FusionFlow
Discrete-Continuous Optimization for Optical Flow Estimation

Victor Lempitsky, MSR Cambridge
Stefan Roth, TU Darmstadt
Carsten Rother, MSR Cambridge

In Proc. IEEE Computer Vision and Pattern Recognition (CVPR), Anchorage, USA, June 2008

Presenter – Ankit Gupta

Outline

- Problem definition
- Previous work
- System overview
- Evaluation
- Conclusion

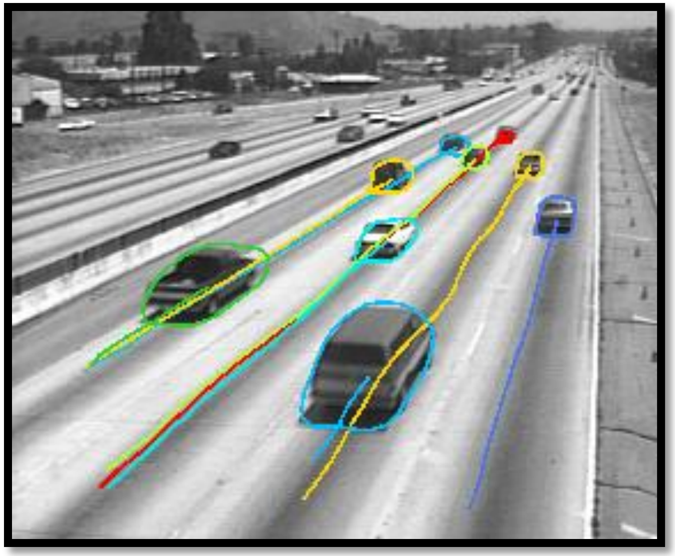
What is optical flow?

Two images I1 and I2

Where did a pixel in I1 go in I2?

Optical flow - Applications

- Tracking for surveillance
- Robotics
- Video editing
- 3D scene structure
- etc .....



Why isn’t it solved yet?

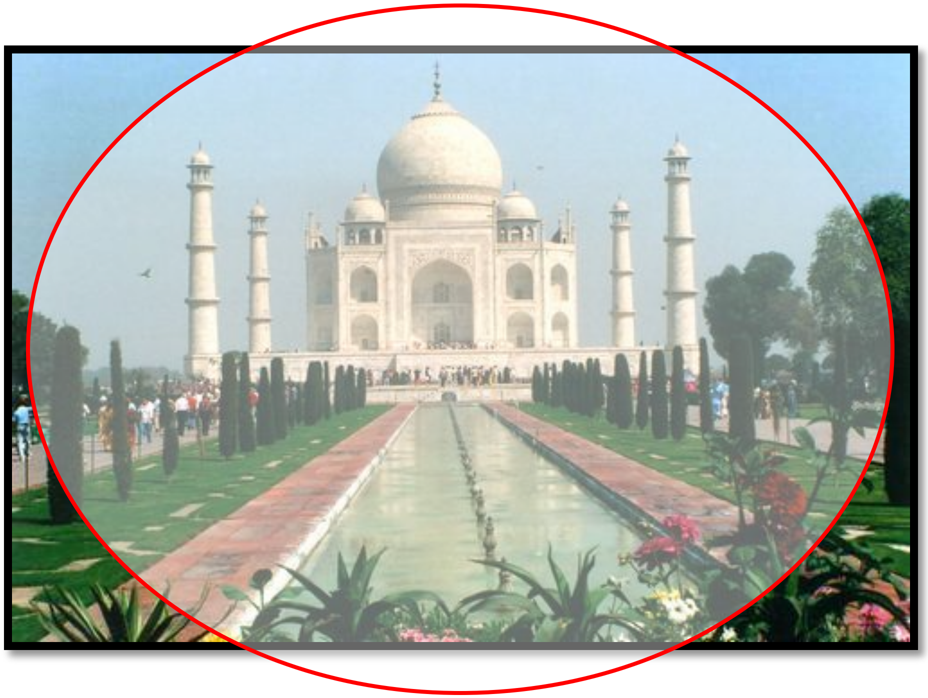
Why is optical flow hard?

Homogenous regions

Why is optical flow hard?

Deformations

Why is optical flow hard?



Lighting changes

Why is optical flow hard?

Occlusions

Outline

- Problem definition
- Previous work
- System overview
- Evaluation
- Conclusion

Lukas-Kanade [1981]

- Given images: F and G
- F(x+h) = G(x)
- To find h

$$E = \sum_x [F(x+h) - G(x)]^2$$

Use Taylor's expansion to linearize in 'h' and differentiate

$$h = \left[\sum_x \left(\frac{\partial F}{\partial x}\right)^T [G(x) - F(x)]\right]\left[\sum_x \left(\frac{\partial F}{\partial x}\right)^T \left(\frac{\partial F}{\partial x}\right)\right]^{-1}$$

Follow Newton-Raphson type iterations

Lukas-Kanade [1981]

$$h = \left[\sum_x \left(\frac{\partial F}{\partial x} \right)^T [G(x) - F(x)] \right] \left[\sum_x \left(\frac{\partial F}{\partial x} \right)^T \left(\frac{\partial F}{\partial x} \right) \right]^{-1}$$

What is summation on?

