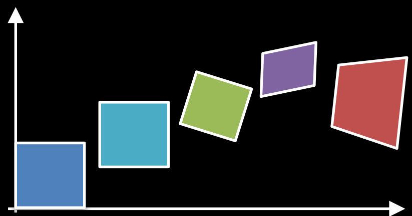
Motion Estimation (II)

Ce Liu
celiu@microsoft.com
Microsoft Research New England

Last time

- Motion perception



- Motion representation
- Parametric motion: *Lucas-Kanade*

$$\begin{bmatrix} du \\ dv \end{bmatrix} = -\begin{bmatrix} \mathbf{I}_x^T\mathbf{I}_x & \mathbf{I}_x^T\mathbf{I}_y \\ \mathbf{I}_x^T\mathbf{I}_y & \mathbf{I}_y^T\mathbf{I}_y \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{I}_x^T\mathbf{I}_t \\ \mathbf{I}_y^T\mathbf{I}_t \end{bmatrix}$$

- Dense optical flow: *Horn-Schunck*

$$\begin{bmatrix} \mathbf{I}_x^2 + \alpha\mathbf{L} & \mathbf{I}_x\mathbf{I}_y \\ \mathbf{I}_x\mathbf{I}_y & \mathbf{I}_y^2 + \alpha\mathbf{L} \end{bmatrix} \begin{bmatrix} U \\ V \end{bmatrix} = -\begin{bmatrix} \mathbf{I}_x I_t \\ \mathbf{I}_y I_t \end{bmatrix}$$

- Robust estimation

$$\begin{bmatrix} \mathbf{\Psi}'_{xx} + \alpha\mathbf{L} & \mathbf{\Psi}'_{xy} \\ \mathbf{\Psi}'_{xy} & \mathbf{\Psi}'_{yy} + \alpha\mathbf{L} \end{bmatrix} \begin{bmatrix} dU \\ dV \end{bmatrix} = -\begin{bmatrix} \mathbf{\Psi}'_{xt} + \alpha\mathbf{L}U \\ \mathbf{I}_y I_t + \alpha\mathbf{L}V \end{bmatrix}$$

- Applications (1)

Who are they?

Berthold K. P. Horn

Takeo Kanade

Today

- Discrete optical flow
- Layer motion analysis
- Contour motion analysis
- Obtaining motion ground truth

Block matching

- Both Horn-Schunk and Lucas-Kanade are sub-pixel accuracy algorithms
- But in practice we may not need sub-pixel accuracy
- MPEG: 16×16 block matching using MMSE (insert a block matching example)

Tracking reliable features

- Idea: no need to work on ambiguous regions pixels (flat regions & line structures)
- Instead, we can track features and then propagate the tracking to ambiguous pixels
- Good features to track [Shi & Tomasi 94]

$$\begin{bmatrix} \mathrm{d}u \\ \mathrm{d}v \end{bmatrix} = -\begin{bmatrix} \mathbf{I}_x^T\mathbf{I}_x & \mathbf{I}_x^T\mathbf{I}_y \\ \mathbf{I}_x^T\mathbf{I}_y & \mathbf{I}_y^T\mathbf{I}_y \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{I}_x^T\mathbf{I}_t \\ \mathbf{I}_y^T\mathbf{I}_t \end{bmatrix}$$

- Block matching + Lucas-Kanade refinement

Feature detection & tracking

From sparse to dense

- Interpolation: given values $\{d_i\}$ at $\{(x_i, y_i)\}$, reconstruct a smooth plane $f(x, y)$
- Membrane model

$$\iint \sum_i w_i (f(x_i, y_i) - d_i)^2 + \alpha \left(f_x^2 + f_y^2\right) dxdy$$

- Thin plate model

$$\iint \sum_i w_i (f(x_i, y_i) - d_i)^2 + \alpha \left(f_{xx}^2 + f_{xy}^2 + f_{yy}^2\right) dxdy$$

Membrane vs. thin plate

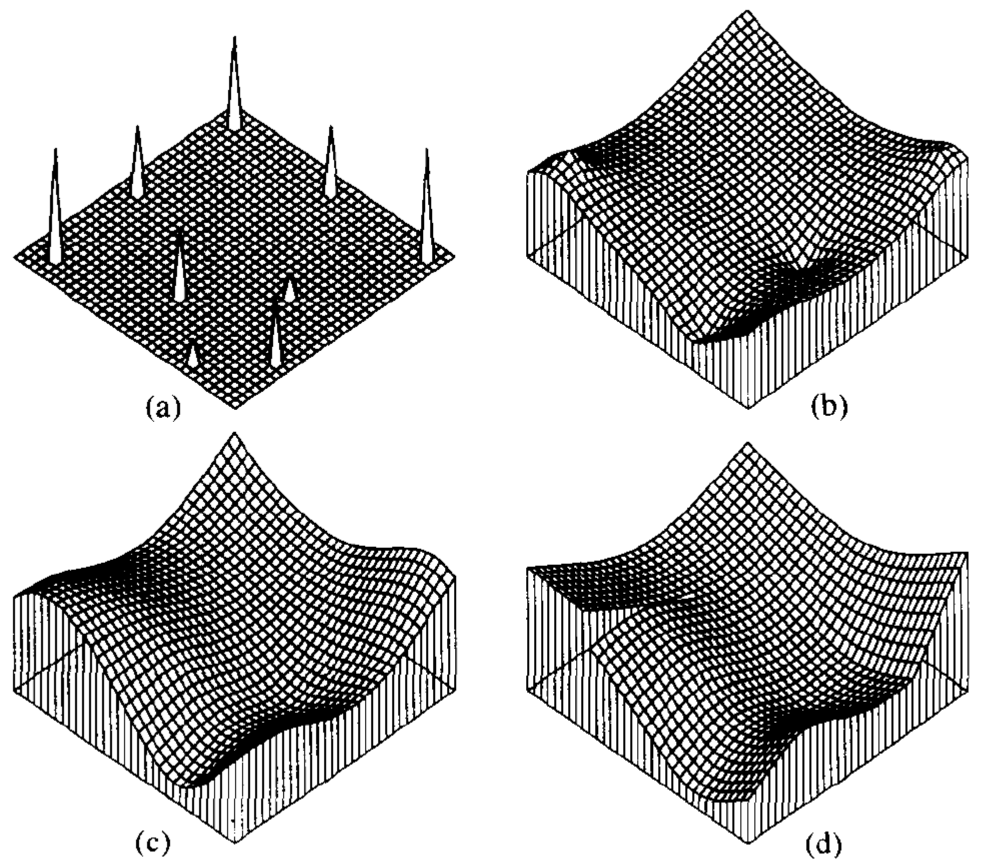


Fig. 1. Sample data points and interpolated solutions: (a) sample data points, (b) membrane interpolant, (c) thin plate interpolant, (d) controlled continuity spline (thin plate with discontinuities and creases).

Dense flow field from sparse tracking

Pros and Cons of Feature Matching

- Pros
  - Efficient (a few feature points vs. all pixels)
  - Reliable (with advanced feature descriptors)
- Cons
  - Independent tracking (tracking can be unreliable)
  - Not all information is used (may not capture weak features)
- How to improve
  - Track every pixel with uncertainty
  - Integrate spatial regularity (neighboring pixels go together)

Discrete Optical Flow

- The objective function is similar to continuous flow

$$E(\mathrm{w}) = \sum_{\mathrm{x}} \min\left(\left|I_1(\mathrm{x}) - I_2\big(\mathrm{w}(\mathrm{x})\big)\right|, t\right) +$$

$$\sum_{\mathrm{x}} \eta(|u(\mathrm{x})| + |v(\mathrm{x})|)$$

$$\sum_{(\mathrm{x}_1,\mathrm{x}_2)\in\varepsilon} \min(\alpha|u(\mathrm{x}_1) - u(\mathrm{x}_2)|, d) + \min(\alpha|v(\mathrm{x}_1) - v(\mathrm{x}_2)|, d)$$

Data term

Small displacement

Spatial regularity

- $\mathrm{x} = (x, y)$ is pixel coordinate, $\mathrm{w} = (u, v)$ is flow vector
- Truncated L1 norms:
  - Account for outliers in the data term
  - Encourage piecewise smoothness in the smoothness term

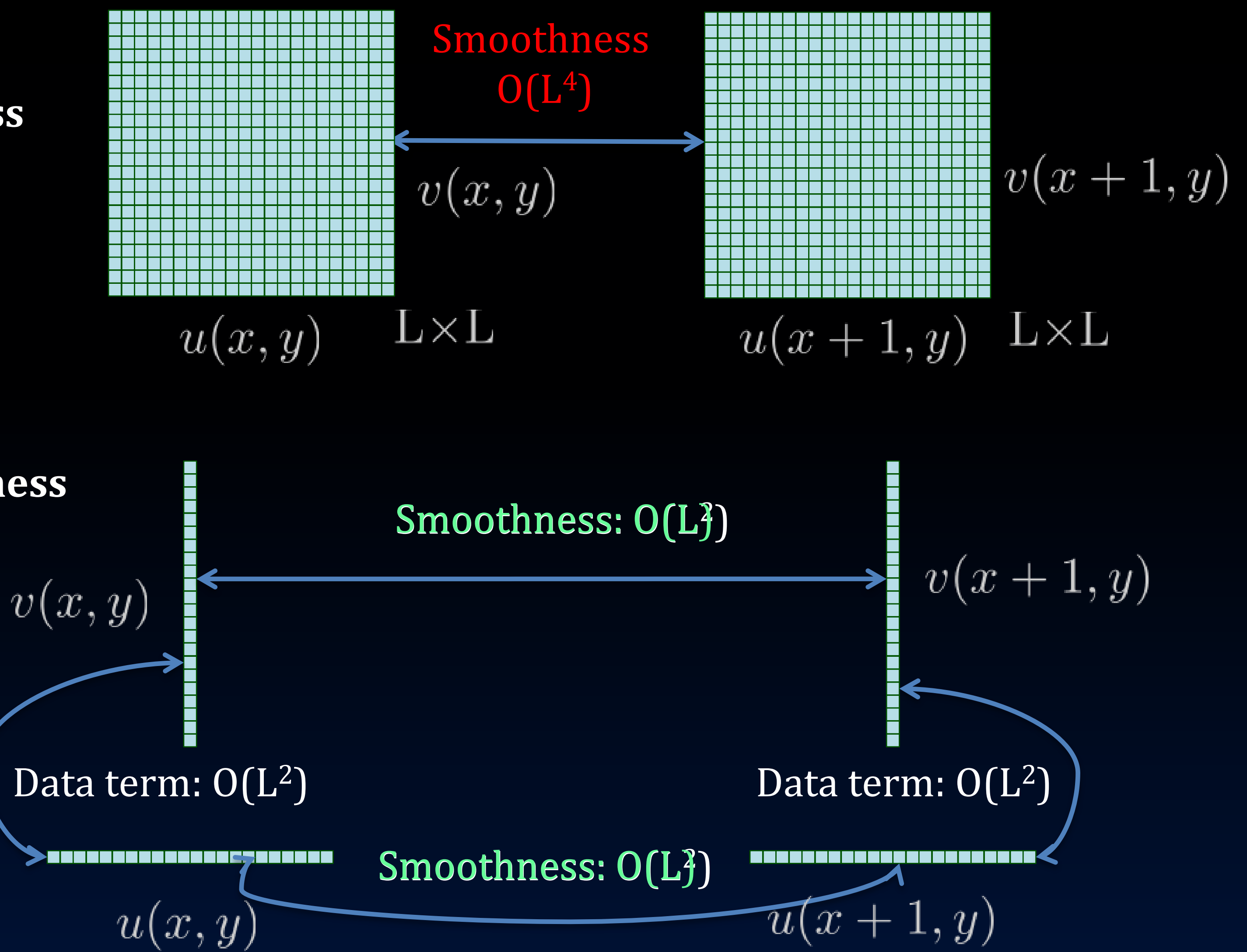
Decoupled smoothness

Coupled smoothness

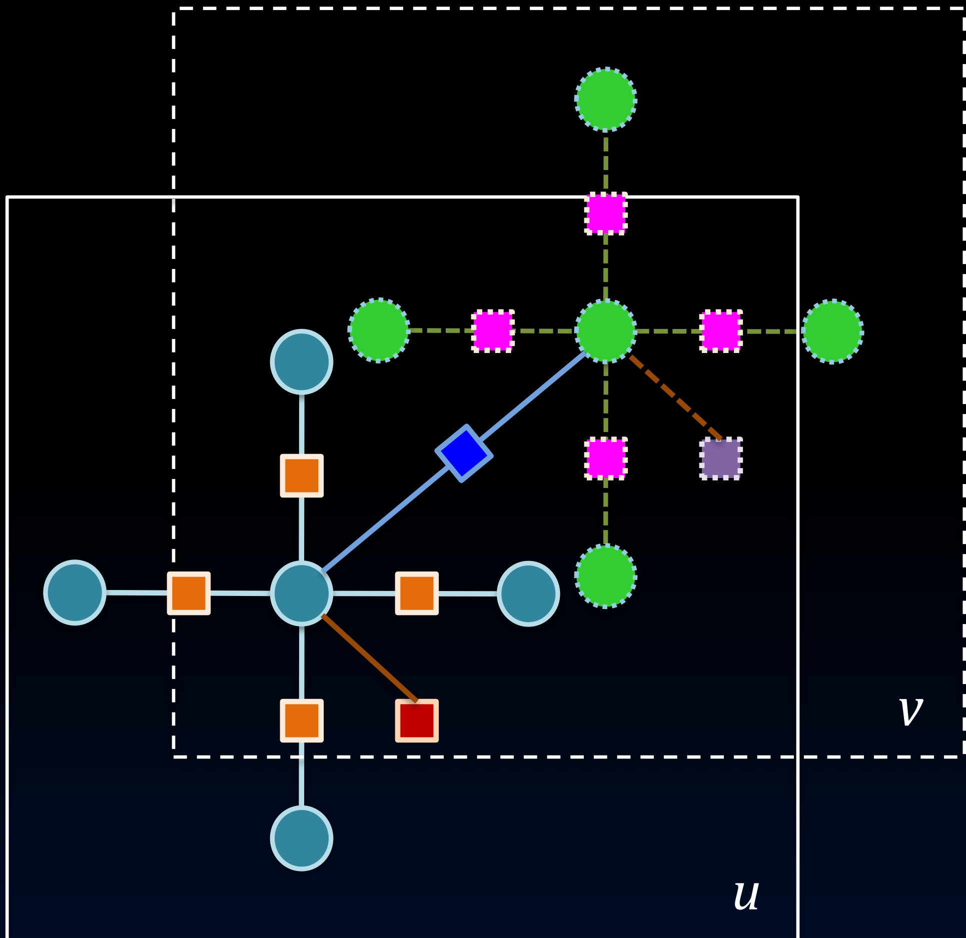
$$\sqrt{u_x^2 + v_x^2}$$

Decoupled smoothness

$$|u_x| + |v_x|$$



Dual-layer belief propagation



[Shekhovtsov et al. CVPR 07]

- Horizontal flow u
- Vertical flow v

$\mathrm{w} = (u, v)$

- Data term

$$||I_1(\mathrm{x}) - I_2(\mathrm{x} + \mathrm{w})||_1$$

- Smoothness term on u

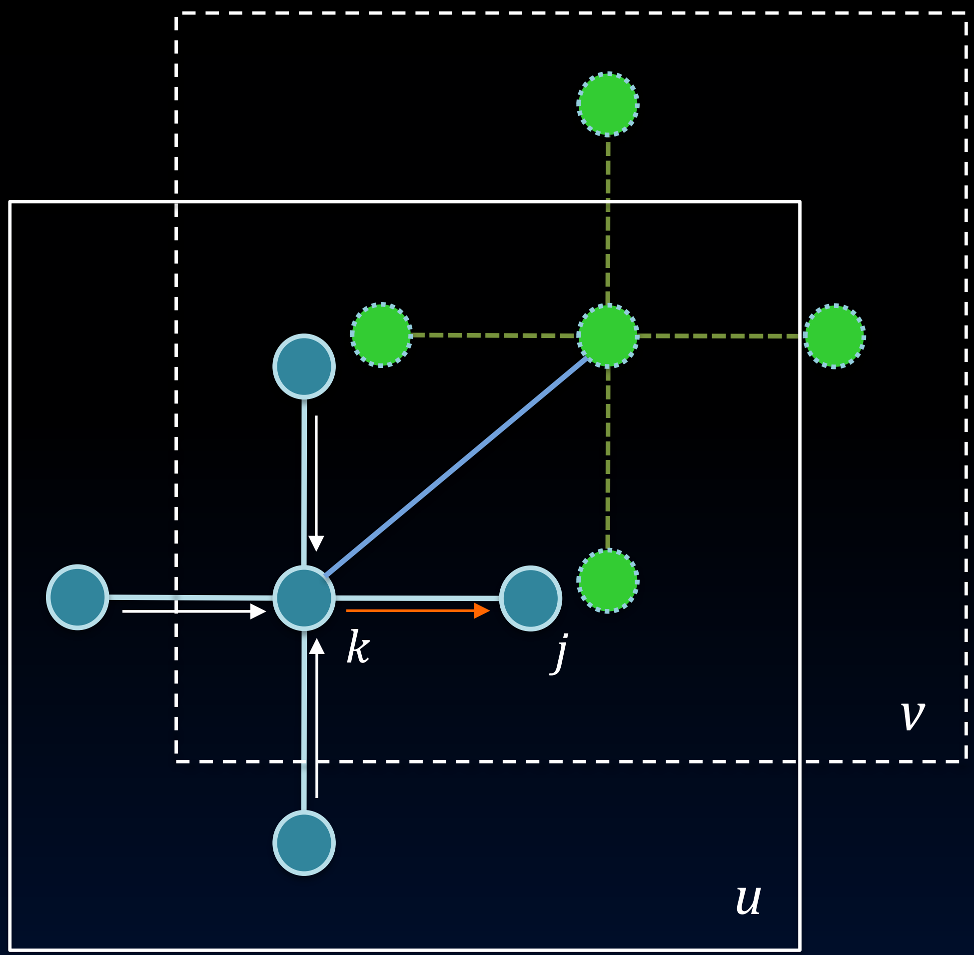
$$\min(\alpha|u(\mathrm{x}_1) - u(\mathrm{x}_2)|, d)$$

- Smoothness term on v

$$\min(\alpha|v(\mathrm{x}_1) - v(\mathrm{x}_2)|, d)$$

- Regularization term on u $\quad \eta|u(\mathrm{x})|$
- Regularization term on v $\quad \eta|v(\mathrm{x})|$

