3D Photography: Stereo
Marc Pollefeys, Torsten Sattler
Spring 2015
http://cvg.ethz.ch/teaching/3dphoto/
Dense Correspondence / Stereo

Tsukuba dataset

http://cat.middlebury.edu/stereo/
Stereo

- Standard stereo geometry
- Stereo matching
  - Correlation
  - Optimization (DP, GC)
- General camera configuration
  - Rectifications
- PatchMatch Stereo
Stereo
Stereo camera configurations

Short baseline:
- Good matches
- Few occlusions
- Poor precision

Long baseline:
- Harder to match
- More occlusions
- Better precision

(Slide from Pascal Fua)
Occlusions

\[ \rightarrow \text{Consistency test} \]
Exploiting scene constraints
Ordering constraint

surface slice

surface as a path

occlusion left

occlusion right
Uniqueness constraint

- In an image pair each pixel has at most one corresponding pixel
  - In general one corresponding pixel
  - In case of occlusion there is none
Disparity constraint

use reconstructed features to determine bounding box
Stereo matching

- **Similarity measure** (SSD or NCC)
- **Optimal path** (dynamic programming)

**Constraints**
- epipolar
- ordering
- uniqueness
- disparity limit

**Trade-off**
- Matching cost (data)
- Discontinuities (prior)

Consider all paths that satisfy the constraints
pick best using dynamic programming
Hierarchical stereo matching

Allows faster computation
Deals with large disparity ranges
Disparity map

image $I(x,y)$

Disparity map $D(x,y)$

image $I'(x',y')$

$(x',y') = (x + D(x,y), y)$
Energy minimization

Disparity continuous in most places,

except at depth discontinuities

1. Matching pixels should have similar intensities.
2. Most nearby pixels should have similar disparities

\[ \text{Minimize} \quad \sum [I_1(x + D(x, y), y) - I_2(x, y)]^2 \]
\[ + \lambda \sum [D(x + 1, y) - D(x, y)]^2 \]
\[ + \mu \sum [D(x, y + 1) - D(x, y)]^2 \]
Graph Cut

1. Stereo is a labeling problem

2. Graph cut corresponds to a labeling.

→ **Assign edge weights cleverly so that the min-weight cut gives the minimum energy!**

(general formulation requires multi-way cut!)

(Slide from Pascal Fua)
Simplified graph cut

\[ V = V^* \cup \{s, t\} \]
\[ E = E^* \cup \{(s, v) : v \in \text{Front}\} \cup \{(u, t) : u \in \text{Back}\} \]

(Roy and Cox ICCV‘98)

(a) initial labeling
(b) standard move
(c) \( \alpha-\beta \)-swap
(d) \( \alpha \)-expansion

(Boykov et al ICCV‘99)
True disparities

11 – GC + occlusions

*2 – Dynamic progr.

16 – Fast Correlation
Semi-global optimization

- Optimize: \( E = E_{\text{data}} + E(|D_p - D_q| = 1) + E(|D_p - D_q| > 1) \) [Hirshmüller CVPR05]
  - Use mutual information as cost
- NP-hard using graph cuts or belief propagation (2-D optimization)
- Instead do dynamic programming along many directions
  - Don’t use visibility or ordering constraints
  - Enforce uniqueness
  - Add costs
Stereo matching with general camera configuration

[Images of two stereo matching photographs with yellow lines indicating matching points]
Image pair rectification
Planar rectification

Bring two views to standard stereo setup
(moves epipole to \( \infty \))
(not possible when in/close to image)

\[
H'^{-\top}FH^{-1} = \begin{bmatrix}
0 & 0 & 0 \\
0 & 0 & 1 \\
0 & -1 & 0
\end{bmatrix}
\]

~ image size
(calibrated)

Distortion minimization
(uncalibrated)
Polar rectification
(Pollefeys et al. ICCV’99)

Polar re-parameterization around epipoles
Requires only (oriented) epipolar geometry
Preserve length of epipolar lines
Choose $\Delta \theta$ so that no pixels are compressed

Works for all relative motions
Guarantees minimal image size
original image pair

planar rectification

polar rectification
Example: Béguinage of Leuven

Does not work with standard Homography-based approaches
Plane-sweep multi-view matching

- Simple algorithm for multiple cameras
- no rectification necessary
- doesn’t deal with occlusions

Collins’ 96; Roy and Cox’ 98 (GC); Yang et al.’ 02/’ 03 (GPU)
Fast GPU-based plane-sweeping stereo

Plane-sweep multi-view depth estimation

(Yang & Pollefeys, CVPR’03)
Slanted Support Windows

fronto-parallel windows vs. slanted support windows
PatchMatch Stereo
(Bleyer et al. BMVC’11)

For a particular plane the disparity at a pixel is given by

\[ d_p = a f_p p_x + b f_p p_y + c f_p \]

The plane with the minimal cost is chosen

\[ f_p = \arg\min_{f \in \mathcal{F}} m(p, f) \]

The dissimilarity cost is calculated as

\[ m(p, f) = \sum_{q \in \mathcal{W}_p} w(p, q) \cdot \rho(q, q - (a f q_x + b f q_y + c f)) \]

with

\[ w(p, q) = e^{-\frac{||I_p - I_q||}{\gamma}} \]

\[ \rho(q, q') = (1 - \alpha) \cdot \min(||I_q - I_{q'}||, \tau_{col}) + \alpha \cdot \min(||\nabla I_q - \nabla I_{q'}||, \tau_{grad}) \]
PatchMatch Stereo
(Bleyer et al. BMVC’11)

Idea: Start with a random initialization of disparities and plane parameters for each pixel and update the estimates by propagating information from the neighboring pixels

- **Spatial propagation**: Check for each pixel the disparities and plane parameters for the left and upper (right and lower) neighbors and replace the current estimates if matching costs are smaller.
- **View propagation**: Warp the point in the other view and check the corresponding estimates in the other image. Replace if the matching costs are lower.
- **Temporal propagation**: Propagate the information analogously by considering the estimates for the same pixel at the preceding and consecutive video frame.
PatchMatch Stereo  
(Bleyer et al. BMVC’11)

- Plane refinement: disparity and plane parameters for each pixel are refined by generating random samples within a certain range interval and updating the current estimates if matching costs are reduced.

- Post-processing: remove outliers with left/right consistency checking and weighted median filter. Gaps are filled by propagating information from the neighborhood.
PatchMatch Stereo
(Bleyer et al. BMVC’11)
Paper Presentation

Amrollah Seifoddini, Matthew Krenik Saurer et al., **Rolling Shutter Stereo.** ICCV 2013
*opponent: Himanshu Jain, Christos Sakaridis*

David Gonon, Yang Shuoran Weinzaepfel et al., **DeepFlow: Large displacement optical flow with deep matching.** ICCV 2013
*opponent: Radek Danecek, Alex Lelidis*

Marco Karrer, Patrick Stahli Richard et al., **Megastereo: Constructing High-Resolution Stereo Panoramas.** CVPR 2013
*opponent: Wolf Vollprecht, Anurag Sai Vempati*