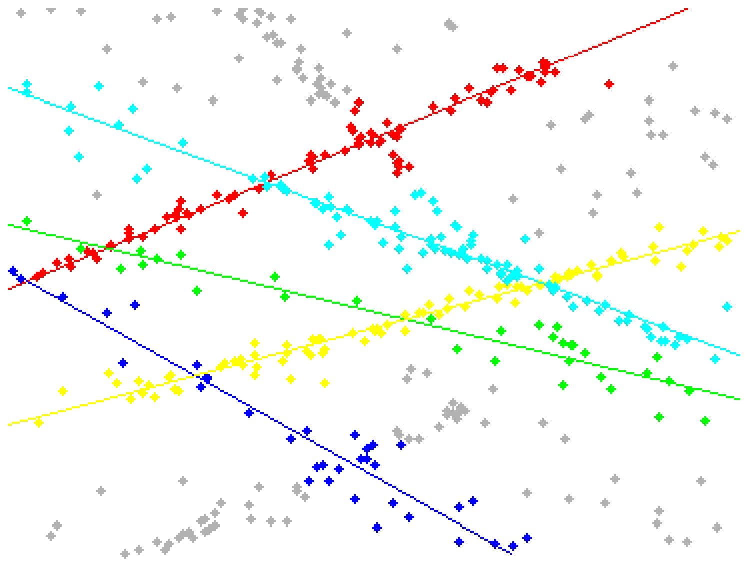
better explores label space

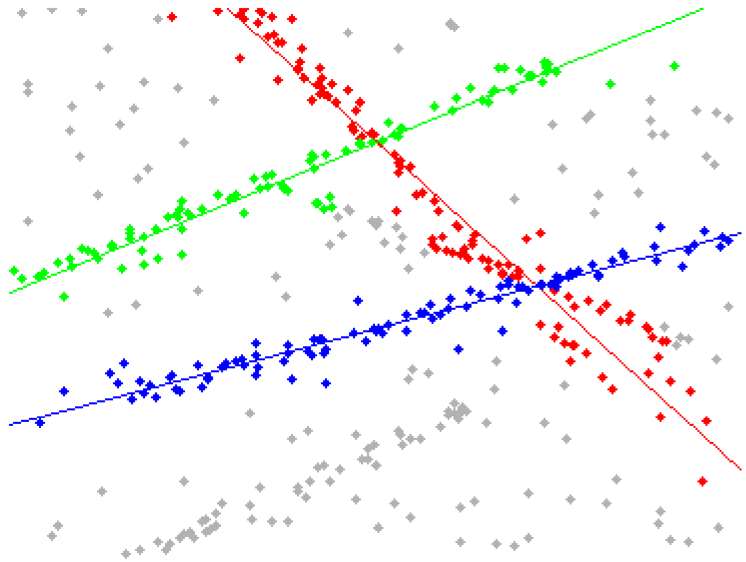
K-means vs. PEARL

$$E(f) = \sum_{p} \| p, f_p \| + \sum_{f \in \Lambda} h_f \cdot \delta_f(f)$$

K-means

Our approach $h_f = 2000$





5 random initial lines + outlier model

gets stuck in local minima

1000 initial lines + outlier model

better explores label space