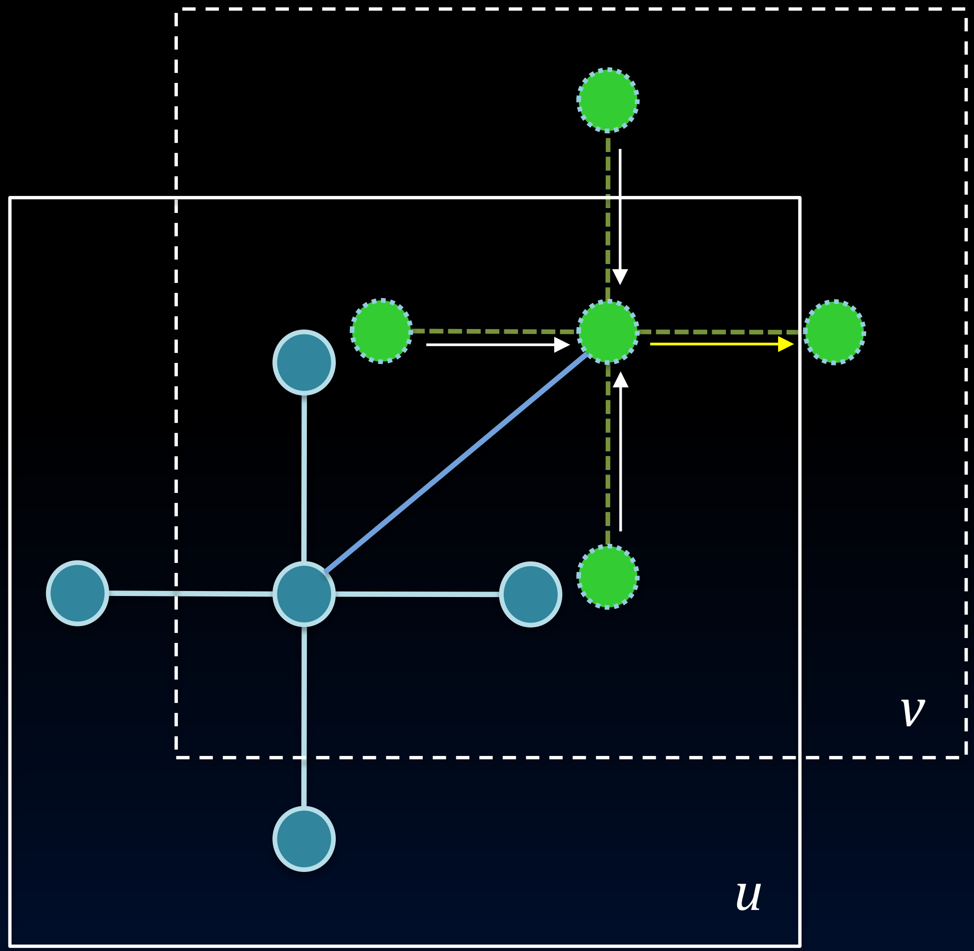
Dual-layer belief propagation



Message M_j^k: given all the information at node k, predict the distribution at node j

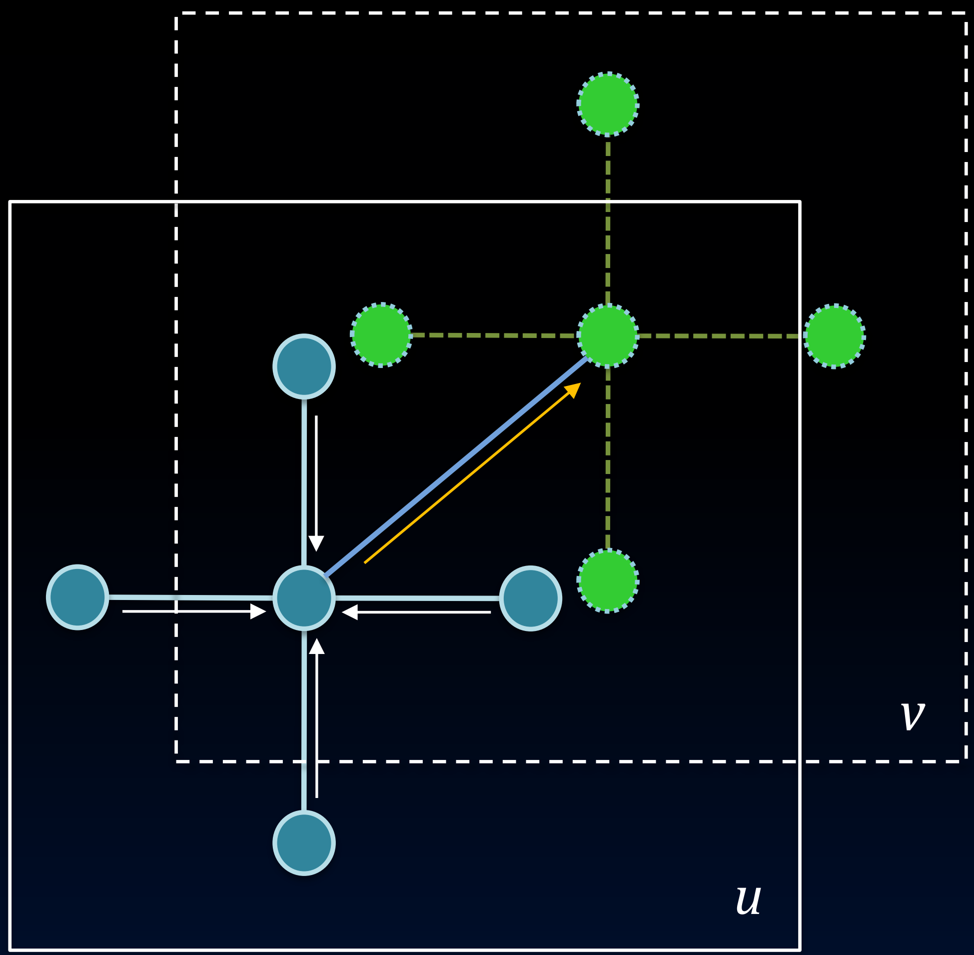
Update within u plane

Dual-layer belief propagation



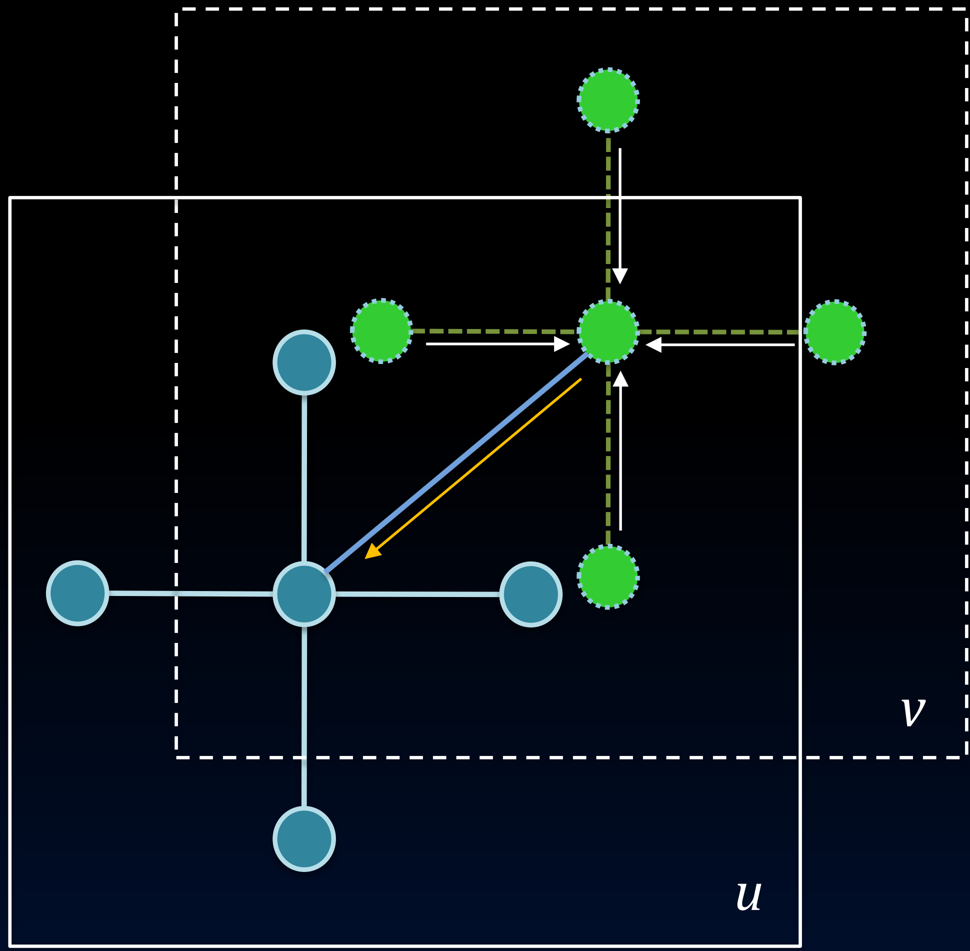
Update within v plane

Dual-layer belief propagation



Update from u plane to v plane

Dual-layer belief propagation

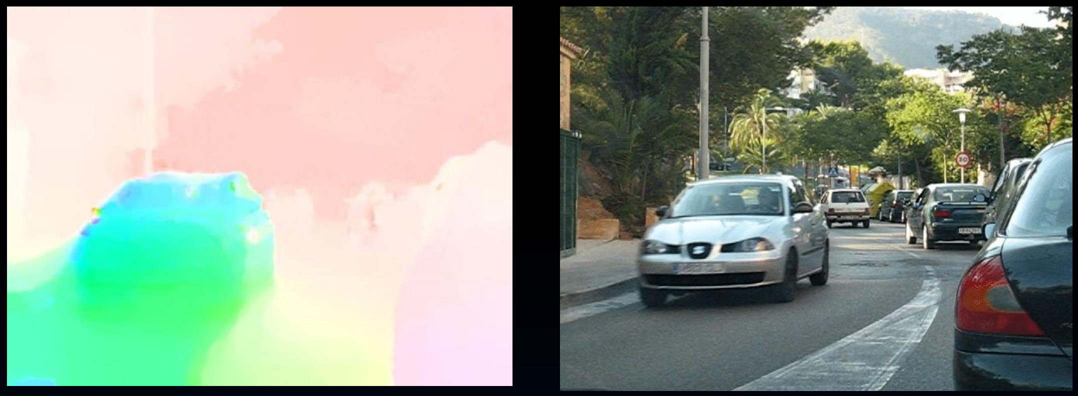


Update from v plane to u plane

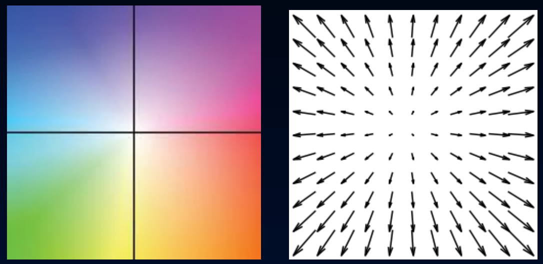
Examples

Discrete optical flow

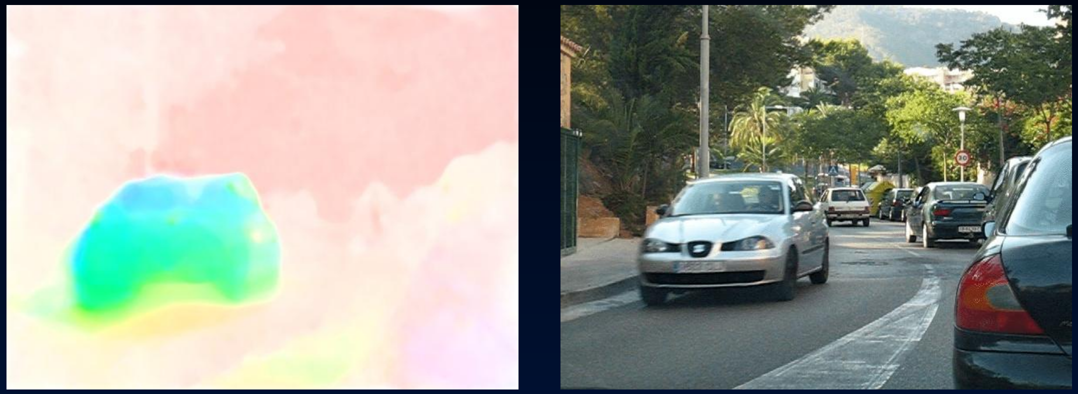
Input two frames



Robust optical flow



Flow visualization



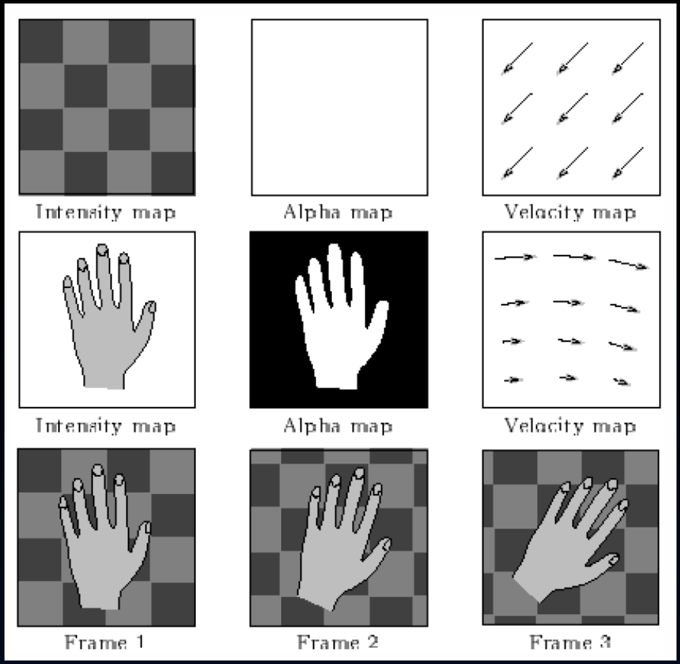
Coarse-to-fine LK with median filtering

Content

- Discrete optical flow
- Layer motion analysis
- Contour motion analysis
- Obtaining motion ground truth

Layer representation

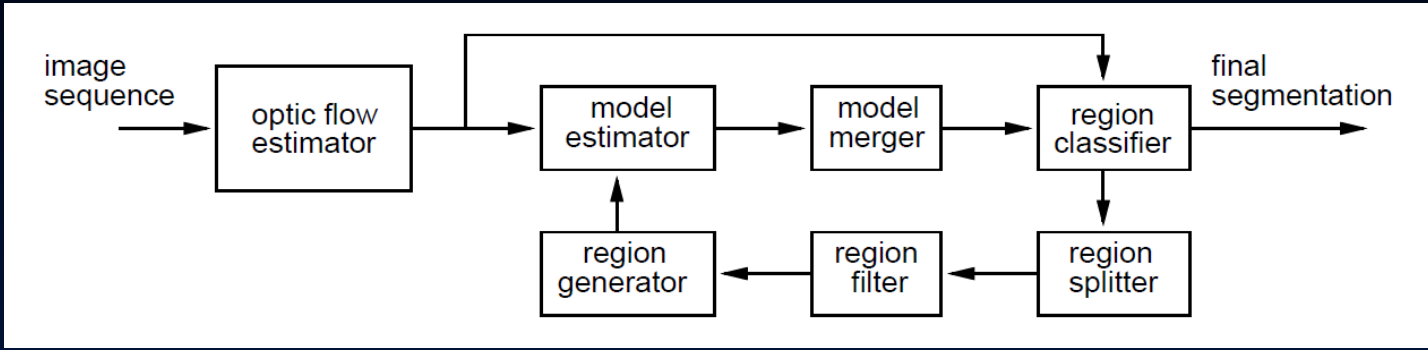
- Optical flow field is able to model complicated motion
- Different angle: a video sequence can be a composite of several moving layers
- Layers have been widely used
  - Adobe Photoshop
  - Adobe After Effect
- Compositing is straightforward, but inference is hard



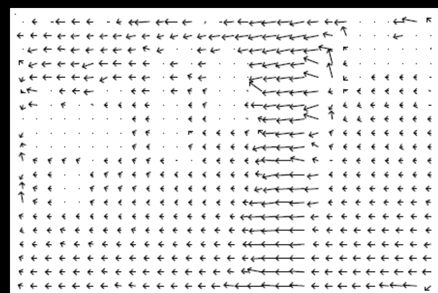
Wang & Adelson, 1994

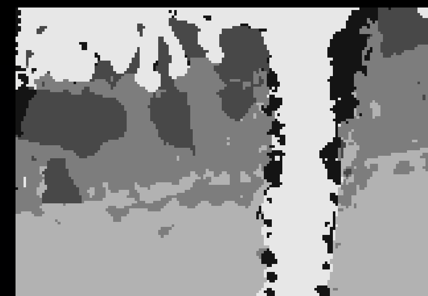
Wang & Adelson, 1994

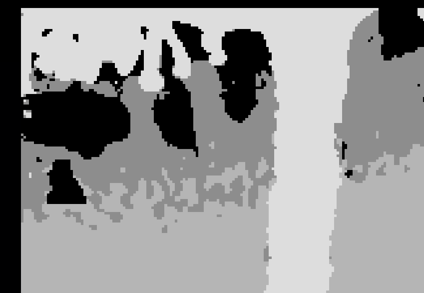
- Strategy
  - Obtaining dense optical flow field
  - Divide a frame into non-overlapping regions and fit affine motion for each region
  - Cluster affine motions by k-means clustering
  - Region assignment by hypothesis testing
  - Region splitter: disconnected regions are separated



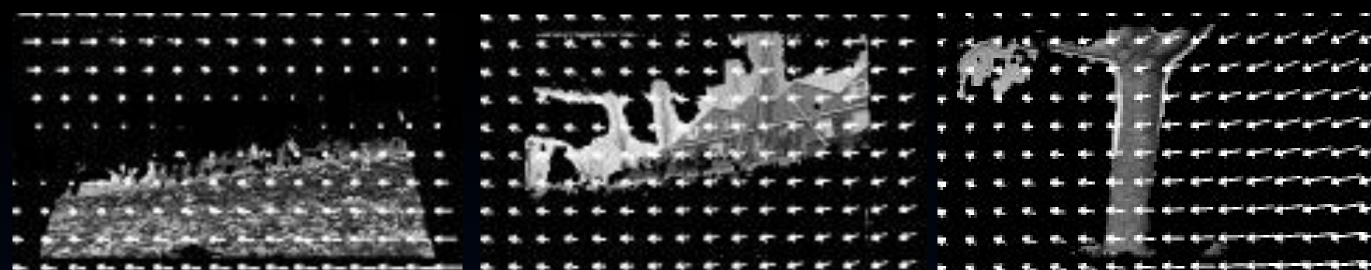
Results


Optical flow field


Clustering to affine regions


Clustering with error metric


Flower garden


Three layers with affine motion superimposed


Reconstructed background layer

Weiss & Adelson, 1996

- Chicken & egg problem
  - Good motion → good segmentation
  - Good segmentation → good motion
- We don't have either of them, so iterate!
- Perceptually organized expectation & maximization (POEM)
  - E-step: estimate the motion parameter of each layer
  - M-step: estimate the likelihood that a pixel belongs to each of the layers (segmentation)

Liu & Torralba et. al. 2005

- Reliable layer segmentation for motion magnification
- Layer segmentation pipeline

→ Feature point tracking → Trajectory clustering → Dense optical flow interpolation → Layer segmentation