- Whole image – limited usefulness
- Small patch – Whole patch has same motion
- Single pixel – Ill conditioned

Horn-Schunck [1981]

Sequence of images as volume: E(x,y,t)

Illumination constancy constraint: dE/dt = 0

Each pixel has its own (u,v) flow vector

One constraint per pixel $(E_x, E_y) \cdot (u, v) = -E_t$

(after linearizing illumination constancy)

<u>**Aperture Problem**</u>

Horn-Schunck [1981]

<u>Countering the aperture problem</u>

Data term: $\mathscr{E}_b = E_x u + E_y v + E_t$

Smoothness term: $\mathscr{E}_c^2 = \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2$

Total energy to be minimized: $\mathscr{E}^2 = \iint (\alpha^2 \mathscr{E}_c^2 + \mathscr{E}_b^2)\, dx\, dy$

Minimized using differential calculus

$$(\alpha^2 + E_x^2 + E_y^2)u = +(\alpha^2 + E_y^2)\bar{u} - E_x E_y \bar{v} - E_x E_t$$
$$(\alpha^2 + E_x^2 + E_y^2)v = -E_x E_y \bar{u} + (\alpha^2 + E_x^2)\bar{v} - E_y E_t$$

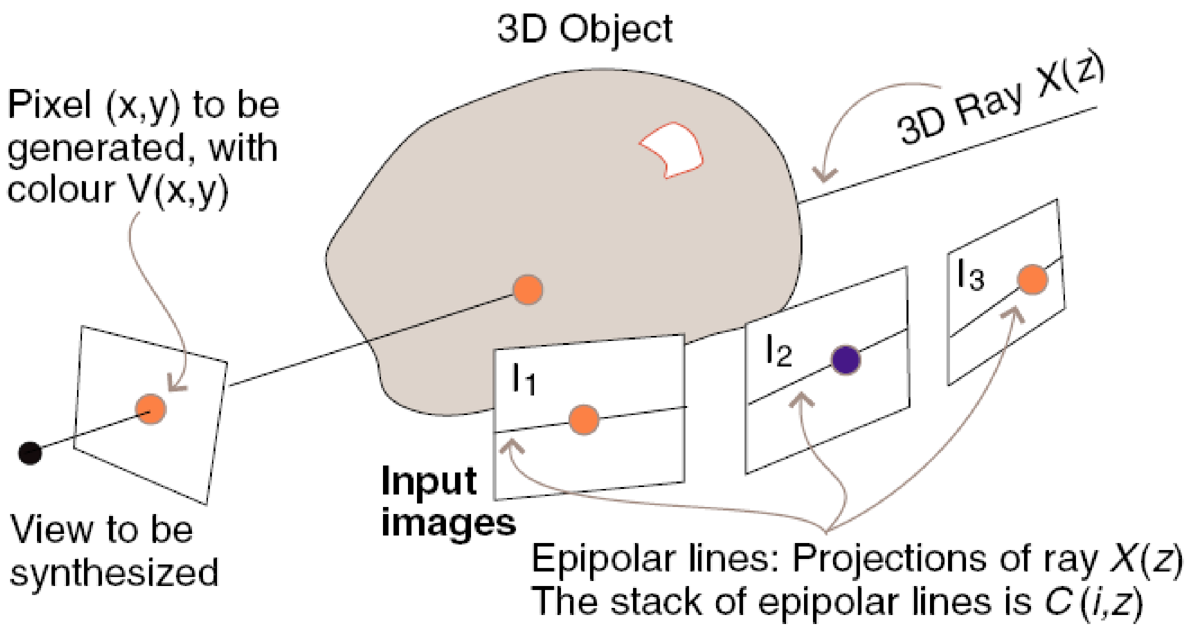
Convex optimization

- Reasons
  - Linearization of constraints
  - L2 norms for data terms
  - Quadratic forms for smoothness
- Problems
  - Large motions not handled
  - Over-smooth motion fields

