Depth Map Fusion with Camera Position Refinement

Computer Vision Winter Workshop 2009

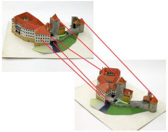
Radim Tyleček and Radim Šára
tylecr1@cmp.felk.cvut.cz

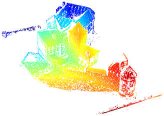
Center for Machine Perception
Department of Cybernetics, Faculty of Electrical Engineering,
Czech Technical University Prague, Czech Republic

Contents

3D Reconstruction Pipeline

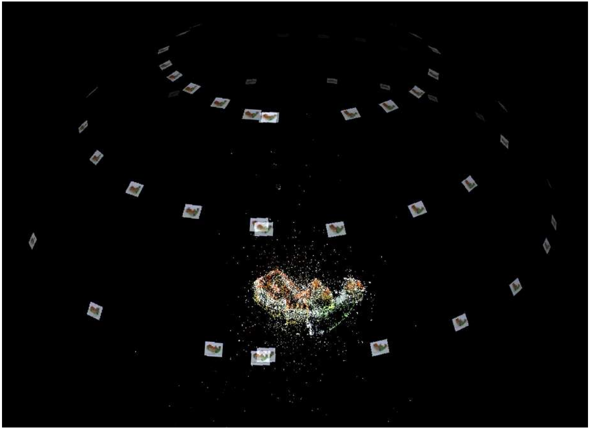
Input images ⇒ Corresponding regions ⇒ Disparity maps ⇒ Point cloud ⇒ **Surface mesh**





- Pair-wise vs. multi-view stereo
- Calibration inaccuracy
- Inconsistency between disparity maps

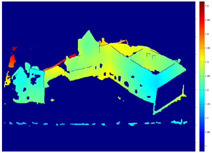
Input 3D point cloud



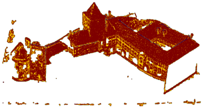
- Noise, outliers in the data
- Redundancy

Surface representation

- Depth maps

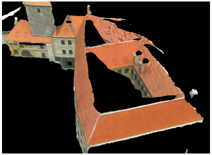


- Visibility maps

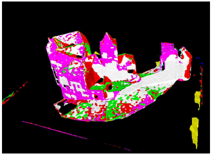


⇒ Back-projection

- Reconstructed surface (scan)