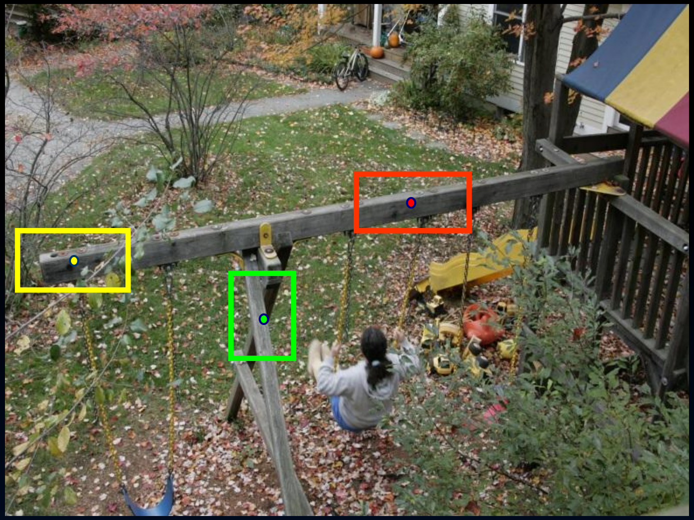
Input video sequence

Feature point tracking

Normalized complex correlation

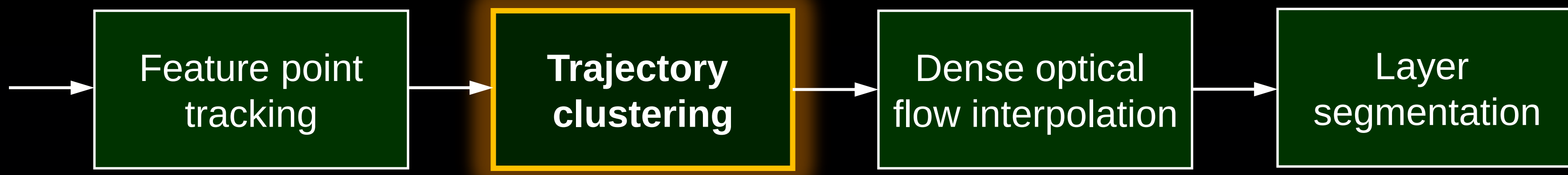
Feature point tracking → **Trajectory clustering** → Dense optical flow interpolation → Layer segmentation

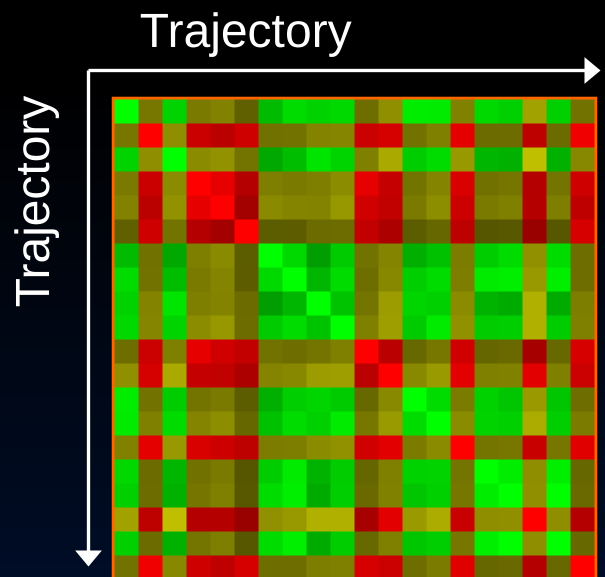
- The similarity metric should be independent of phase and magnitude
- Normalized complex correlation



$$S(C_1, C_2) = \frac{|\sum_t C_1(t)\overline{C}_2(t)|^2}{\sqrt{\sum_t C_1(t)\overline{C}_1(t)}\sqrt{\sum_t C_2(t)\overline{C}_2(t)}}$$

Spectral clustering

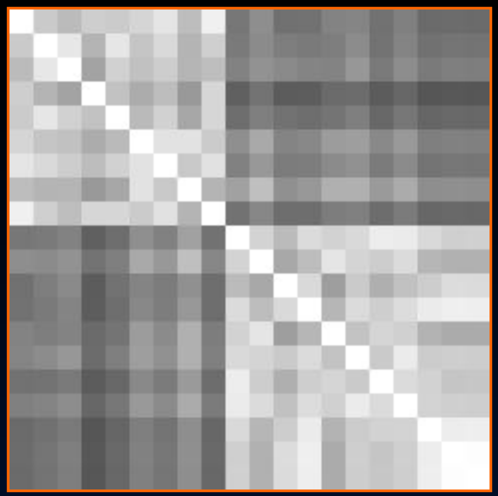




Affinity matrix

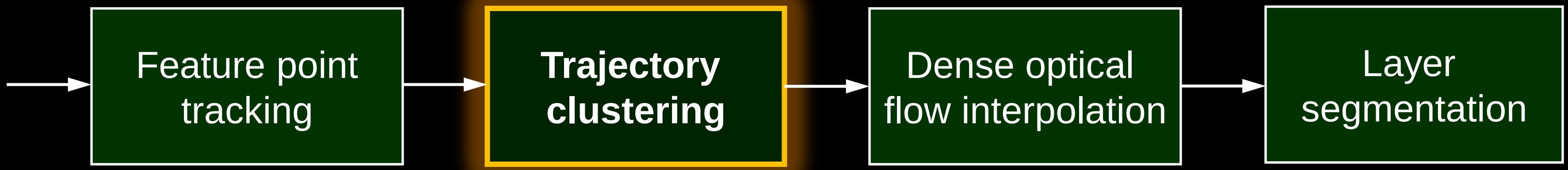
Two clusters

Clustering



Reordering of affinity matrix

Clustering results



From sparse feature points to dense optical flow fields

Feature point tracking → Trajectory clustering → **Dense optical flow interpolation** → Layer segmentation

- Interpolate dense optical flow field using locally weighted linear regression

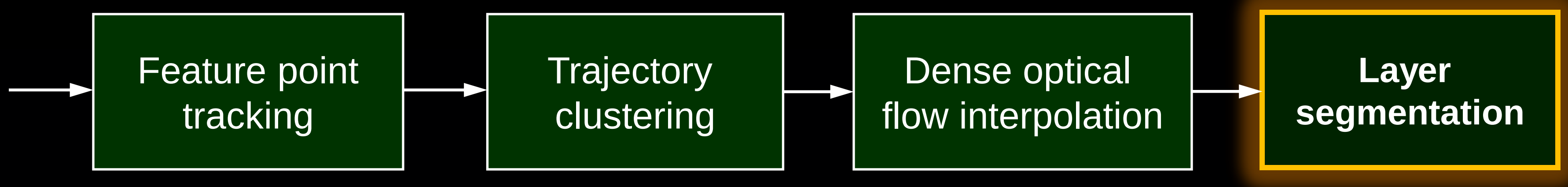
Flow vectors of clustered sparse feature points

Dense optical flow field of cluster 2 (swing)

Cluster 1: leaves

Cluster 2: swing

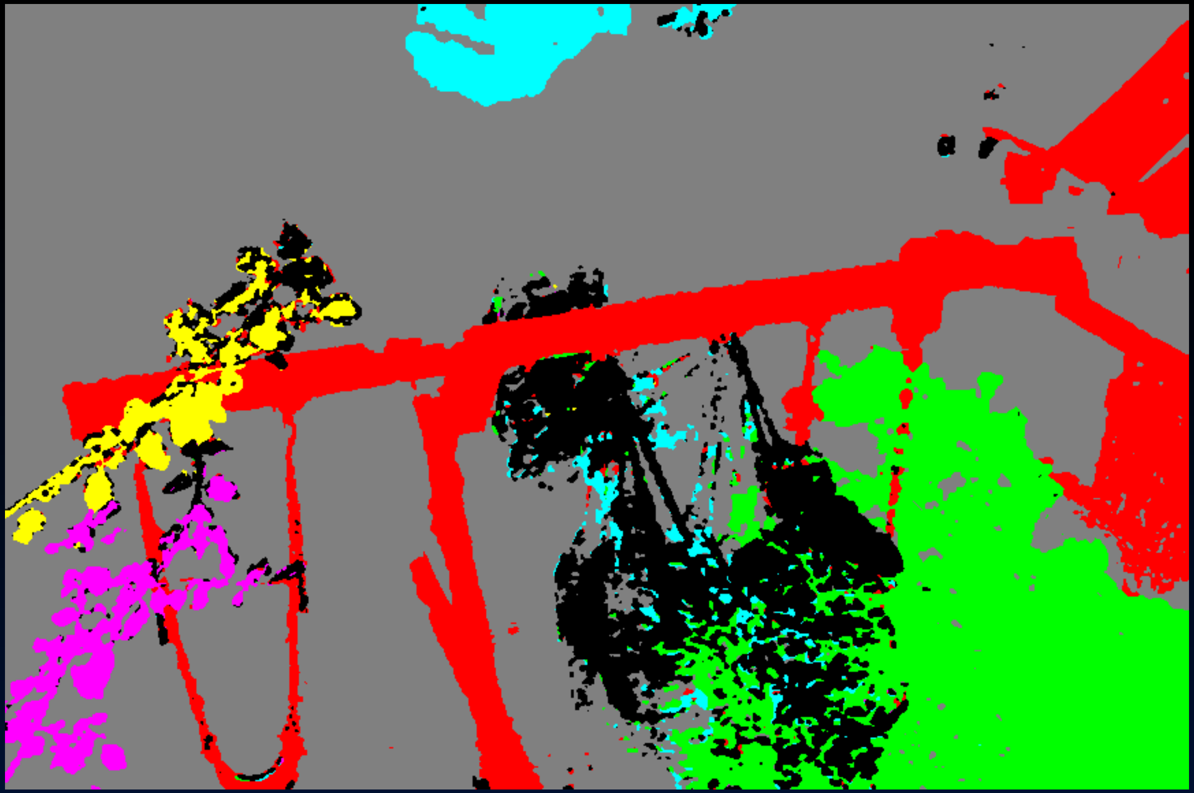
Motion layer assignment



- Assign each pixel to a motion cluster layer, using four cues:
  - **Motion likelihood**—consistency of pixel's intensity if it moves with the motion of a given layer (dense optical flow field)
  - **Color likelihood**—consistency of the color in a layer
  - **Spatial connectivity**—adjacent pixels favored to belong the same group
  - **Temporal coherence**—label assignment stays constant over time
- Energy minimization using graph cuts

Segmentation results

Feature point tracking → Trajectory clustering → Dense optical flow interpolation → **Layer segmentation**



Motion Magnification

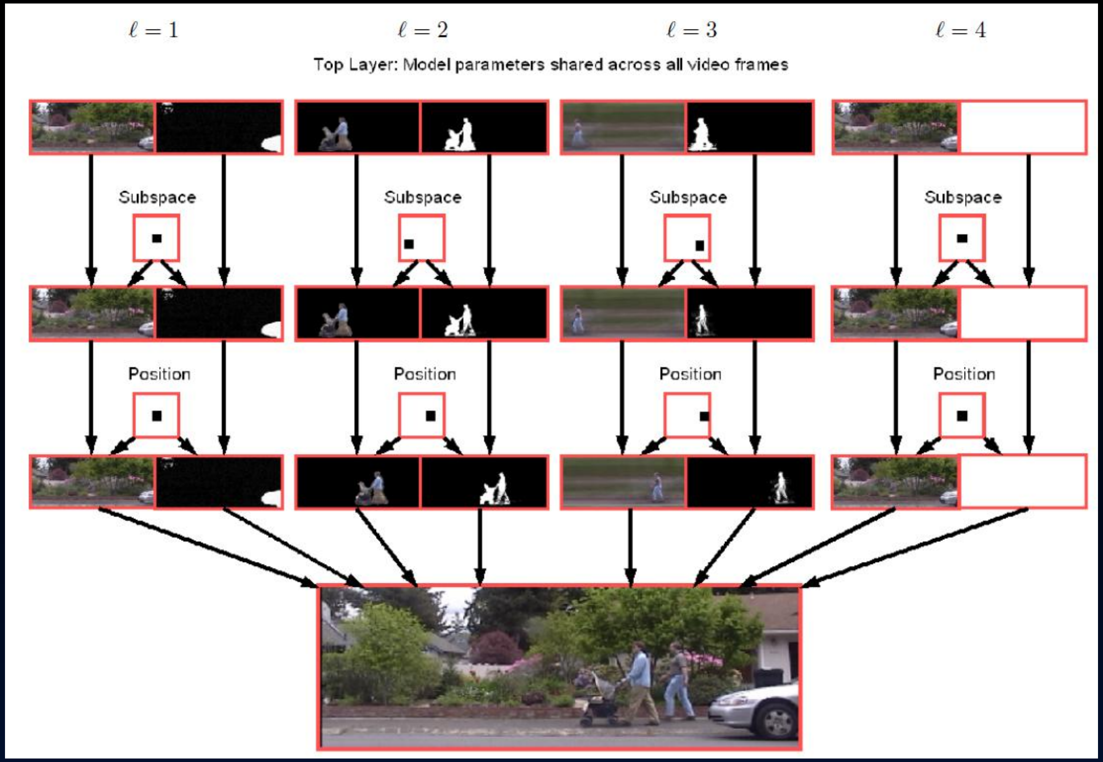
Ce Liu
Antonio Torralba
William T. Freeman
Fredo Durand
Edward H. Adelson

Massachusetts Institute of Technology
Computer Science and Artificial Intelligence Laboratory

SIGGRAPH2005
The 32nd International Conference on Computer Graphics and Interactive Techniques

Generative models

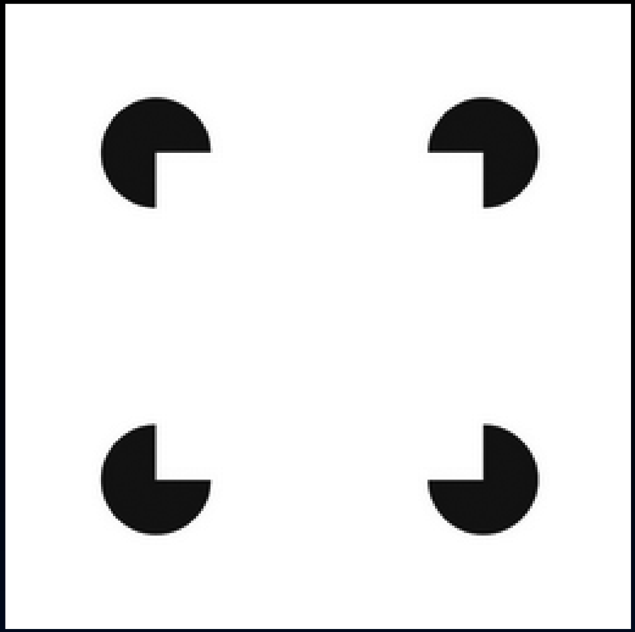
- Learning flexible sprites [Frey & Jojic 2001, 2003]



Content

- Discrete optical flow
- Layer motion analysis
- Contour motion analysis
- Obtaining motion ground truth