Optimizing non-convex functions is hard

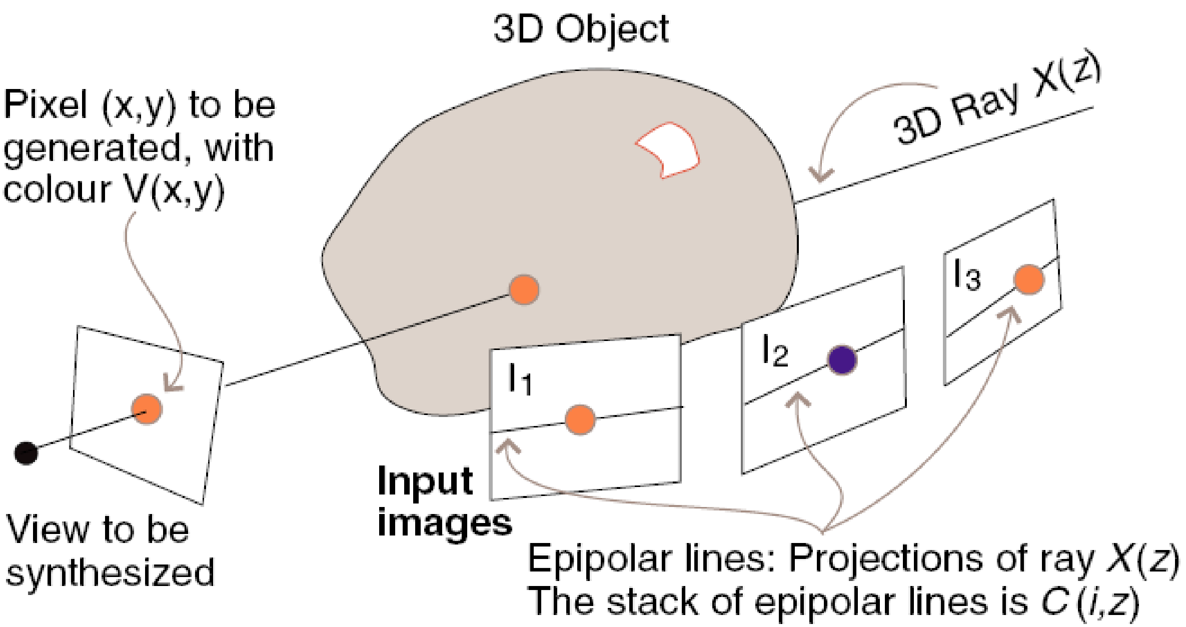
Let’s move away from this a bit

A similar optimization in stereo



Depths for novel view generation

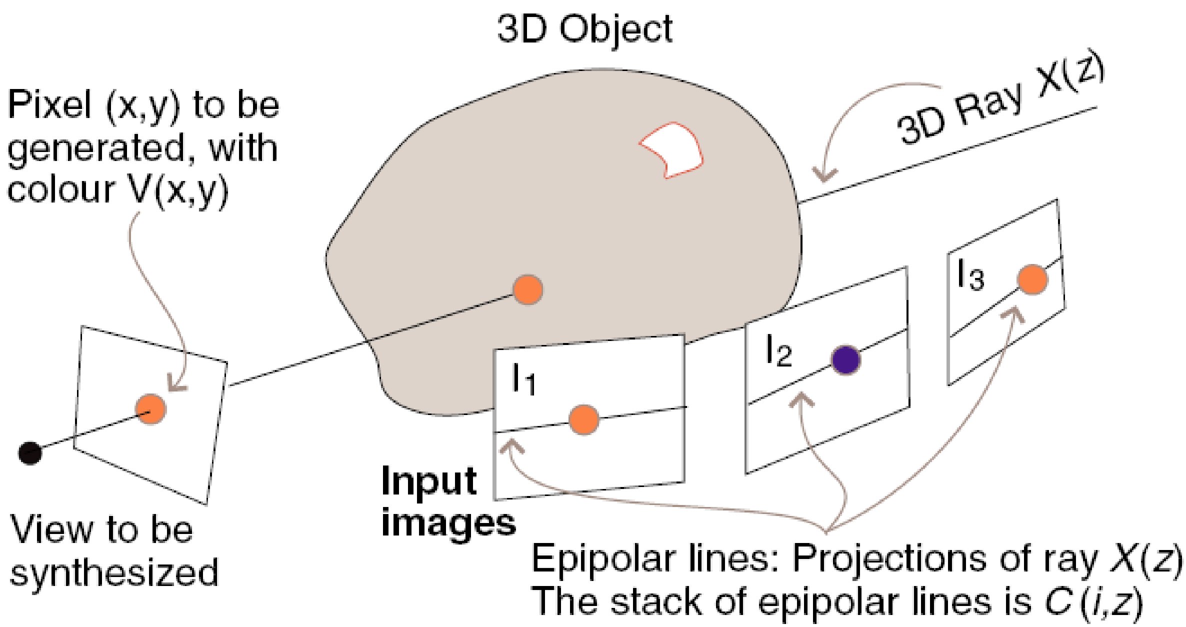
A similar optimization in stereo



<u>Depths for novel view generation</u>

Every pixel in novel view to be assigned a depth and rendered

A similar optimization in stereo



<u>Depths for novel view generation</u>

DISCRETE DEPTH LABELING PROBLEM

Discrete labeling model for optical flow

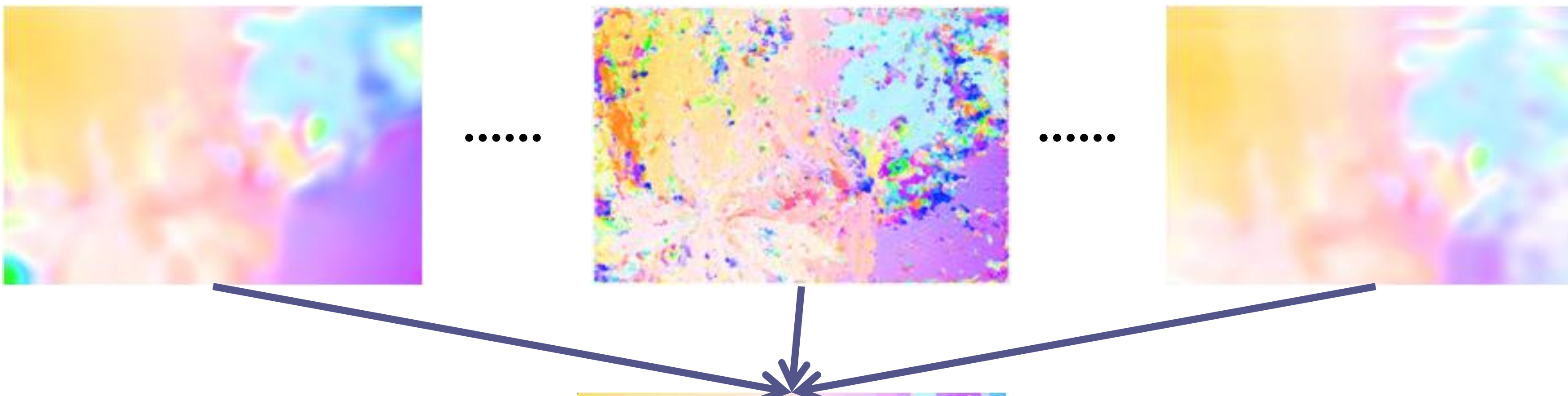
- Each pixel assigned a flow vector
- Problem – too many possible labels
- Can we limit the set of labels?
  - Cues from existing optical flow algorithms
  - **Core idea behind current paper**

