3D Vision: Stereo

Marc Pollefeys, Torsten Sattler

Spring 2016

http://cvg.ethz.ch/teaching/3dvision/
Dense Correspondence / Stereo

Tsukuba dataset

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

- Standard stereo geometry
- Stereo matching
  - Correlation
  - Optimization (DP, GC)
- General camera configurations
  - Rectification
- PatchMatch Stereo
Stereo
Disparity map

image $I(x,y)$

Disparity map $D(x,y)$

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

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

→ Consistency test

(Slide from Pascal Fua)
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

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

**Similarity measure**
(SSD or NCC)

**Optimal path**
(dynamic programming)

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

**Trade-off**
- Matching cost (data)
- Discontinuities (prior)
Hierarchical stereo matching

Allows faster computation
Deals with large disparity ranges

Downsampling (Gaussian pyramid)
True disparities

*2 – Dynamic progr.

16 – Fast Correlation

(Scharstein & Szeliski, IJCV‘02)
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

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

(Slide from Pascal Fua)
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

(Scharstein & Szeliski, IJCV‘02)
Semi-global optimization

- Optimize:
  \[ E = E_{\text{data}} + E(|D_p - D_q| = 1) + E(|D_p - D_q| > 1) \]
  [Hirschmüller 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 of all paths
Stereo matching with general camera configurations
Image pair rectification
Planar rectification

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

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

\(~\text{image size}~\)
(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
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

(Bleyer et al. BMVC’11)
PatchMatch Stereo
(Bleyer et al. BMVC’ 11)

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

\[ d_p = a_f p_x + b_f p_y + c_f \]

- 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

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

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

Left to right:
- Fronto-parallel, discrete disparities
- Fronto-parallel, continuous disparities
- PatchMatch Stereo (slanted, continuous disparities)
Next week:
Bundle Adjustment & SLAM