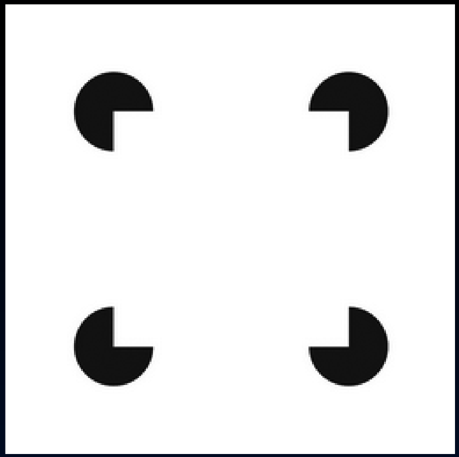
Seemingly Simple Examples



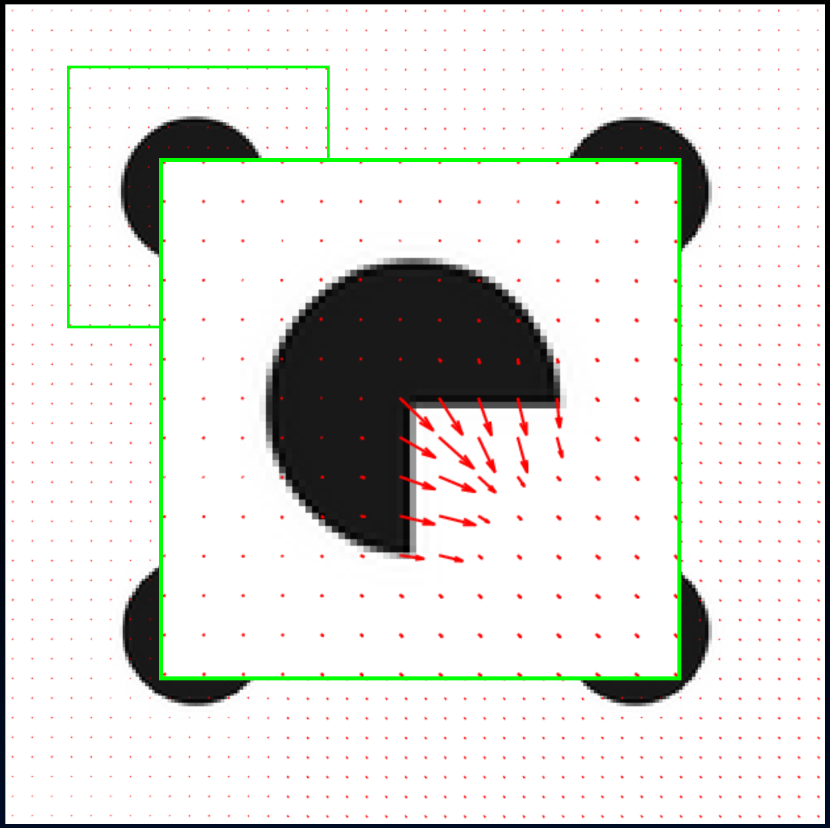
Kanizsa square

From real video

Output from the State-of-the-Art Optical Flow Algorithm



Kanizsa square

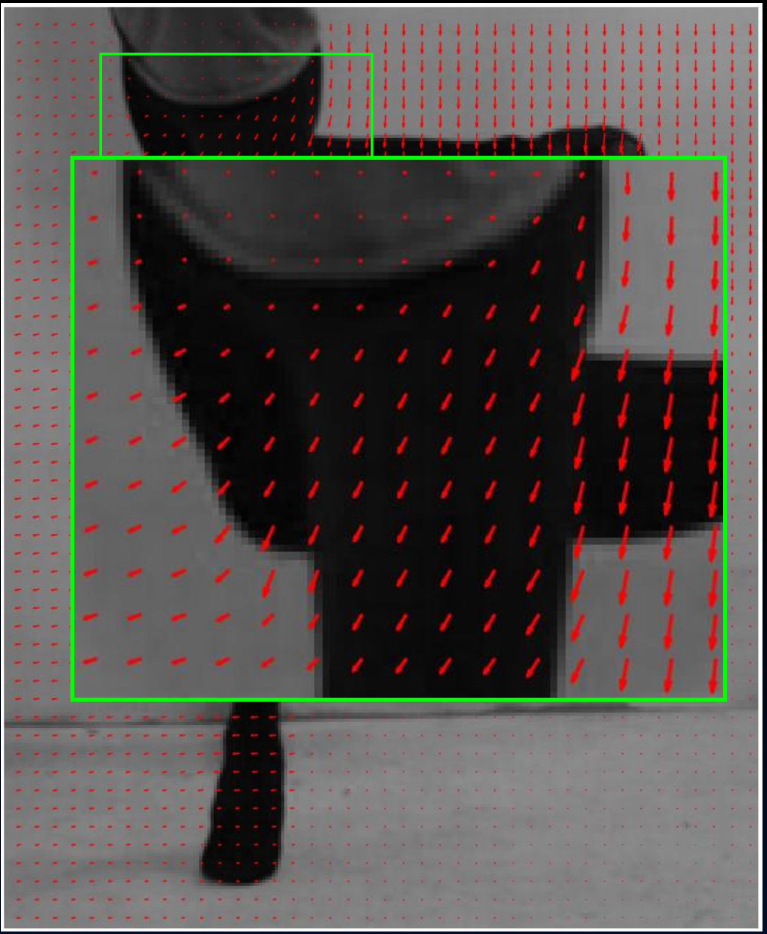


Optical flow field

T. Brox *et al*. High accuracy optical flow estimation based on a theory for warping. *ECCV* 2004

Output from the State-of-the-Art Optical Flow Algorithm

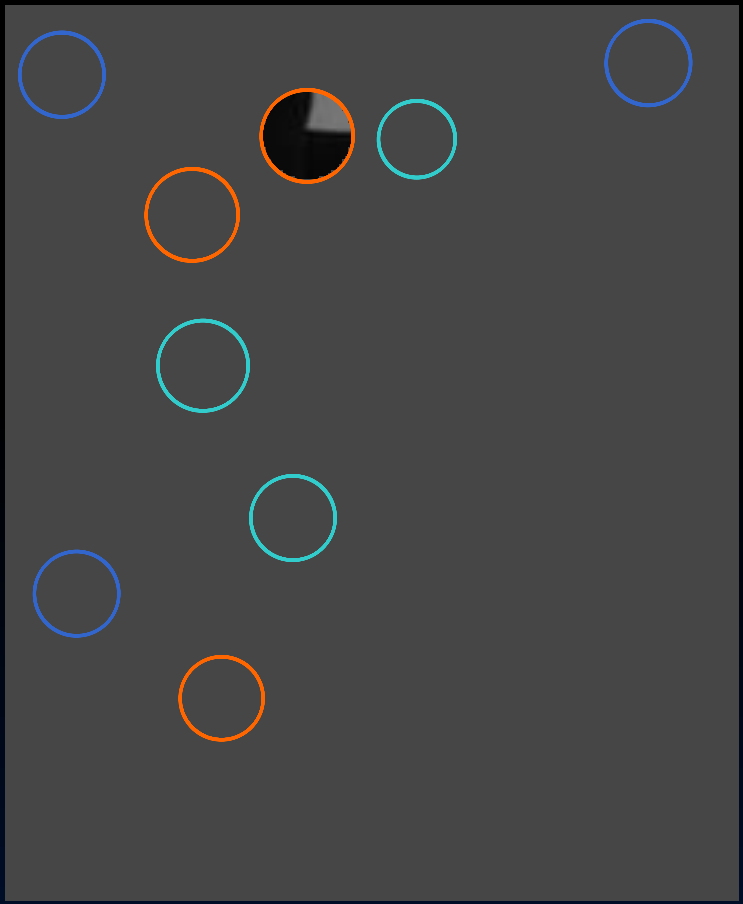
Dancer

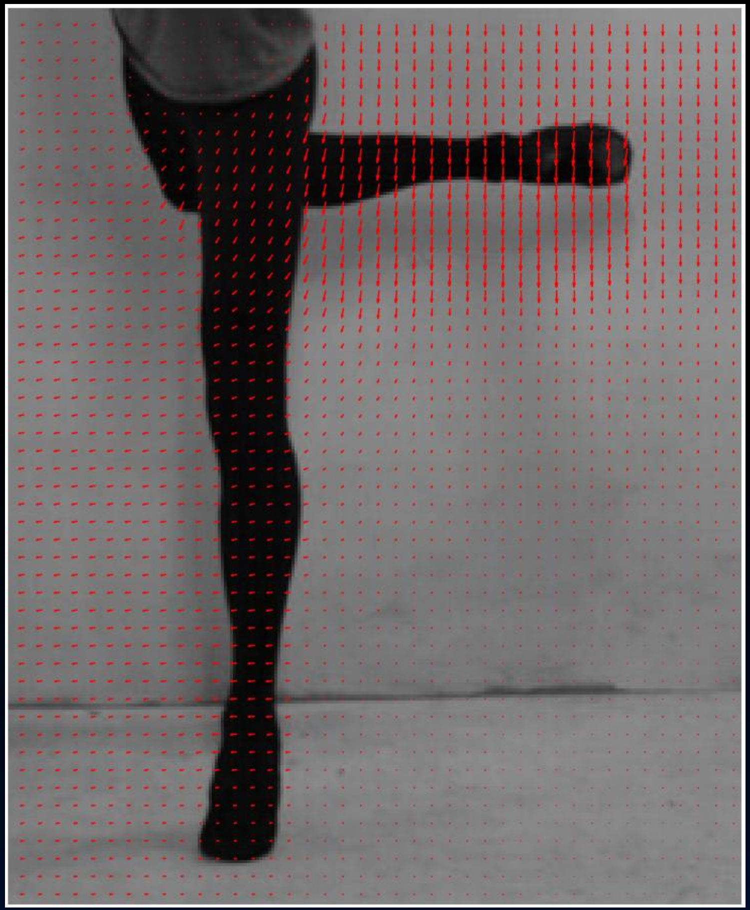


Optical flow field

T. Brox *et al*. High accuracy optical flow estimation based on a theory for warping. *ECCV* 2004

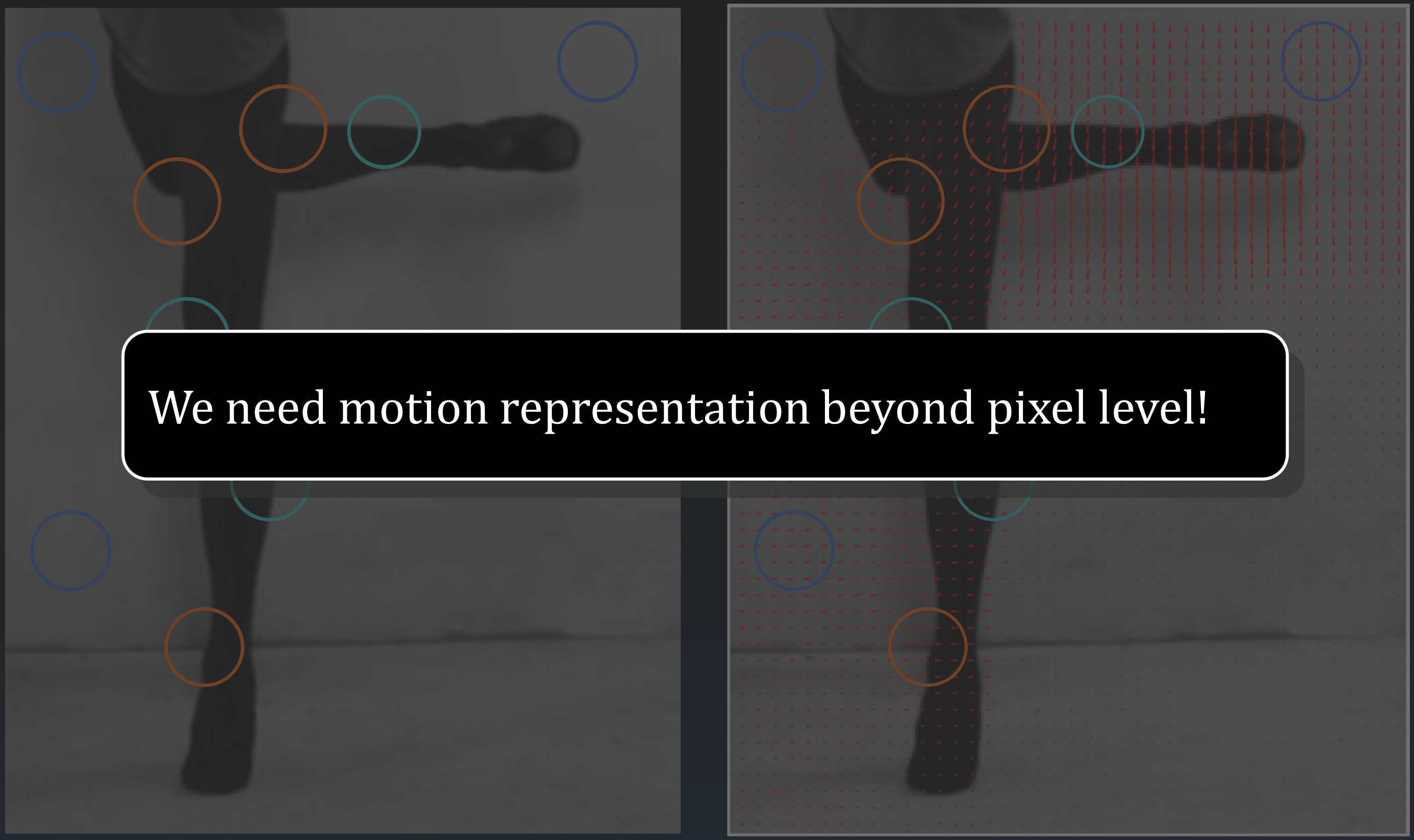
Optical flow representation: aperture problem





Corners	Lines	Flat regions
Spurious junctions	Boundary ownership	Illusory boundaries

Optical Flow Representation



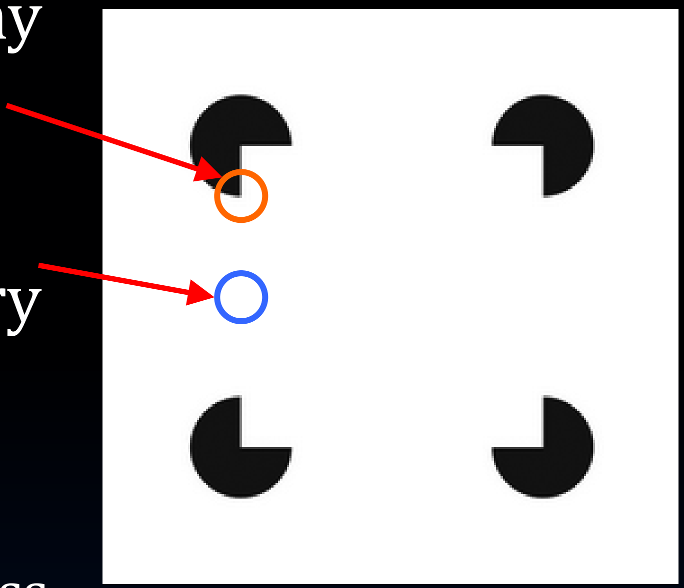
We need motion representation beyond pixel level!

Corners | Lines | Flat regions

Spurious junctions | Boundary ownership | Illusory boundaries

Challenge: Textureless Objects under Occlusion

- Corners are not always trustworthy (junctions)
- Flat regions do not always move smoothly (discontinuous at illusory boundaries)
- How about boundaries?
  - Easy to detect and track for textureless objects
  - Able to handle junctions with illusory boundaries



Analysis of Contour Motions

- Our approach: simultaneous grouping and motion analysis
  - Multi-level contour representation
  - Junctions are appropriated handled
  - Formulate graphical model that favors good contour and motion criteria
  - Inference using importance sampling
- Contribution
  - An important component in motion analysis toolbox for textureless objects under occlusion

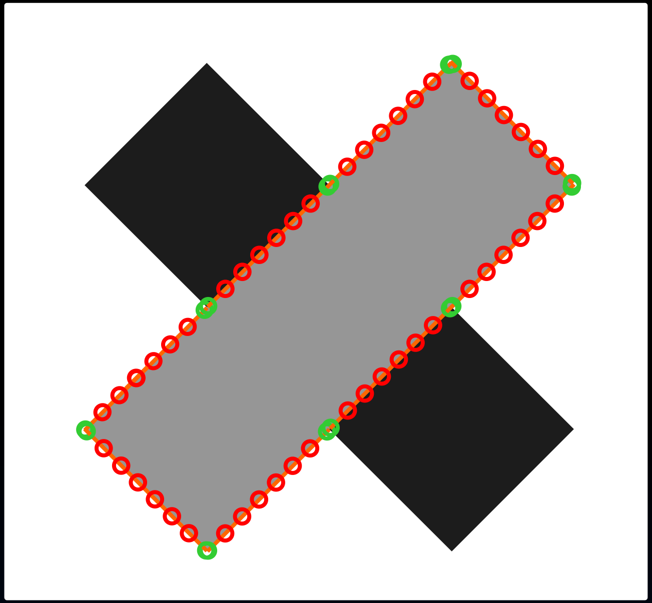
C. Liu, W. T. Freeman and E. H. Adelson. NIPS 2006

Three Levels of Contour Representation

– ***Edgelets***: edge particles

– ***Boundary fragments***: a chain of edgelets with small curvatures

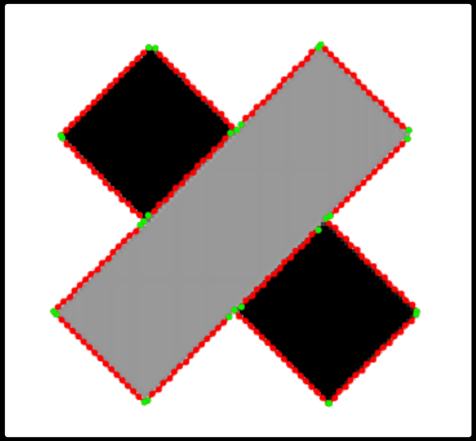
– ***Contours***: a chain of boundary fragments



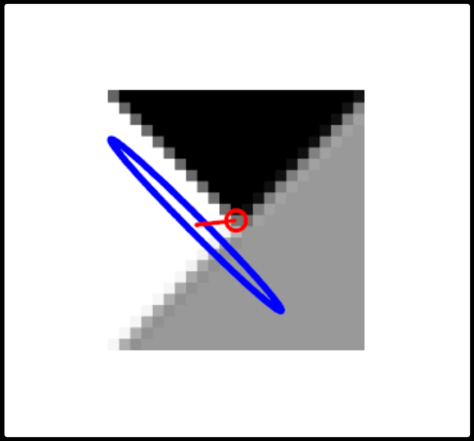
Forming boundary fragments: easy (for textureless objects)

Forming contours: hard (the focus of our work)