Outline

- Problem definition
- Previous work
- System overview
- Evaluation
- Conclusion

Labeling pixels with flow fields

Possible flow fields from existing algorithms

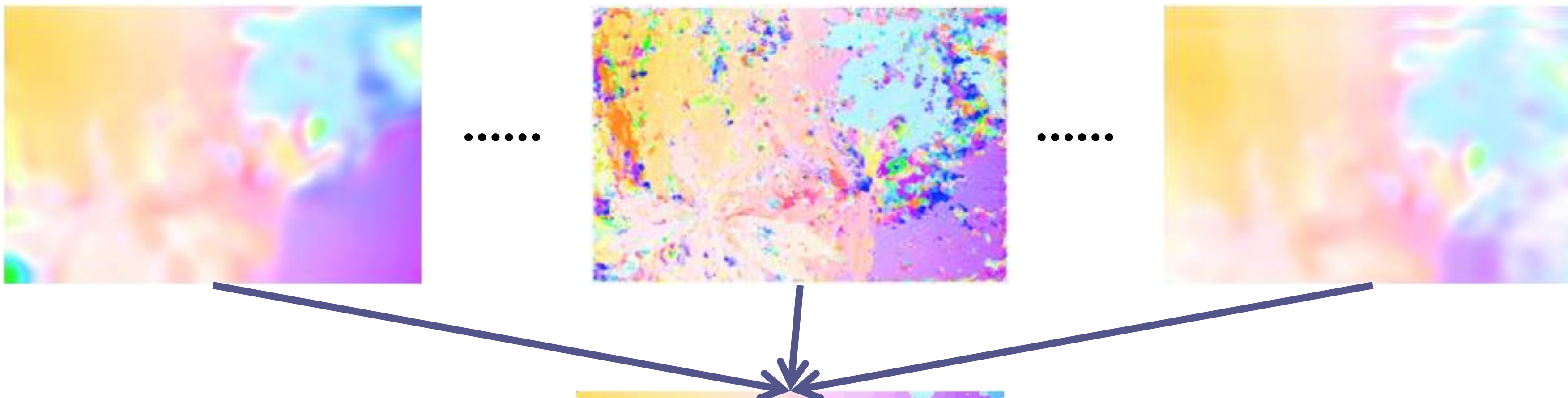


Fused solution

Every pixel chooses flow vector from one flow field which minimizes the overall energy

Labeling pixels with flow fields

Possible flow fields from existing algorithms



Fused solution

Every pixel chooses flow vector from one flow field which minimizes the **overall energy** ?

Objective Energy

$$E(\mathbf{f}) = \sum_{\mathbf{p}\in\Omega} D_{\mathbf{p}}(f_{\mathbf{p}};\ I^0, I^1) + \sum_{(\mathbf{p},\mathbf{q})\in\mathcal{N}} S_{\mathbf{p},\mathbf{q}}(f_{\mathbf{p}}, f_{\mathbf{q}})$$

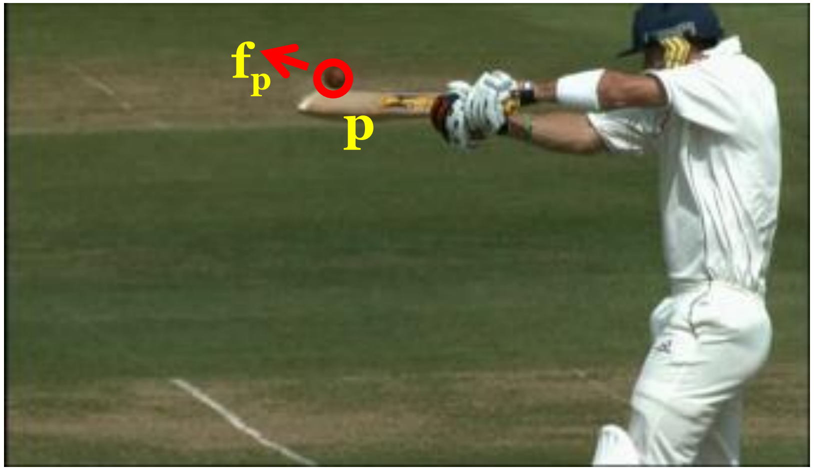
$f_p = (u_p, v_p)$

Objective Energy – Data term

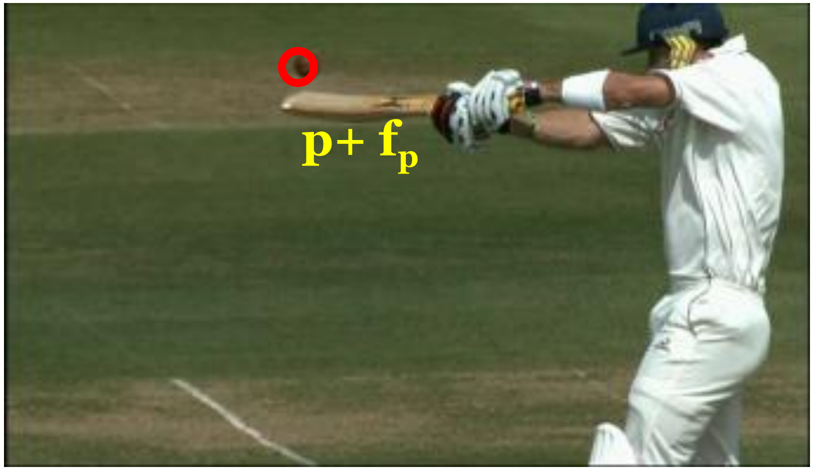
$$E(\mathbf{f}) = \sum_{\mathbf{p}\in\Omega} D_{\mathbf{p}}(f_{\mathbf{p}};\ I^0, I^1) + \sum_{(\mathbf{p},\mathbf{q})\in\mathcal{N}} S_{\mathbf{p},\mathbf{q}}(f_{\mathbf{p}}, f_{\mathbf{q}})$$