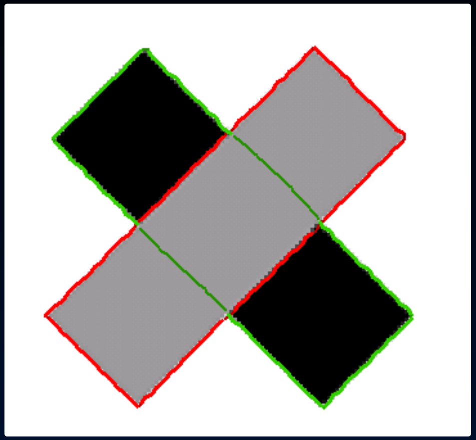
Overview of our system



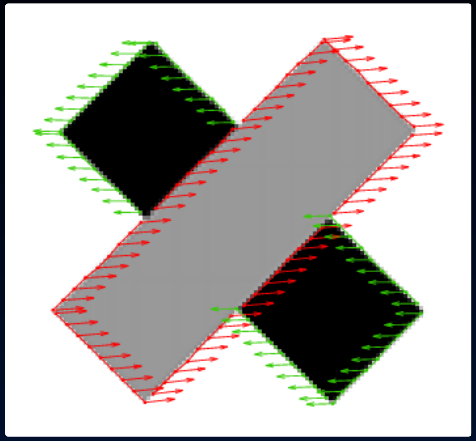
1. Extract boundary fragments



2. Edgelet tracking with uncertainty.

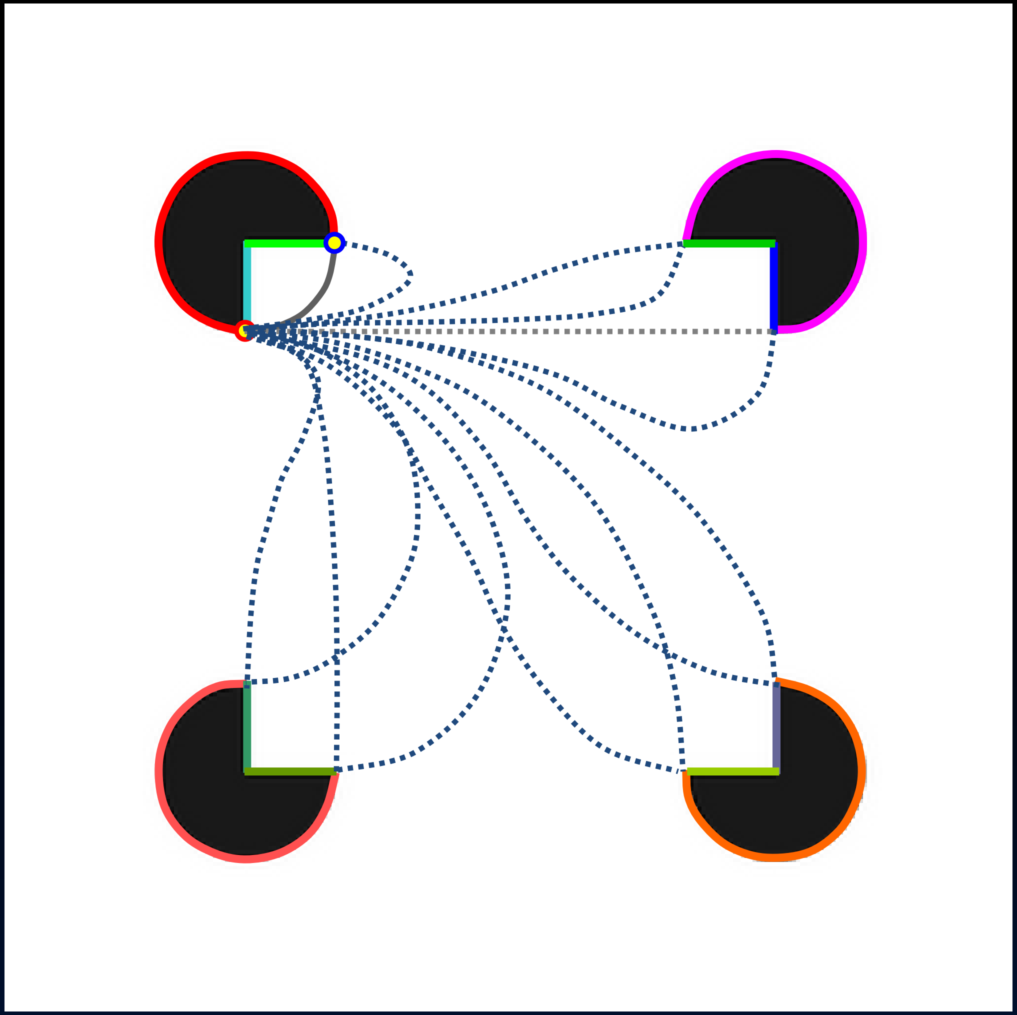


3. Boundary grouping and illusory boundary

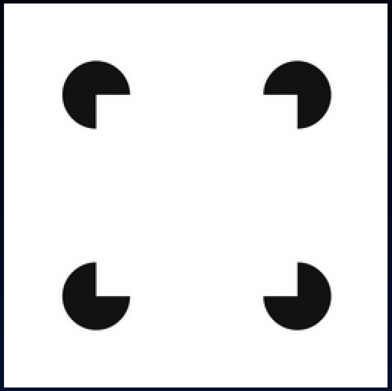


4. Motion estimation based on the grouping

Local Spatial-Temporal Cues for Grouping

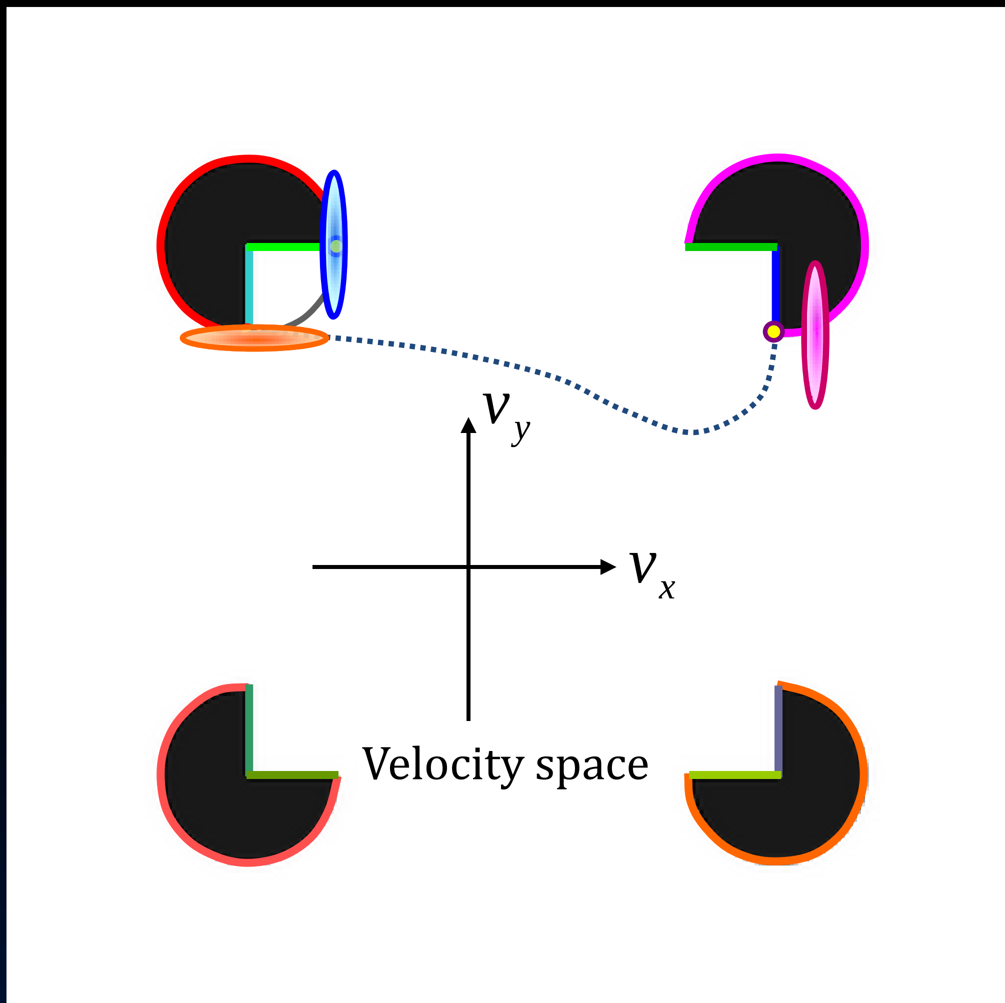


Illusory boundaries corresponding to the groupings (generated by spline interpolation)

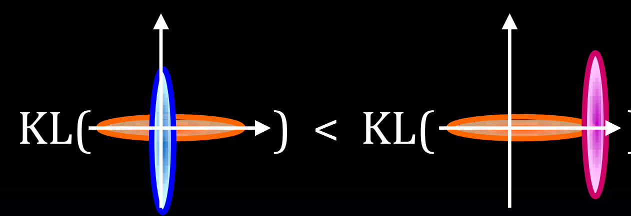


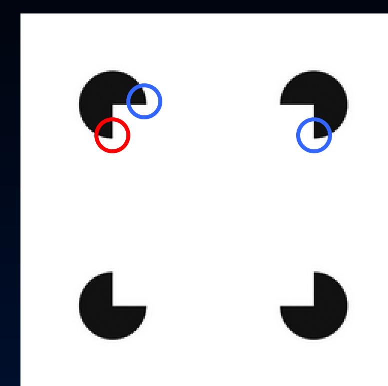
Motion stimulus

Local spatial-temporal cues for grouping: (a) Motion similarity



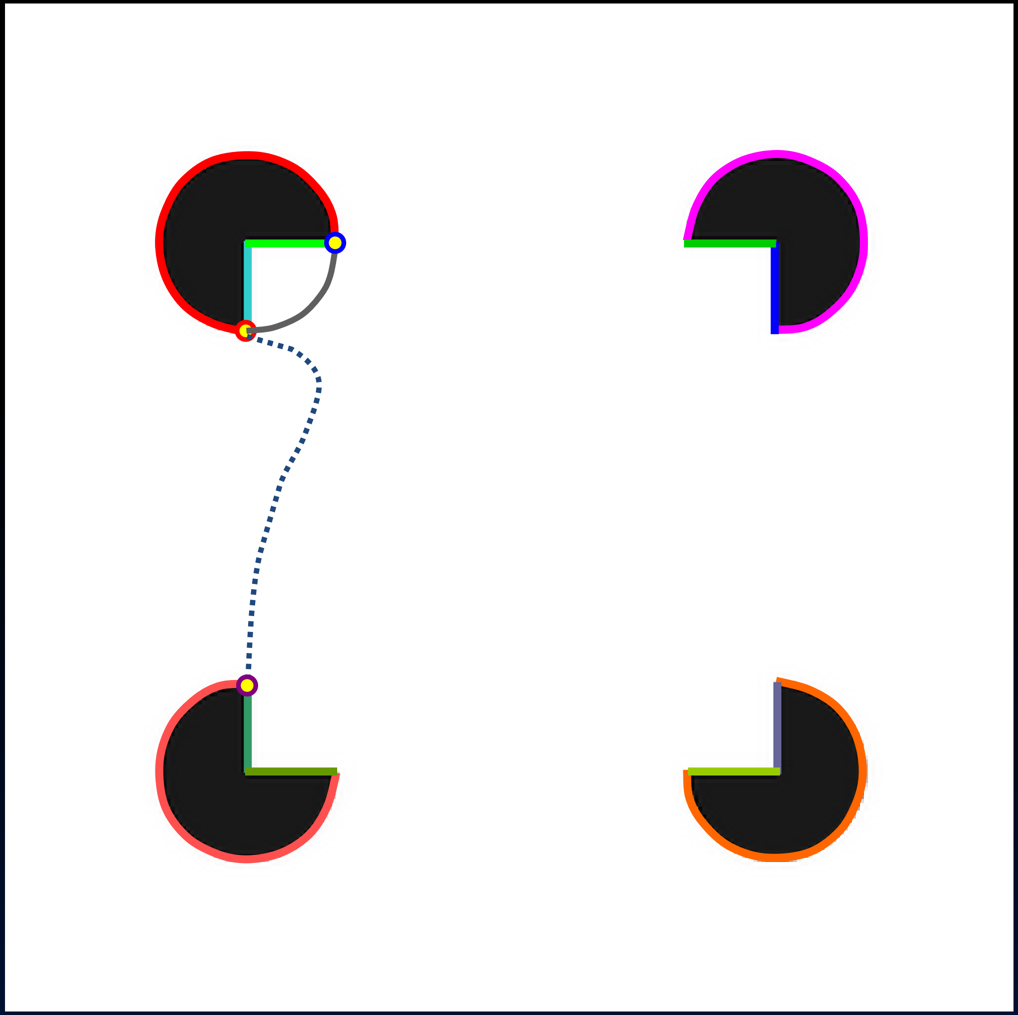
The grouping with higher motion similarity is favored



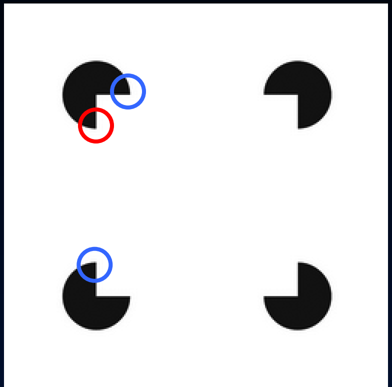


Motion stimulus

Local spatial-temporal cues for grouping: (b) Curve smoothness

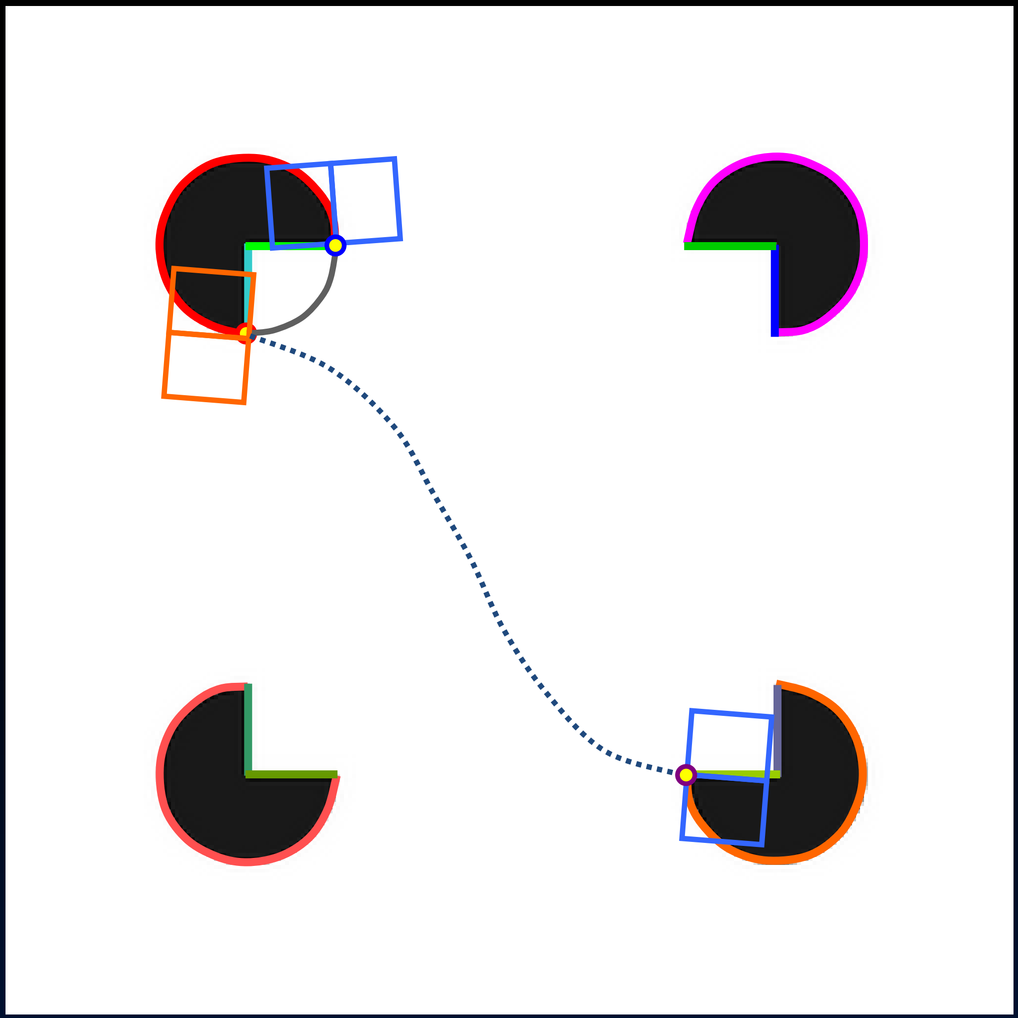


The grouping with smoother and shorter illusory boundary is favored

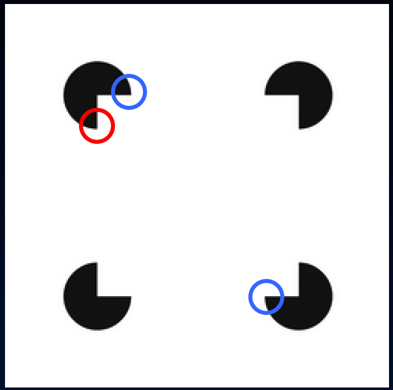


Motion stimulus

Local spatial-temporal cues for grouping: (c) Contrast consistency



The grouping with consistent local contrast is favored



Motion stimulus

The Graphical Model for Grouping

- Affinity metric $\lambda(S(i,t_i);\mathbf{B},O)$ terms
  - (a) Motion similarity

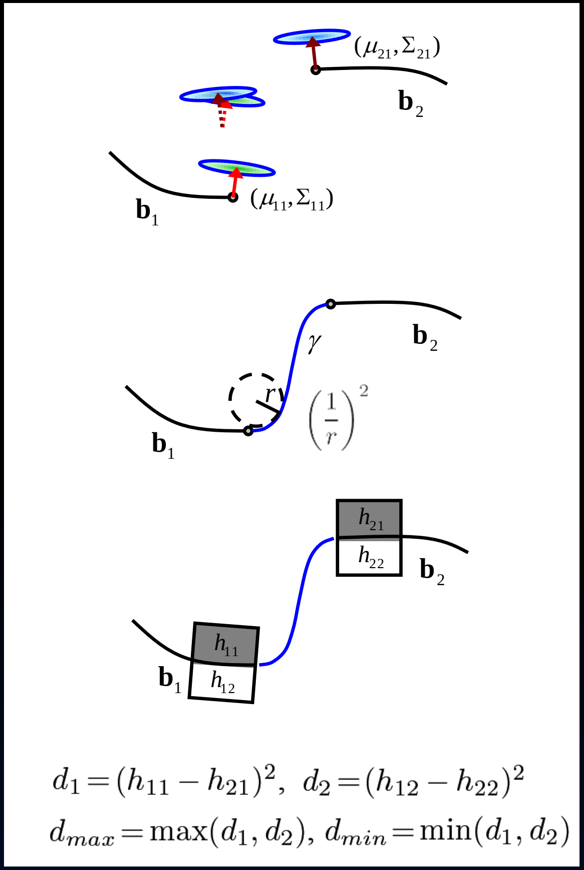
    $$\exp\{-\alpha_{KL} KL\big(N(\mu_{11},\Sigma_{11}),N(\mu_{21},\Sigma_{21})\big)\}$$

  - (b) Curve smoothness

    $$\exp\left\{-\alpha_r \int_r \left(\frac{d\theta}{ds}\right)^2 ds\right\}$$

  - (c) Contrast consistency

    $$\exp\left\{-\frac{d_{max}}{2\sigma_{max}^2}-\frac{d_{min}}{2\sigma_{min}^2}\right\}$$



- The graphical model for grouping

$$\Pr(\mathbf{S};\mathbf{B},O)=\frac{1}{Z_S}\prod_{i=1}^{N}\prod_{t_i=0}^{1}\underbrace{\lambda(S(i,t_i);\mathbf{B},O)}_{\text{affinity}}\underbrace{\delta\big[S\big(S(i,t_i)\big)-(i,t_i)\big]}_{\text{reversibility}}$$

affinity reversibility no self-intersection

Motion estimation for grouped contours

- Gaussian MRF (GMRF) within a boundary fragment

$$\varphi(v_i; b_i) = \prod_{k=1}^{n_i} \exp\left\{-(v_{ik} - \mu_{ik})^T \sum_{ik}^{-1} (v_{ik} - \mu_{ik})\right\} \prod_{k=1}^{n_i - 1} \exp\left\{-\frac{1}{2\sigma^2} \left|v_{ik} - v_{i,k+1}\right|^2\right\}$$

- The motions of two end edgelets are similar if they are grouped together

$$\phi\left(\mathbf{V}(i, t_i), \mathbf{V}\big(S(i, t_i)\big)\right) = \begin{cases} 1 & \text{if } S(i, t_i) = (i, t_i) \\ \exp\left\{-\frac{1}{2\sigma^2} \left|\mathbf{V}(i, t_i) - \mathbf{V}\big(S(i, t_i)\big)\right|^2\right\} & \text{otherwise} \end{cases}$$

- The graphical model of motion: joint Gaussian given the grouping

$$\Pr(\mathbf{V}|\mathbf{S}; \mathbf{B}) = \frac{1}{Z_V} \prod_i^N \varphi(\mathbf{v}_i; \mathbf{b}_i) \prod_{t_i}^1 \phi\left(\mathbf{V}(i, t_i), \mathbf{V}\big(S(i, t_i)\big)\right)$$

This problem is solved in early work: Y. Weiss, Interpreting images by propagating Bayesian beliefs, NIPS, 1997.