$f_p = (u_p, v_p)$

$$D_{\mathbf{p}}(f_{\mathbf{p}};\ I^0, I^1) = \rho_d(||H^1(\mathbf{p} + f_{\mathbf{p}}) - H^0(\mathbf{p})||)$$



I^0



I^1

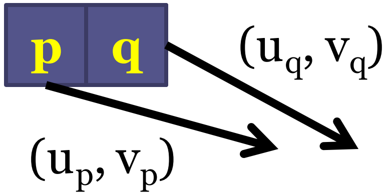
Objective Energy- Regularization

$$E(\mathbf{f}) = \sum_{\mathbf{p}\in\Omega} D_{\mathbf{p}}(f_{\mathbf{p}};\ I^0, I^1) + \sum_{(\mathbf{p},\mathbf{q})\in\mathcal{N}} S_{\mathbf{p},\mathbf{q}}(f_{\mathbf{p}}, f_{\mathbf{q}})$$

$f_p = (u_p, v_p)$

8-connected neighbourhood system

$$S_{\mathbf{p},\mathbf{q}} = \rho_{\mathbf{p},\mathbf{q}}\left(\frac{u_{\mathbf{p}} - u_{\mathbf{q}}}{||\mathbf{p} - \mathbf{q}||}\right) + \rho_{\mathbf{p},\mathbf{q}}\left(\frac{v_{\mathbf{p}} - v_{\mathbf{q}}}{||\mathbf{p} - \mathbf{q}||}\right)$$



- Neighbors have similar flow vectors
- Use of robust functions

Robust functions for smoothness

$$S_{\mathbf{p},\mathbf{q}} = \rho_{\mathbf{p},\mathbf{q}}\left(\frac{u_{\mathbf{p}} - u_{\mathbf{q}}}{||\mathbf{p}-\mathbf{q}||}\right) + \rho_{\mathbf{p},\mathbf{q}}\left(\frac{v_{\mathbf{p}} - v_{\mathbf{q}}}{||\mathbf{p}-\mathbf{q}||}\right)$$

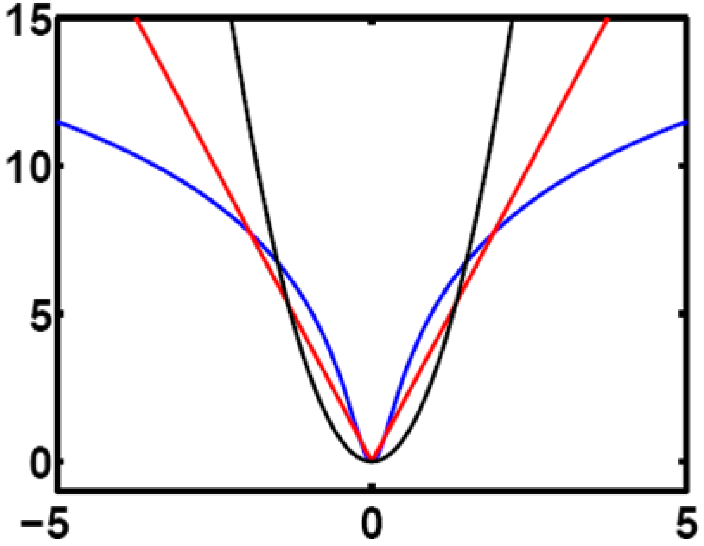
- $\rho_1(x) = x^2$

  [Horn & Schunck]

- $\rho_2(x) = |x|$

- $\rho_3(x) = \lambda_{p,q}\log(1+x^2/2v^2)$

  [Rother et al IJCV 2006]



Energy optimization

$$E(\mathbf{f}) = \sum_{\mathbf{p}\in\Omega} D_{\mathbf{p}}(f_{\mathbf{p}};\ I^0, I^1) + \sum_{(\mathbf{p},\mathbf{q})\in\mathcal{N}} S_{\mathbf{p},\mathbf{q}}(f_{\mathbf{p}}, f_{\mathbf{q}})$$

How do we optimize ?

Energy optimization

$$E(\mathbf{f}) = \sum_{\mathbf{p}\in\Omega} D_{\mathbf{p}}(f_{\mathbf{p}};\ I^0, I^1) + \sum_{(\mathbf{p},\mathbf{q})\in\mathcal{N}} S_{\mathbf{p},\mathbf{q}}(f_{\mathbf{p}}, f_{\mathbf{q}})$$

How do we optimize ?

Step 1: Discrete optimization

- Labeling over candidate flow fields

Step 2: Continuous optimization

- Gradient descent over flow vectors

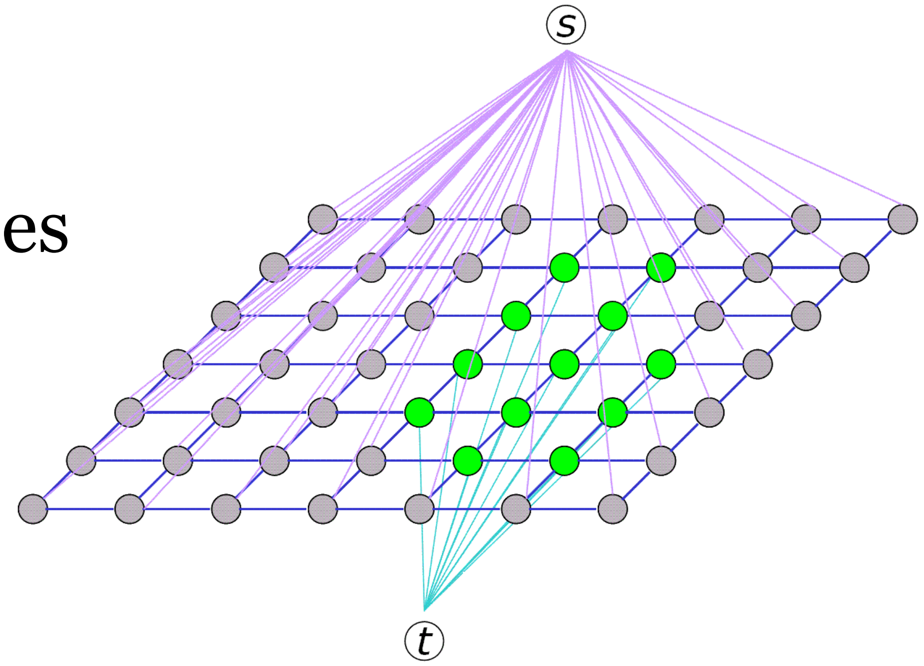
Discrete Optimization Step

- Candidate solutions as labels
  - Horn & Schunck [1981]
  - Lukas Kanade [1981]
  - Varying hierarchy levels and smoothness, shifted copies etc.
  - Constant flow fields from the fused solution
- Multi-label graph-cuts

Graph Cuts

<u>Two-label problem</u>

- Label affinities
- Neighborhood affinities

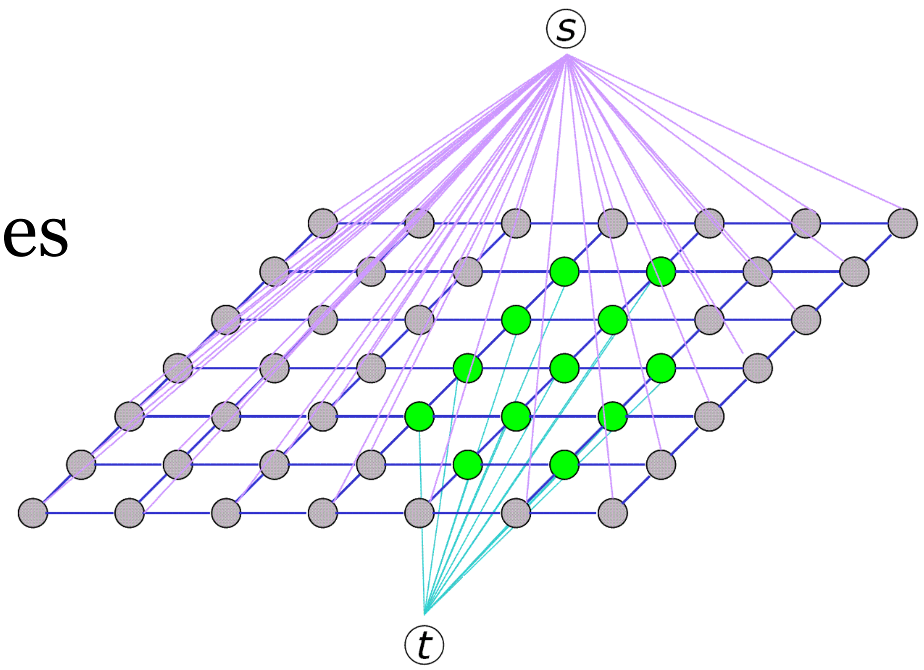


$$E(L) = \sum_{p} E_p(L_p) + \sum_{pq \in N} E(L_p, L_q) \quad L_p \in \{s, t\}$$

Graph Cuts

<u>Two-label problem</u>

- Label affinities
- Neighborhood affinities
- Exactly solvable by max-flow-min-cut algorithm

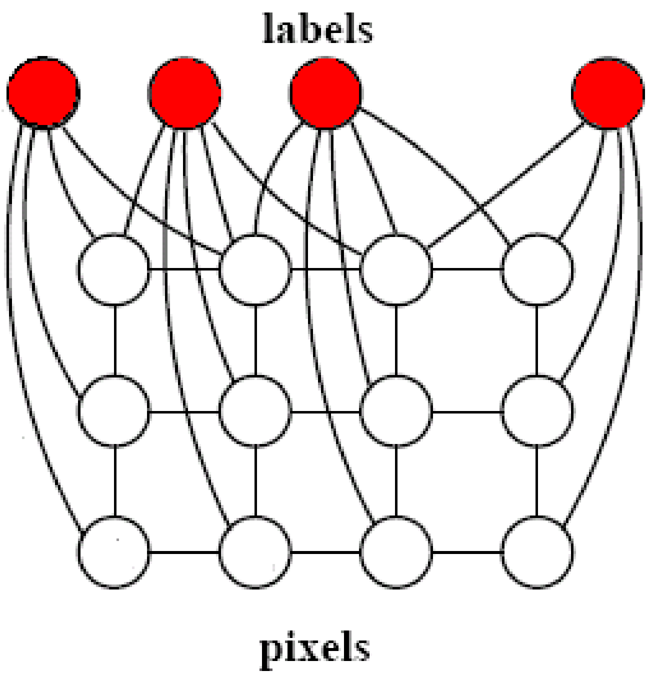


$$E(L) = \sum_{p} E_p(L_p) + \sum_{pq \in N} E(L_p, L_q) \quad L_p \in \{s, t\}$$

Graph Cuts

<u>Multi-label problem</u>

- Many algorithms
  - Belief propagation
    - Local moves
  - Alpha-expansion
    - Global moves