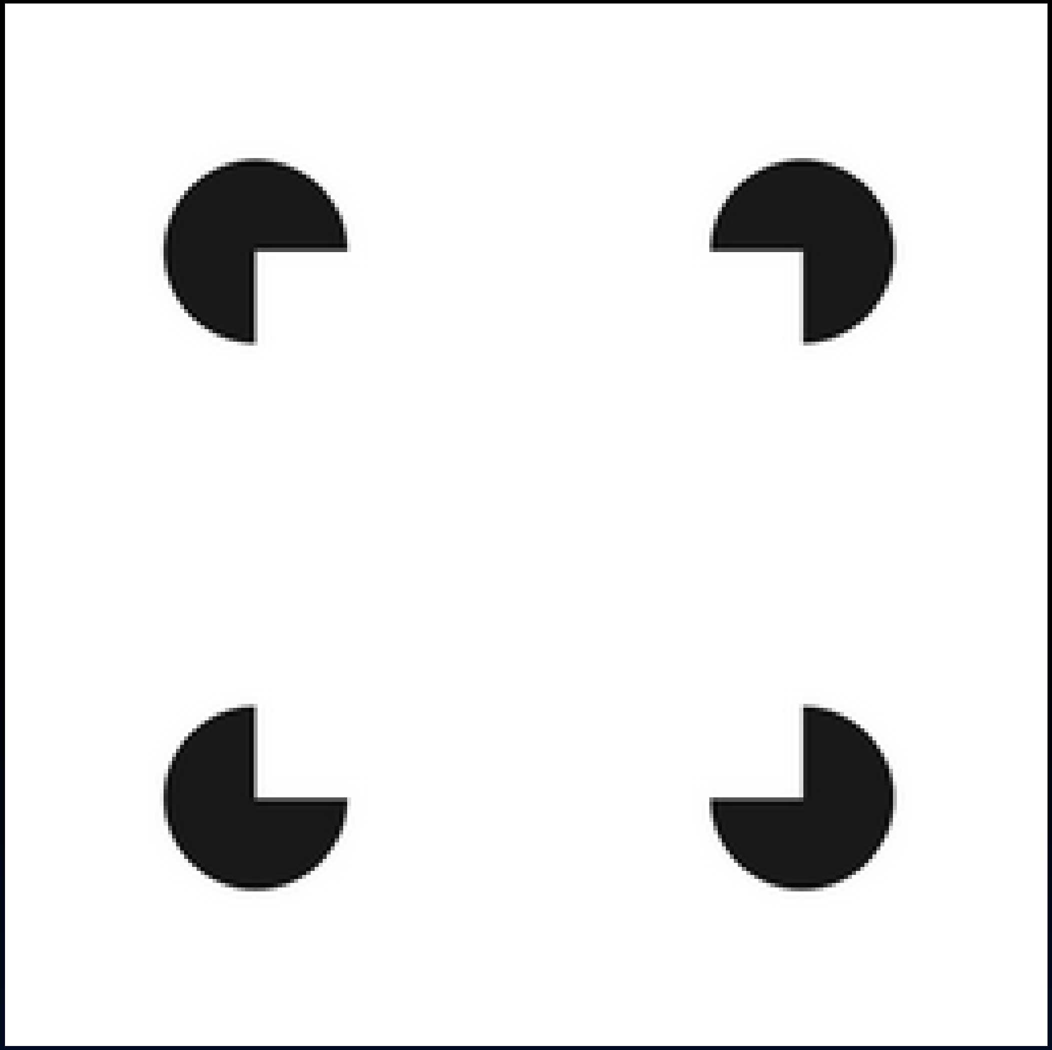
Inference

- Two-step inference
  - Grouping (switch variables)
  - Motion based on grouping (easy, least square)
- Grouping: importance sampling to estimate the marginal of the switch variables
  - Bidirectional proposal density

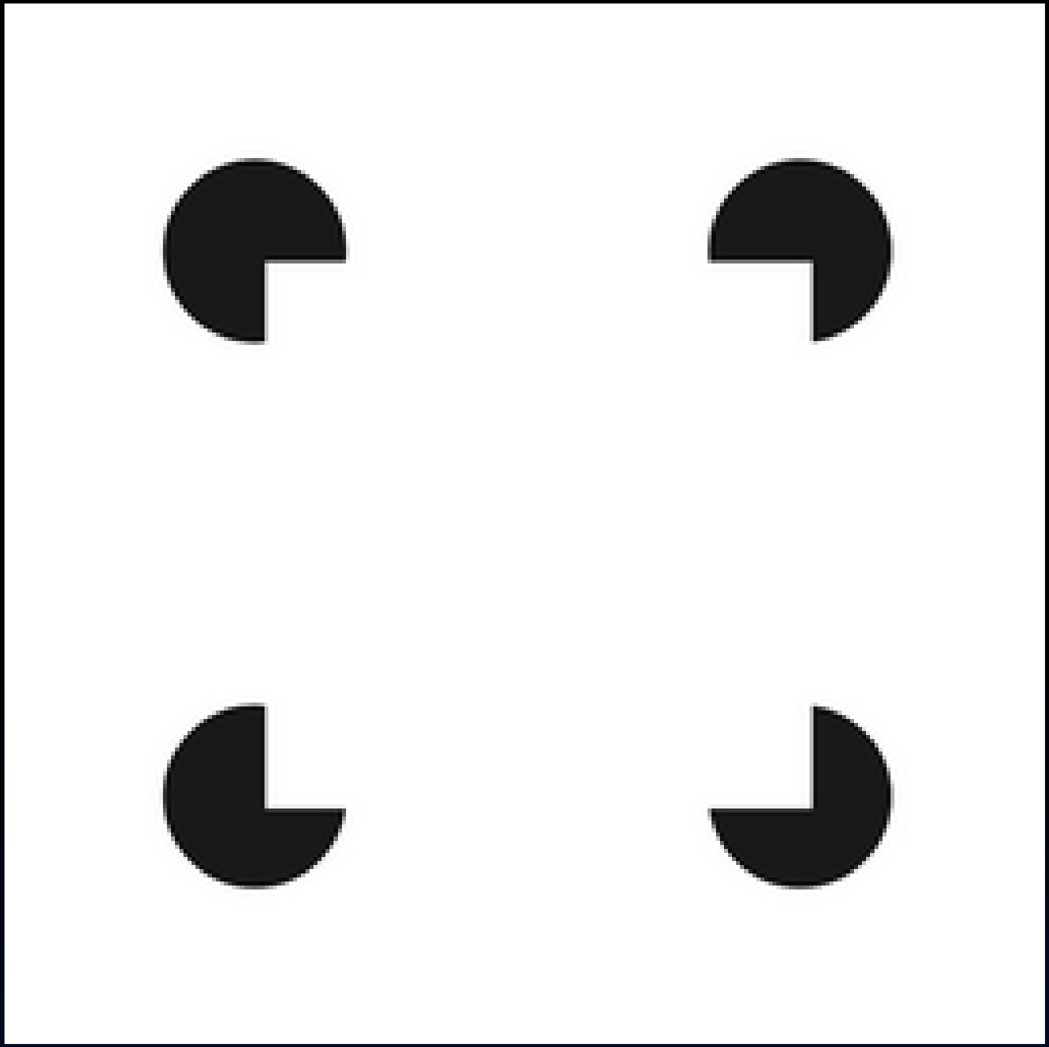
$$q\left(S(i,t_i)=(j,t_j)\right) \propto \frac{1}{Z_q}\lambda\left(S(i,t_i)=(j,t_j)\right)\lambda\left(S(j,t_j)=(i,t_i)\right)$$

  - Toss the sample if self-intersection is detected
- Obtain the optimal grouping from the marginal

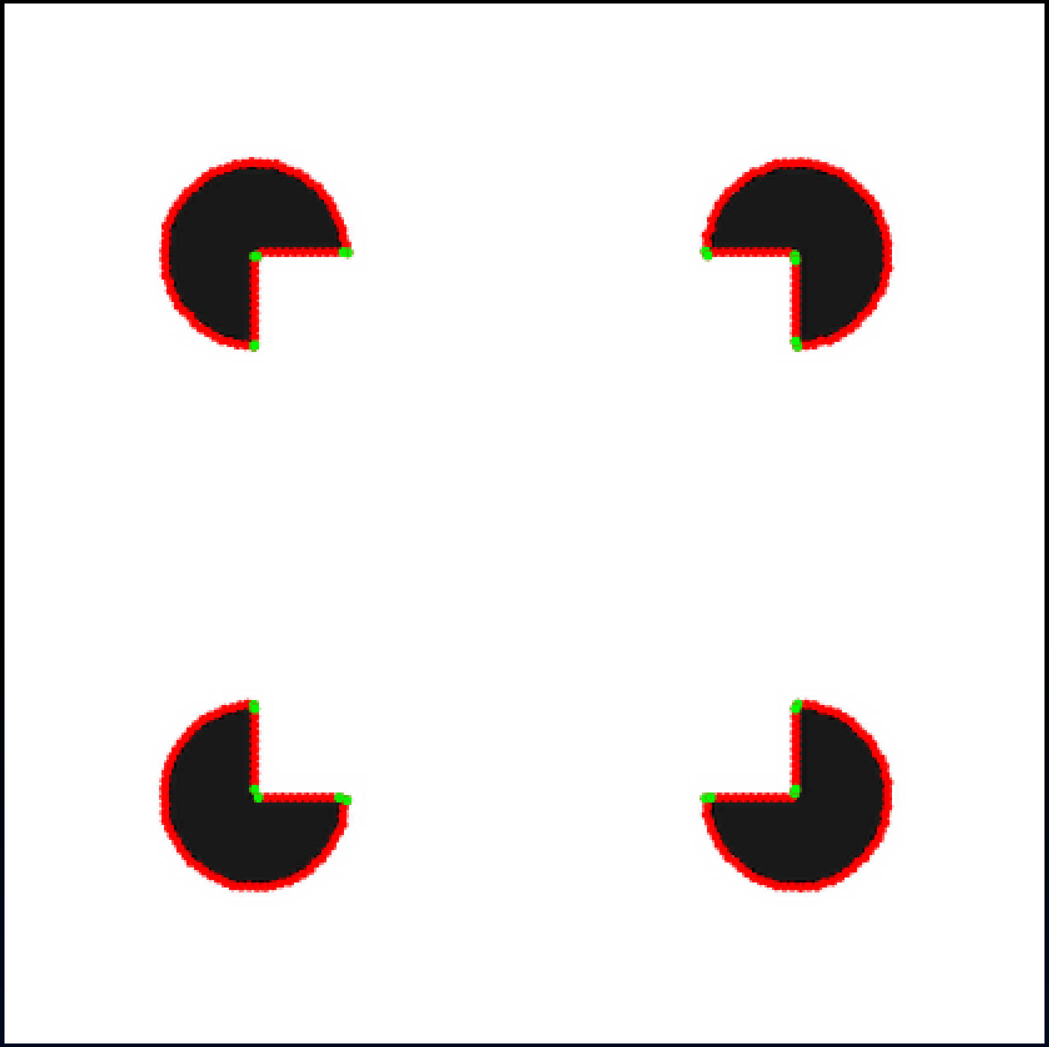
Kanizsa Square



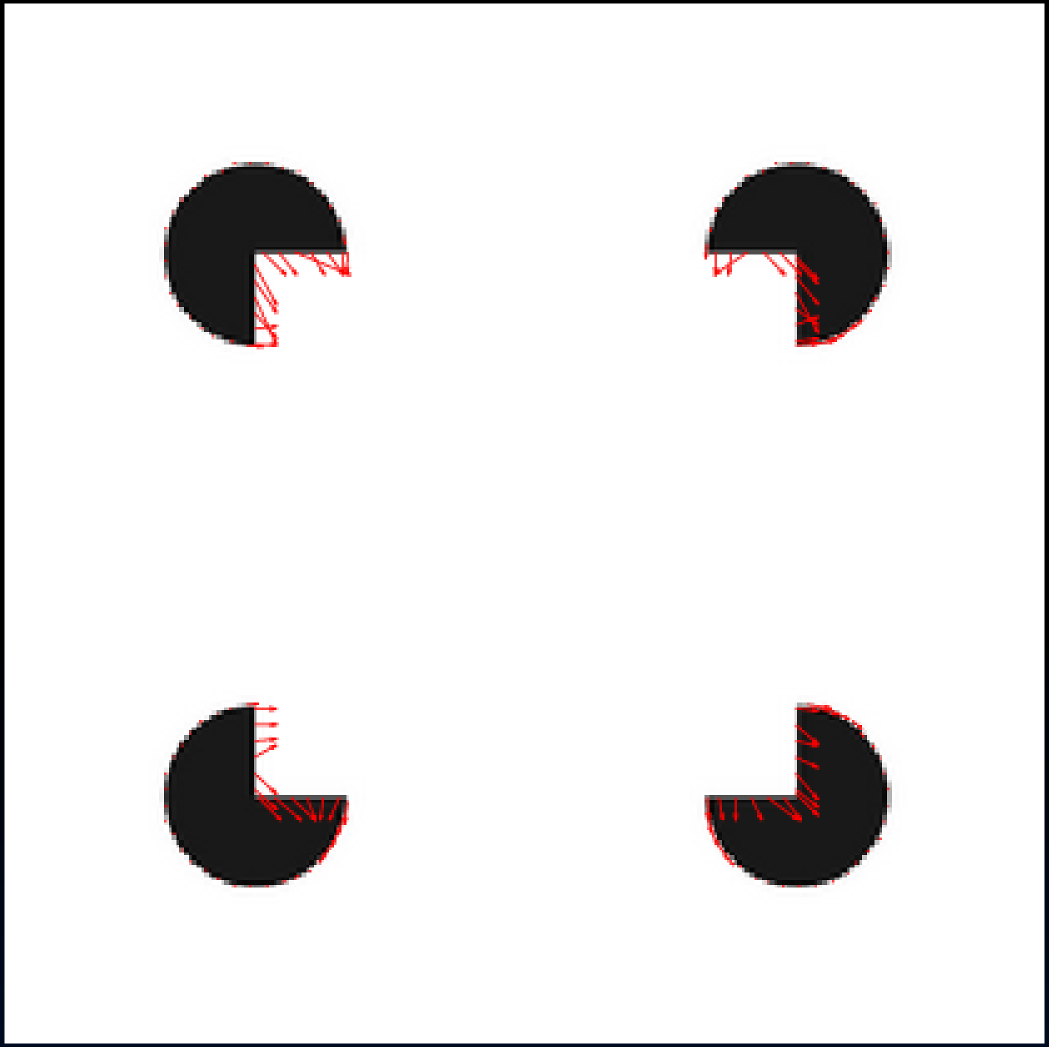
Frame 1



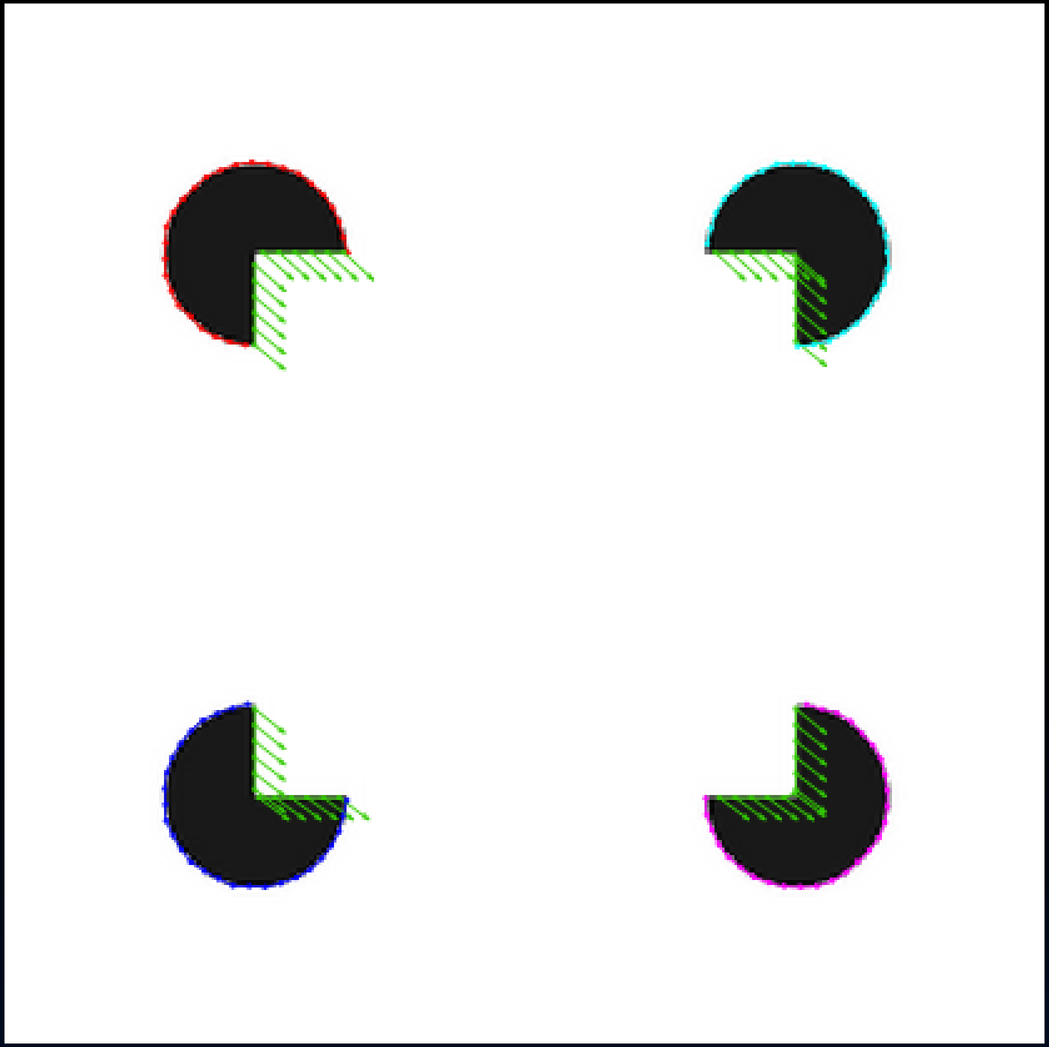
Frame 2



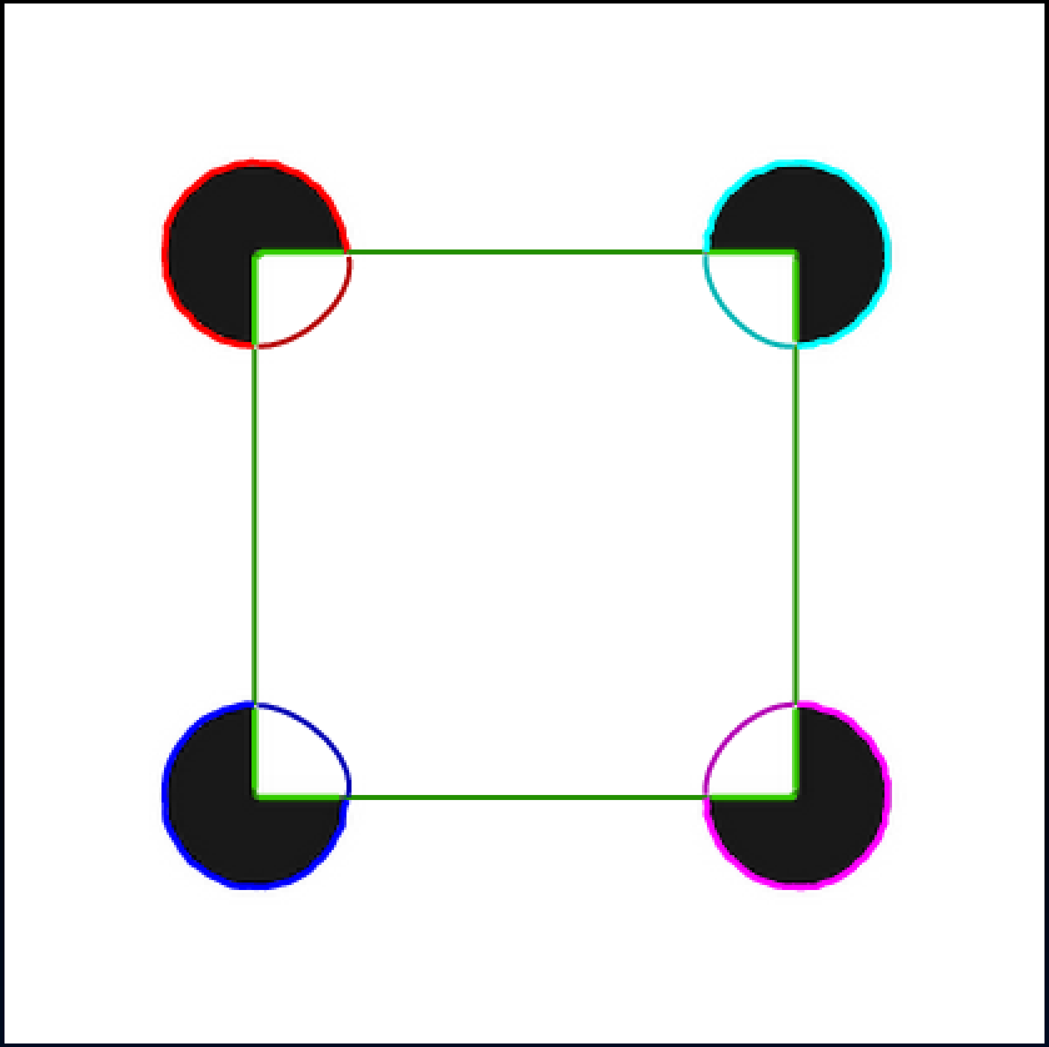
Extracted boundary fragments



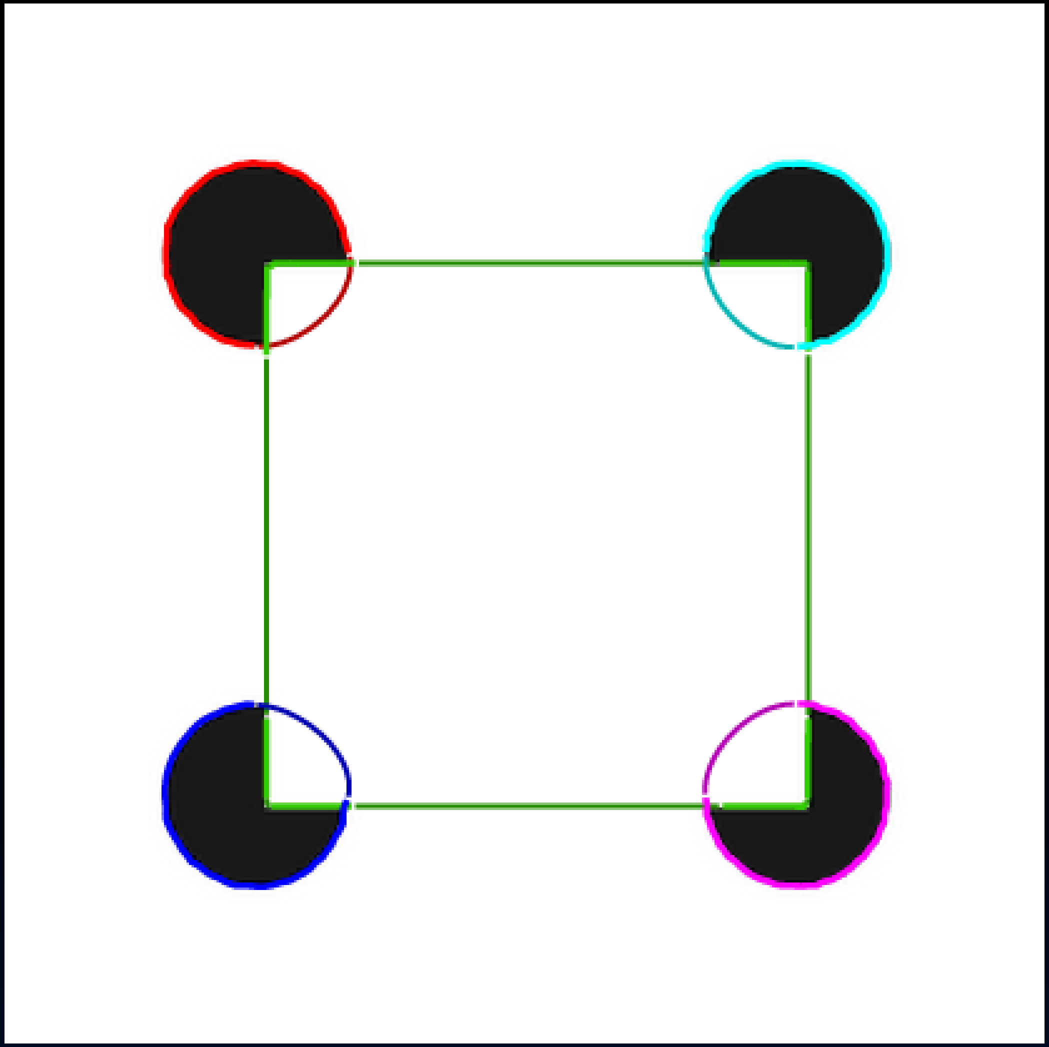
Optical flow from Lucas-Kanade algorithm



Estimated motion by our system, after grouping



Boundary grouping and illusory boundaries (frame 1)



Boundary grouping and illusory boundaries (frame 2)

Rotating Chair

Frame 1

Frame 2

Extracted boundary fragments

Estimated flow field from Brox et al.

Estimated motion by our system, after grouping

Boundary grouping and illusory boundaries (frame 1)

Boundary grouping and illusory boundaries (frame 2)

Content

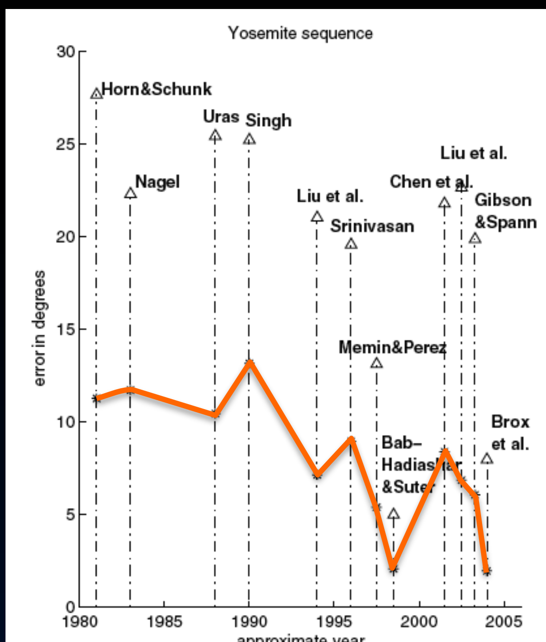
- Discrete optical flow
- Layer motion analysis
- Contour motion analysis
- Obtaining motion ground truth

How to evaluate optical flow?

- Assume the ground truth motion is known
- Average Angular Error (AAE)
  - Let $w = (u, v, 1)$
  - Angular error: $\arccos(\frac{w^T w_0}{\|w\| \|w_0\|})$
- Error in flow endpoint (EP)
  - EP: $\sqrt{(u - u_0)^2 + (v - v_0)^2}$
- Other metrics

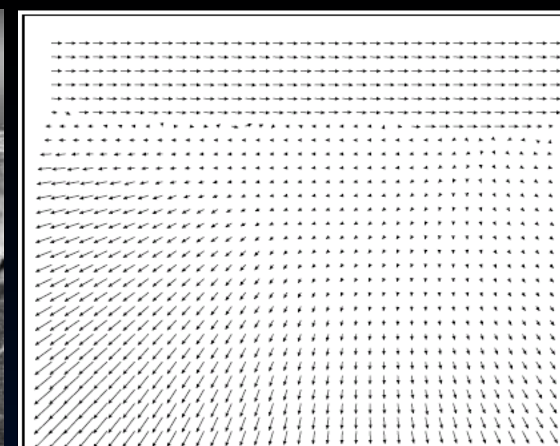
Is optical flow solved

- The AAE (average angular error) race on the *Yosemite* sequence for over 15 years



Improvement#

Yosemite sequence



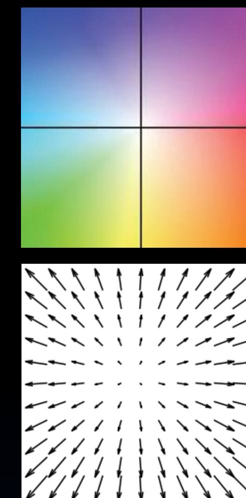
State-of-the-art optical flow*

#I. Austvoll. Lecture Notes in Computer Science, 2005
*Brox *et al. ECCV*, 2004.

But when optical flow is applied to real-life videos…

A sample sequence

State-of-the-art optical flow

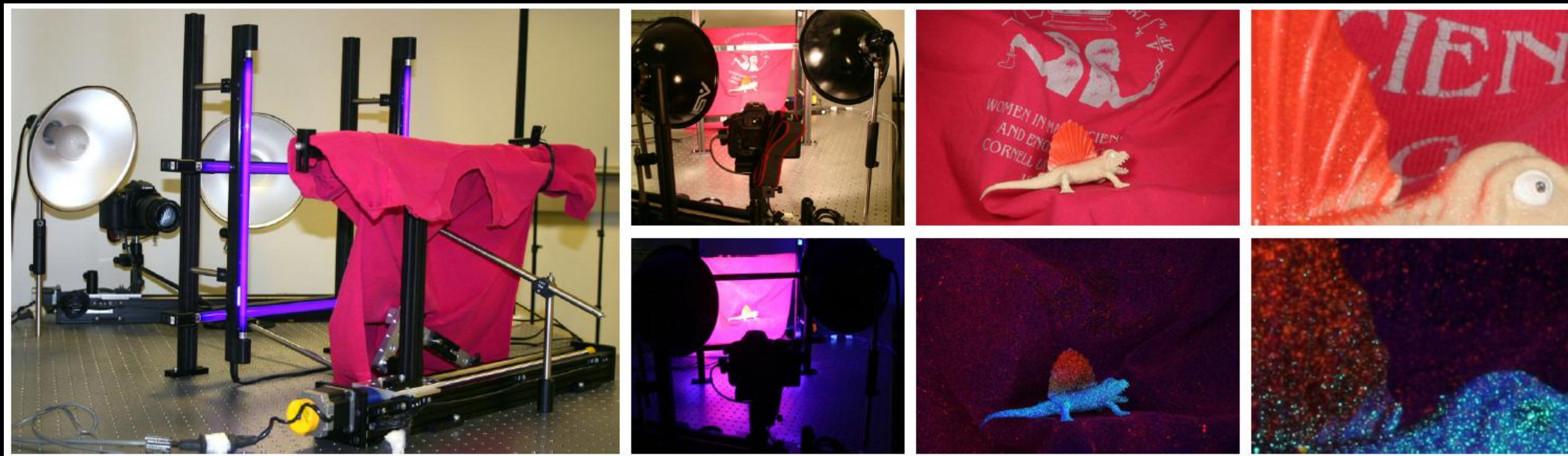


Flow visualization color map

Optical flow is far from being solved:

- Often fails to capture occluding boundaries correctly
- Puzzles on the right choice of smoothness

Middlebury flow database



Baker et. al. A Database and Evaluation Methodology for Optical Flow. ICCV 2007

Middlebury flow database

Optical flow evaluation results

Statistics: Average SD R0.5 R1.0 R2.0 A50 A75 A95

Error type: endpoint angle interpolation normalized interpolation

Average endpoint error	avg.	Army (Hidden texture) GT im0 im1			Mequon (Hidden texture) GT im0 im1			Schefflera (Hidden texture) GT im0 im1			Wooden (Hidden texture) GT im0 im1			Grove (Synthetic) GT im0 im1			Urban (Synthetic) GT im0 im1			Yosemite (Synthetic) GT im0 im1			Teddy (Stereo) GT im0 im1		
	rank	all	disc	untext	all	disc	untext	all	disc	untext	all	disc	untext	all	disc	untext	all	disc	untext	all	disc	untext	all	disc	untext
Adaptive [20]	4.4	**0.09** 1	**0.26** 1	**0.06** 1	0.23 5	0.78 4	0.18 5	0.54 8	[illegible] 10	0.21 3	**0.18** 1	0.91 3	**0.10** 1	0.88 3	1.25 3	0.73 5	0.50 3	1.28 3	0.31 3	0.14 10	0.16 12	0.22 10	0.65 3	1.37 3	0.79 4
Complementary OF [21]	5.7	0.11 5	0.28 3	0.10 9	**0.18** 1	**0.63**	**0.12** 1	0.31 3	0.75	**0.18** 1	0.19 2	0.97 5	0.12 3	0.97 10	1.31 6	1.00 11	1.78 20	1.73 7	0.87 14	0.11 4	0.12 2	0.22 10	0.68 4	1.48 4	0.95 8
Aniso. Huber-L1 [22]	5.8	0.10 3	0.28 3	0.08 3	0.31 11	0.88	0.28 12	[illegible] 10	1.13	0.29 12	0.20 4	0.92 4	0.13 5	0.84 2	1.20 2	0.70 2	**0.39** 1	**1.23** 1	**0.28** 1	0.17 15	0.15 9	0.27 16	0.64 2	1.36 2	0.79 4
DPOF [18]	6.1	0.13 12	0.35 12	0.09 4	0.25 6	0.79 5	[illegible]	**0.24**	**0.49** 1	0.21 3	0.19 2	**0.62** 1	0.15 11	**0.74** 1	**1.09** 1	**0.49** 1	0.66 7	1.80 10	0.63 8	0.19 17	0.17 14	0.35 20	**0.50** 1	**1.08** 1	**0.55** 1
TV-L1-improved [17]	7.2	**0.09** 1	**0.26** 1	0.07 2	0.20 3	0.71 3	0.16 2	0.53 7	1.18 9	0.22 5	0.21 7	1.24 11	0.11 2	0.90 4	1.31 6	0.72 3	1.51 14	1.93 11	0.84 11	0.18 16	0.17 14	0.31 17	0.73 8	1.62 9	0.87 7
CBF [12]	7.8	0.10 3	0.28 3	0.09 4	0.34 12	0.80 6	0.37 13	0.43 5	0.95 5	0.26 8	0.21 7	1.14 8	0.13 5	0.90 4	1.27 4	0.82 7	0.41 2	**1.23** 1	0.30 2	0.23 22	0.19 20	0.39 21	0.76 9	1.56 6	1.02 9
Brox et al. [5]	8.4	0.11 5	0.32 8	0.11 12	0.27 9	0.93 10	0.22 9	0.39 4	0.94 4	0.24 7	0.24 9	1.25 12	0.13 5	1.10 13	1.39 12	1.43 17	0.89 8	1.77 8	0.55 7	0.10 2	0.13 4	**0.11** 1	0.91 11	1.83 12	1.13 12
Rannacher [23]	8.5	0.11 5	0.31 6	0.09 4	0.25 6	0.84 7	0.21 8	0.57 12	1.27 15	0.26 8	0.24 9	1.32 14	0.13 5	0.91 7	1.33 8	0.72 3	1.49 13	1.95 13	0.78 9	0.15 12	0.14 7	0.26 13	0.69 6	1.58 8	0.86 6
F-TV-L1 [15]	8.8	0.14 13	0.35 12	0.14 15	0.34 12	0.98 12	0.26 11	0.59 14	1.19 10	0.26 8	0.27 13	1.36 15	0.16 12	0.90 4	1.30 5	0.76 6	0.54 4	1.62 6	0.36 4	0.13 6	0.15 9	0.20 9	0.68 4	1.56 6	0.66 2
Second-order prior [8]	9.0	0.11 5	0.31 6	0.09 4	0.26 8	0.93 10	0.20 7	0.57 12	1.25 14	0.26 8	0.20 4	1.04 6	0.12 3	0.94 8	1.34 9	0.83 9	0.61 6	1.93 11	0.47 6	0.20 18	0.16 12	0.34 19	0.77 10	1.64 10	1.07 10
Fusion [6]	9.4	0.11 5	0.34 10	0.10 9	0.19 2	0.69 2	0.16 2	0.29 2	0.66 2	0.23 6	0.20 4	1.19 10	0.14 9	1.07 11	1.42 13	1.22 13	1.35 10	1.49 5	0.86 13	0.20 18	0.20 21	0.26 13	1.07 14	2.07 16	1.39 16
Dynamic MRF [7]	11.1	0.12 11	0.34 10	0.11 12	0.22 4	0.89 9	0.16 2	0.44 6	1.13 7	0.20 2	0.24 9	1.29 13	0.14 9	1.11 14	1.52 17	1.13 12	1.54 15	2.37 20	0.93 15	0.13 6	0.12 2	0.31 17	1.27 18	2.33 20	1.66 17
SegOF [10]	11.7	0.15 14	0.36 14	0.10 9	0.57 15	1.16 15	0.59 19	0.68 15	1.24 12	0.64 14	0.32 15	0.86 2	0.26 15	1.18 17	1.50 16	1.47 18	1.63 18	2.09 14	0.96 16	**0.08** 1	0.13 4	0.12 2	0.70 7	1.50 5	0.69 3
Learning Flow [11]	13.3	0.11 5	0.32 8	0.09 4	0.29 10	0.99 13	0.23 10	0.55 9	1.24 12	0.29 12	0.36 16	1.56 17	0.25 14	1.25 19	1.64 21	1.41 16	1.55 17	2.32 19	0.85 12	0.14 10	0.18 18	0.24 12	1.09 15	2.09 18	1.27 13
Filter Flow [19]	14.3	0.17 16	0.39 16	0.13 14	0.43 14	1.09 14	0.38 14	0.75 16	1.34 16	0.78 19	0.70 19	1.54 16	0.68 19	1.13 16	1.38 11	1.51 19	0.57 5	1.32 4	0.44 5	0.22 20	0.23 23	0.26 13	0.96 12	1.66 11	1.12 11
GraphCuts [14]	14.5	0.16 15	0.38 15	0.14 15	0.59 18	1.36 19	0.46 15	0.56 10	1.07 6	0.64 14	0.26 12	1.14 8	0.17 13	0.96 9	1.35 10	0.84 10	2.25 23	1.79 9	1.22 21	0.22 20	0.17 14	0.43 22	1.22 17	2.05 15	1.78 19
Black & Anandan [4]	15.0	0.18 17	0.42 17	0.19 18	0.58 17	1.31 17	0.50 16	0.95 19	1.58 18	0.70 16	0.49 17	1.59 18	0.45 17	1.08 12	1.42 13	1.22 13	1.43 11	2.28 17	0.83 10	0.15 12	0.17 14	0.17 6	1.11 16	1.98 14	1.30 14
SPSA-learn [13]	15.7	0.18 17	0.45 18	0.17 17	0.57 15	1.32 18	0.51 17	0.84 17	1.50 17	0.72 17	0.52 18	1.64 19	0.49 18	1.12 15	1.42 13	1.39 15	1.75 19	2.14 15	1.06 20	0.13 6	0.13 4	0.19 7	1.32 19	2.08 17	1.73 18
GroupFlow [9]	15.9	0.21 19	0.51 19	0.21 19	0.79 21	1.69 21	0.72 21	0.86 18	1.64 19	0.74 18	0.30 14	1.07 7	0.26 15	1.29 22	1.81 22	0.82 7	1.94 21	2.30 18	1.36 22	0.11 4	0.14 7	0.19 7	1.06 13	1.96 13	1.35 15
2D-CLG [1]	17.4	0.28 21	0.62 22	0.21 19	0.67 20	1.21 16	0.70 20	1.12 21	1.80 21	0.99 22	1.07 22	2.06 21	1.12 22	1.23 18	1.52 17	1.62 22	1.54 15	2.15 16	0.96 16	0.10 2	**0.11** 1	0.16 4	1.38 20	2.26 19	1.83 20
Horn & Schunck [3]	18.6	0.22 20	0.55 20	0.22 21	0.61 19	1.53 20	0.52 18	1.01 20	1.73 20	0.80 20	0.78 20	2.02 20	0.77 20	1.26 20	1.58 19	1.55 20	1.43 11	2.59 22	1.00 18	0.16 14	0.18 18	0.15 3	1.51 21	2.50 21	1.88 21
TI-DOFE [24]	19.6	0.38 23	0.64 23	0.47 23	1.16 22	1.72 22	1.26 22	1.39 23	2.06 24	1.17 23	1.29 23	2.21 23	1.41 23	1.27 21	1.61 20	1.57 21	1.28 9	2.57 21	1.01 19	0.13 6	0.15 9	0.16 4	1.87 22	2.71 22	2.53 22
FOLKI [16]	22.6	0.29 22	0.73 24	0.33 22	1.52 23	1.96 24	1.80 23	1.23 22	2.04 23	0.95 21	0.99 21	2.20 22	1.08 21	1.53 23	1.85 23	2.07 23	2.14 22	3.23 24	1.60 23	0.26 23	0.21 22	0.68 23	2.67 23	3.27 23	4.32 23
Pyramid LK [2]	23.7	0.39 24	0.61 21	0.61 24	1.67 24	1.78 23	2.00 24	1.50 24	1.97 22	1.38 24	1.57 24	2.39 24	1.78 24	2.94 24	3.72 24	2.98 24	3.33 24	2.74 23	2.43 24	0.30 24	0.24 24	0.73 24	3.80 24	5.08 24	4.88 24

Move the mouse over the numbers in the table to see the corresponding images. Click to compare with the ground truth.