Alpha-expansion algorithm

1. Start with any initial solution
2. For each label "*a*" in any (e.g. random) order

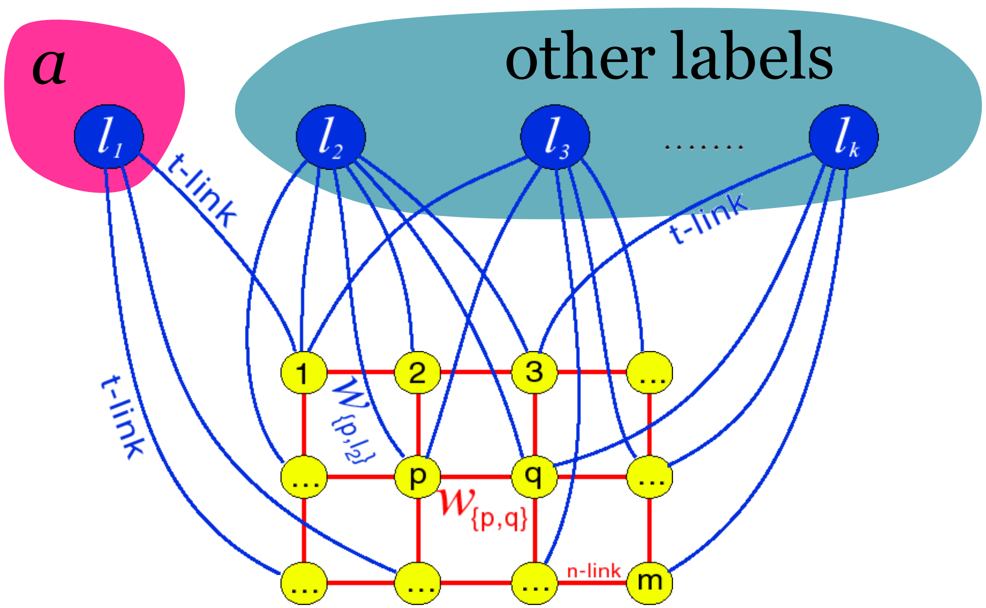
   1. *Compute optimal a-expansion move (s-t graph cuts)*
   2. *Decline the move if there is no energy decrease*

3. *Stop when no expansion move would decrease energy*

Taken from Yuri Boykov's ICCV 2007 tutorial

Alpha-expansion move

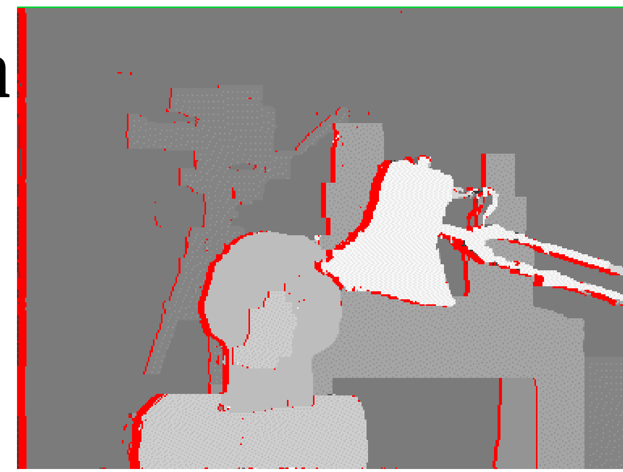
Basic idea: break multi-way cut computation into a **sequence of binary *s-t* cuts**



Taken from Yuri Boykov's ICCV 2007 tutorial

Multi-label graph cuts

stereo vision



depth map ??

original pair of “stereo” images

Taken from Yuri Boykov’s ICCV 2007 tutorial

Alpha-expansion moves

initial solution

-expansion

-expansion

-expansion

-expansion

-expansion

-expansion

-expansion

For each move we choose expansion that gives the largest decrease in the energy: **binary optimization problem**

Taken from Yuri Boykov's ICCV 2007 tutorial

Multi-label graph cuts for optical flow

$$E(\mathbf{f}) = \sum_{\mathbf{p}\in\Omega} D_{\mathbf{p}}(f_{\mathbf{p}};\ I^0, I^1) + \sum_{(\mathbf{p},\mathbf{q})\in\mathcal{N}} S_{\mathbf{p},\mathbf{q}}(f_{\mathbf{p}}, f_{\mathbf{q}})$$

TOO MANY

- Labels ⇔ Algorithms instead of flow vectors
- Energy term ⇔ Flow vectors from algorithms

Multi-label graph cuts for optical flow

$$E(\mathbf{f}) = \sum_{\mathbf{p}\in\Omega} D_{\mathbf{p}}(f_{\mathbf{p}};\ I^0, I^1) + \sum_{(\mathbf{p},\mathbf{q})\in\mathcal{N}} S_{\mathbf{p},\mathbf{q}}(f_{\mathbf{p}}, f_{\mathbf{q}})$$

TOO MANY

- Labels ⇔ Algorithms instead of flow vectors
- Energy term ⇔ Flow vectors from algorithms
- Essentially fusing fields together
- Alpha expansion - Expand a **flow field label**

Multi-label graph cuts for optical flow

$$E(\mathbf{f}) = \sum_{\mathbf{p}\in\Omega} D_{\mathbf{p}}(f_{\mathbf{p}};\ I^0, I^1) + \sum_{(\mathbf{p},\mathbf{q})\in\mathcal{N}} S_{\mathbf{p},\mathbf{q}}(f_{\mathbf{p}}, f_{\mathbf{q}})$$

<u>Fusion move</u> [Lempitsky et al, ICCV 2007]

- Expand a flow field label (fusion)
- Problem - Non-submodular energy

Multi-label graph cuts for optical flow

Submodularity condition

- L and M be two labels assigned to neighbors p and q
- $E_{p,q}(L,L) + E_{p,q}(M,M) <= E_{p,q}(L,M) + E_{p,q}(M,L)$
- Cannot be guaranteed to hold true when L and M are flow fields