Measuring motion for real-life videos

- Challenging because of occlusion, shadow, reflection, motion blur, sensor noise and compression artifacts

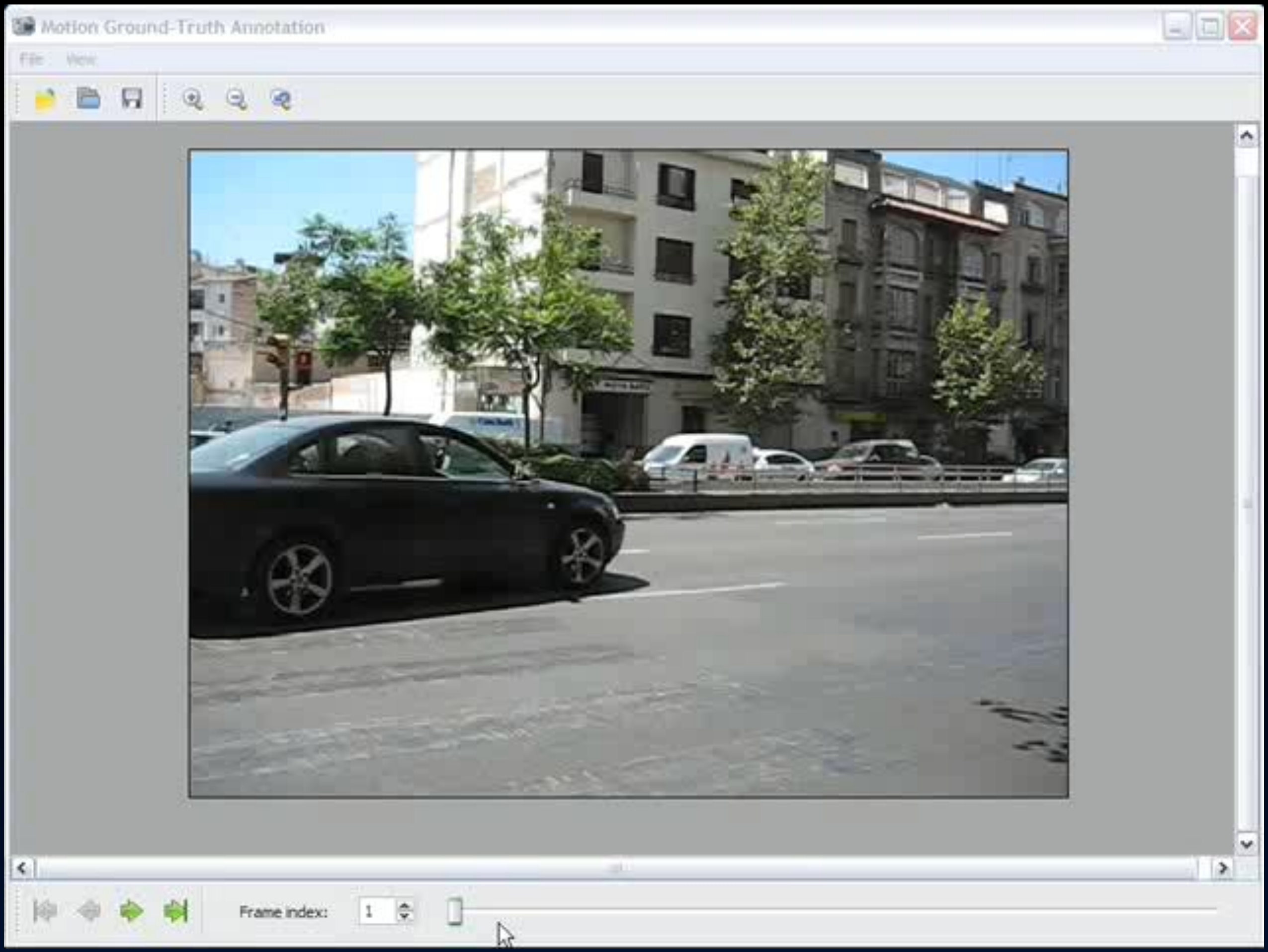
[Video courtesy: Antonio Torralba]

- Accurately measuring motion also has great impact in scientific measurement and graphics applications
- Humans are experts in perceiving motion. Can we use human expertise to annotate motion?

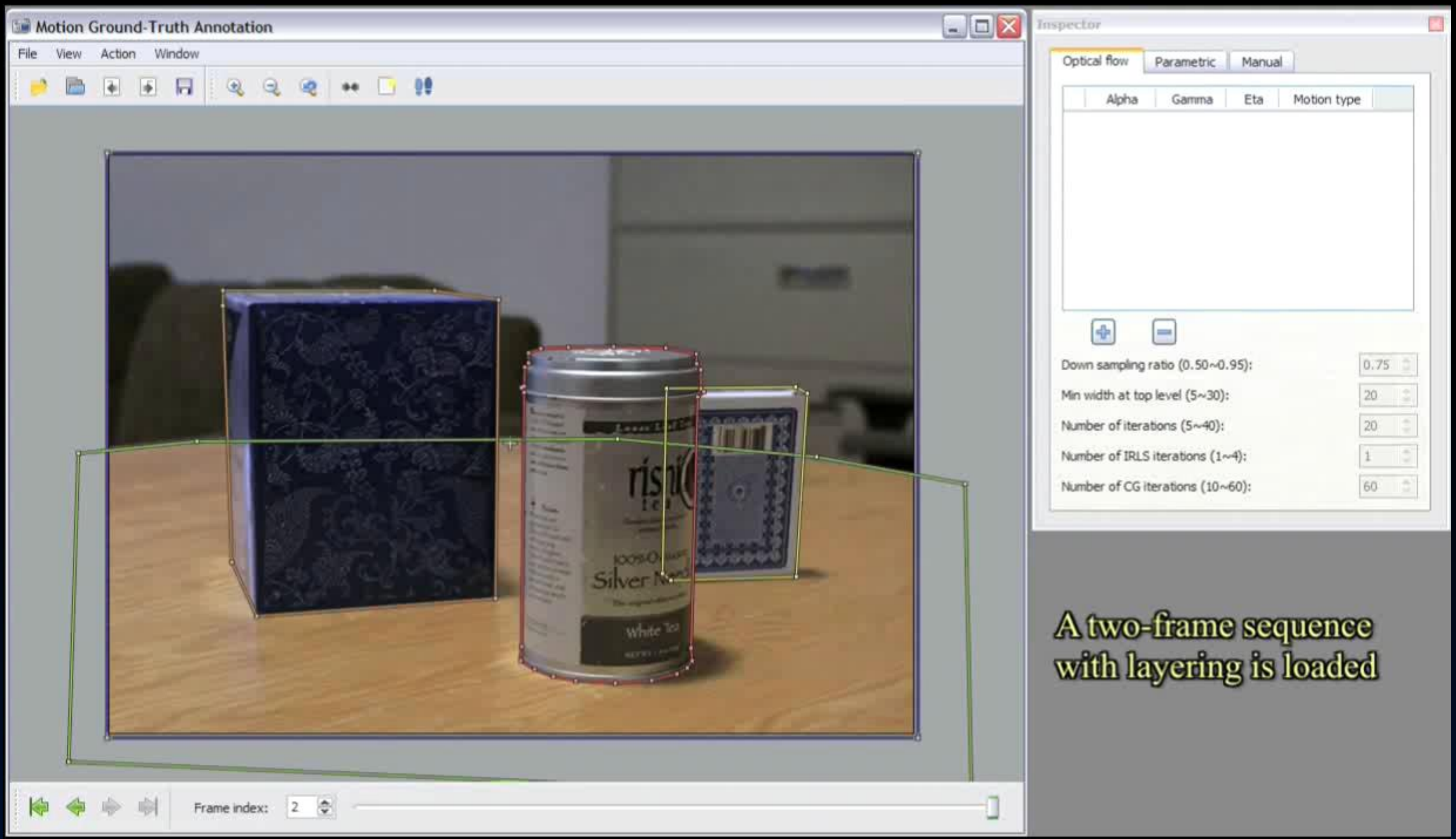
Human-assisted motion annotation

- Our approach: an interactive system to combine human perception and the state-of-the-art computer vision algorithms to annotate motion
- User layers as the interface for user interaction
  - Decompose a video sequence into layers
  - Motion analysis for each layer

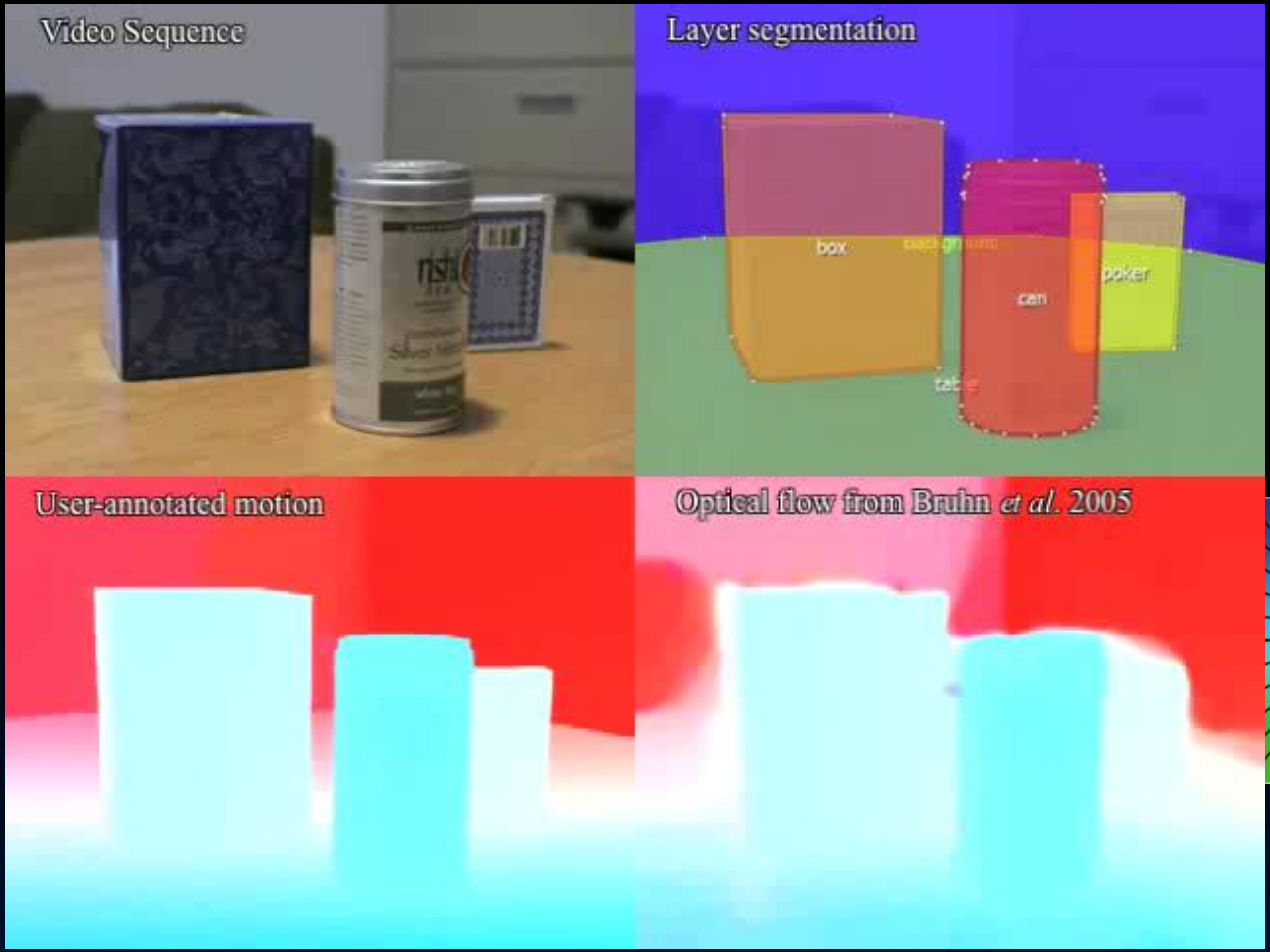
Demo: interactive layer segmentation

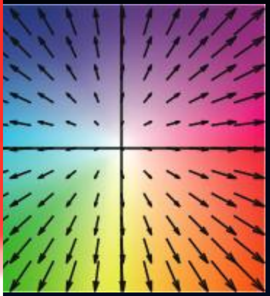


Demo: interactive motion labeling



Motion database of natural scenes

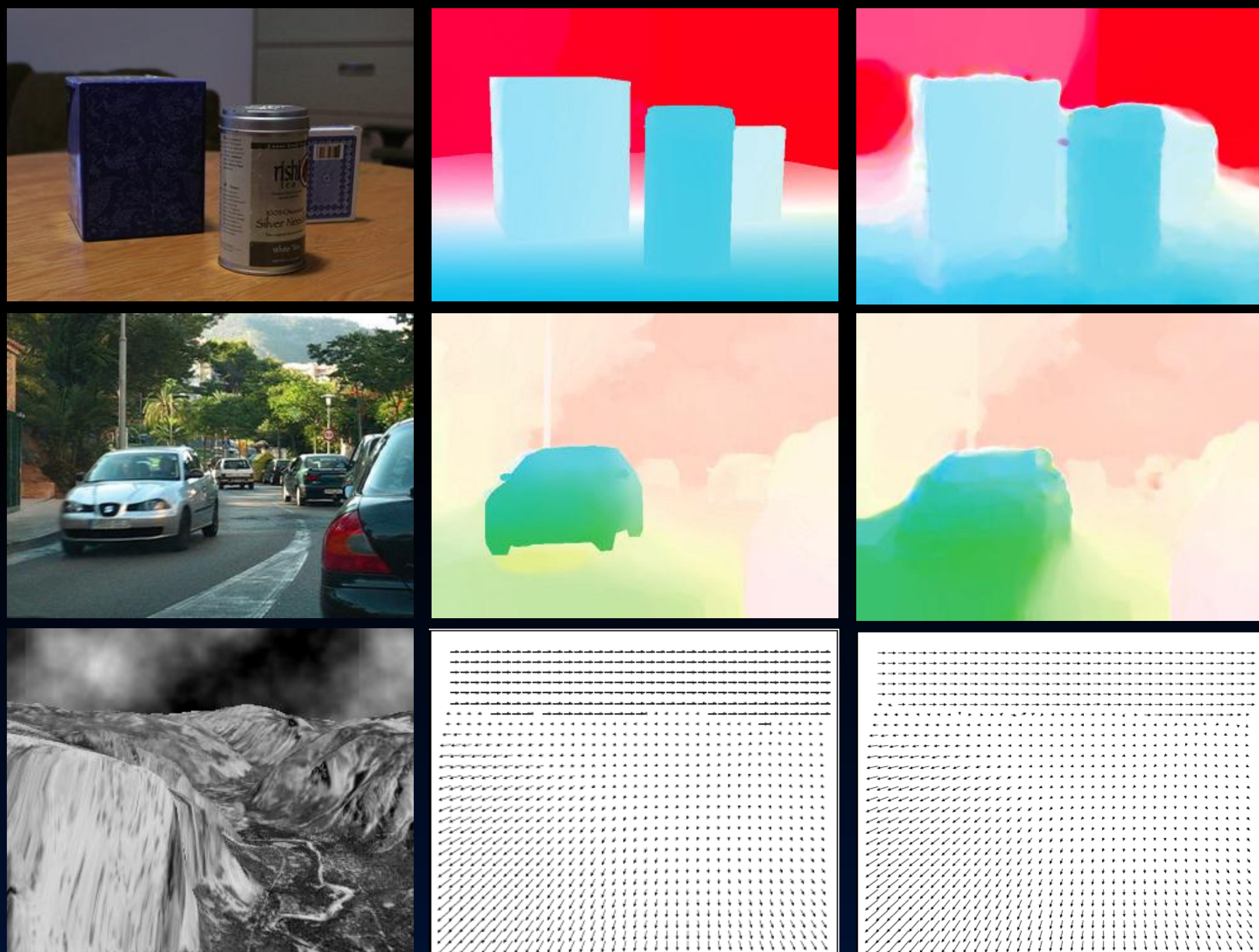




Color map

Bruhn *et al*. Lucas/Kanade meets Horn/Schunck: combining local and global optical flow methods. *IJCV*, 2005

Optical flow is far from being solved



AAE=8.99°

AAE=5.24°

AAE=1.94°

Frame Ground-truth motion Optical flow

LabelMe Video

J. Yuen, B. Russell, C. Liu and A. Torralba. ICCV, 2009

Summary

- Discrete optical flow matching
  - Tracking & motion interpolation
  - Belief propagation
- Other representations
  - Layer motion analysis
  - Contour motion analysis
- Obtaining motion ground truth
  - Human assisted motion annotation