Multi-label graph cuts for optical flow

$$E(\mathbf{f}) = \sum_{\mathbf{p}\in\Omega} D_{\mathbf{p}}(f_{\mathbf{p}};\ I^0, I^1) + \sum_{(\mathbf{p},\mathbf{q})\in\mathcal{N}} S_{\mathbf{p},\mathbf{q}}(f_{\mathbf{p}}, f_{\mathbf{q}})$$

<u>Fusion move</u> [Lempitsky et al, ICCV 2007]

- Expand a flow field label (fusion)
- Non-submodular energy ➔ Alpha-expansion not possible
- QPBO (Quadratic Pseudo-Boolean Optimization) instead of graph cuts [Boros&Hummer, 2002]

Discrete Optimization Step

Input images

Discrete Optimization Step

Initial (energy=30288)

Final (energy = 7483)

Energy optimization

$$E(\mathbf{f}) = \sum_{\mathbf{p}\in\Omega} D_{\mathbf{p}}(f_{\mathbf{p}};\ I^0, I^1) + \sum_{(\mathbf{p},\mathbf{q})\in\mathcal{N}} S_{\mathbf{p},\mathbf{q}}(f_{\mathbf{p}}, f_{\mathbf{q}})$$

How do we optimize ?

Step 1: Discrete optimization

- Labeling over candidate flow fields

Step 2: Continuous optimization

- Gradient descent over flow vectors

Continuous Optimization

- Why another step
  - Good candidates not available in some regions

- Same energy function

$$E(\mathbf{f}) = \sum_{\mathbf{p}\in\Omega} D_{\mathbf{p}}(f_{\mathbf{p}};\ I^0, I^1) + \sum_{(\mathbf{p},\mathbf{q})\in\mathcal{N}} S_{\mathbf{p},\mathbf{q}}(f_{\mathbf{p}}, f_{\mathbf{q}})$$

- Use of conjugate gradients

Continuous Optimization

After discrete step
(energy = 7483)

Finally
(energy = 5788)

Outline

- Problem definition
- Previous work
- System overview
- Evaluation
- Conclusion

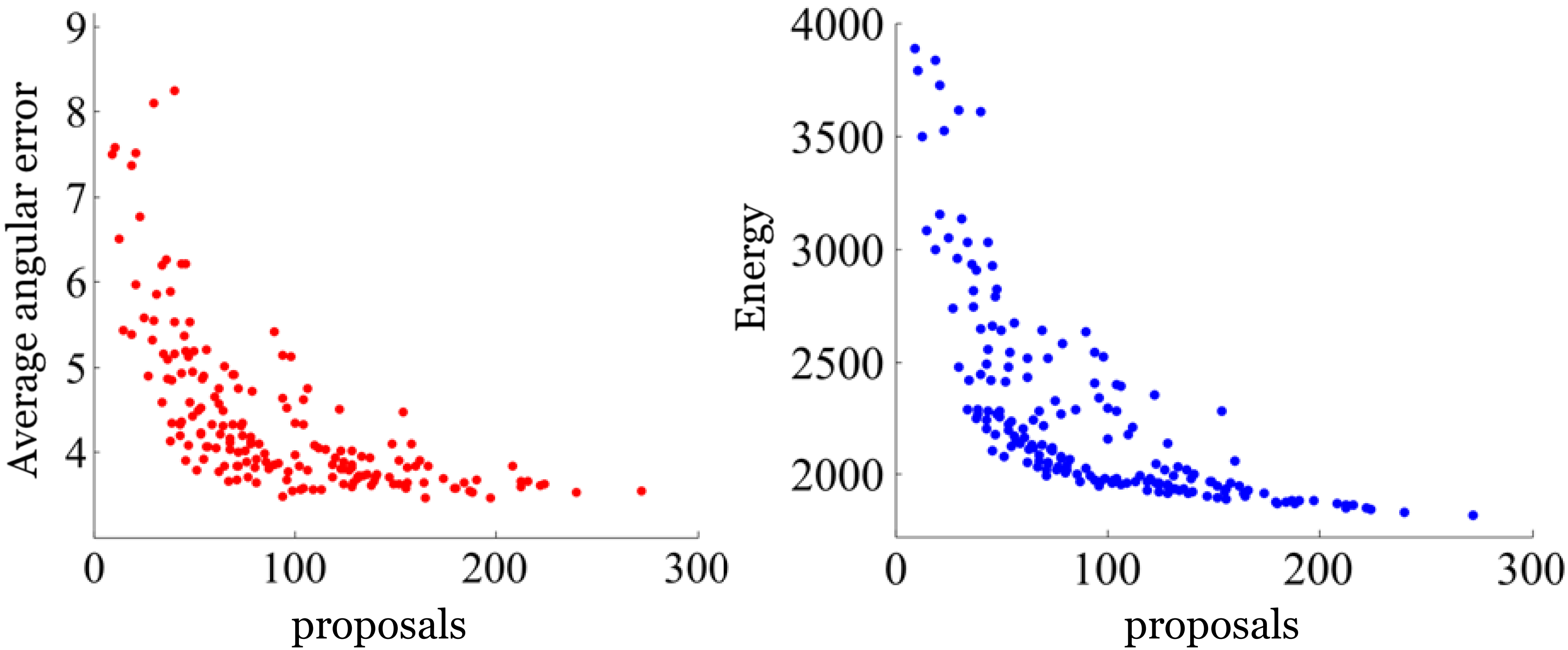
Evaluation

- Talk about Middlebury dataset [Baker et al ICCV 2007]

  (switch to web page)

Evaluation

Evaluation – number of proposals



Conclusion

- Discrete labeling to prevent local minima
  - Followed by continuous optimization
- Use of optical flow statistics
- Spatially varying smoothness weight

- Slow (speed not mentioned in paper)
- What is the limit to improvement?

Thank you

Aperture problem