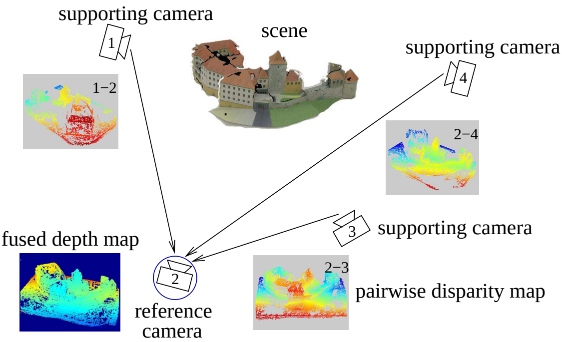
- Surface composition



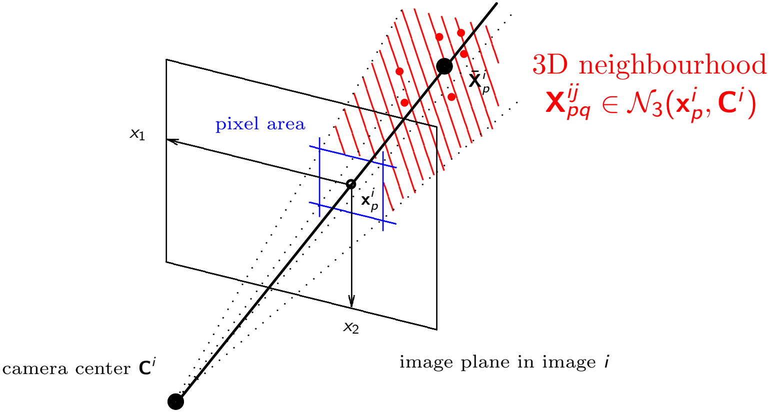
- Linear complexity

Idea of Depth Map Fusion

- Representation with a set of reference cameras

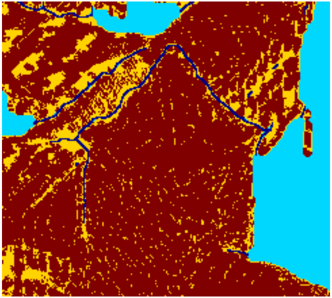


Idea of Depth Map Fusion



Visibility estimation

Visibility map

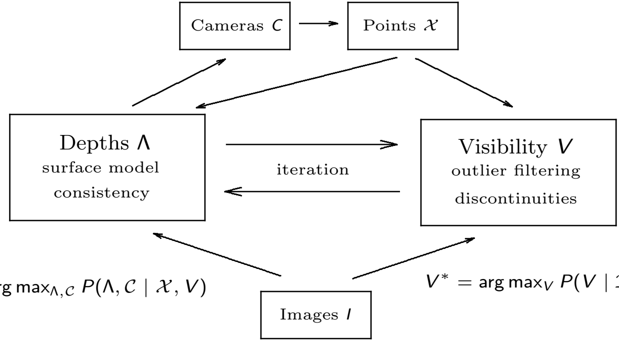


- Visibility labels
  - $v = 0$ surface not visible
  - $v = 1$ visible, no data
    $\rightarrow$ interpolation
  - $v = 2$ visible, data present
- Discontinuity – line of pixels with $v = 0$

Algorithm design

Global Structure and Motion optimisation problem

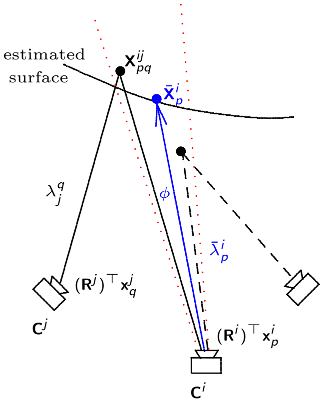
$$(\mathcal{X}^*, \Lambda^*, V^*, \mathcal{C}^*) = \arg\max_{\mathcal{X},\Lambda,V,\mathcal{C}} P(\mathcal{X}, \Lambda, V, \mathcal{C} \mid \mathcal{I})$$



$$(\Lambda^*, \mathcal{C}^*) = \arg\max_{\Lambda,\mathcal{C}} P(\Lambda, \mathcal{C} \mid \mathcal{X}, V)$$

$$V^* = \arg\max_V P(V \mid \mathcal{I}, \Lambda, \mathcal{X})$$

Depth estimation



Geometric constraints

$$\bar{\lambda}_p^i = \arg\min_{\bar{\lambda}} \sum_{(j,q)} \|\bar{\mathbf{X}}_p^i - \mathbf{X}_{pq}^{ij}\|^2 \qquad (1)$$

$$\bar{\mathbf{X}}_p^i = \Phi(\bar{\lambda}_p^i) \ldots \text{backprojection} \qquad (2)$$

Depth estimation

Depth task:

geometric constraints $\longrightarrow$ system of linear equations:

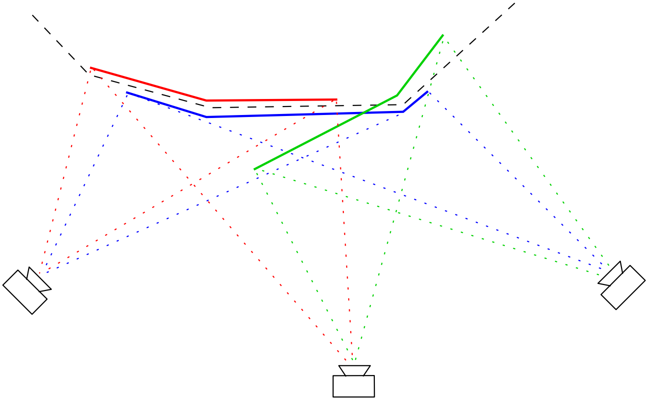
$$\mathbf{R}^{j(3)}\mathbf{C}^{i} + \mathbf{R}^{j(3)}\,\mathbf{R}^{i\top}\mathbf{K}^{i-1}\mathbf{x}_{p}^{i}\,\bar{\lambda}_{p}^{i} - \lambda_{q}^{j} = \mathbf{R}^{j(3)}\mathbf{C}^{j} \tag{3}$$

surface model $\longrightarrow$ system of linear equations:

$$\frac{1}{\sigma_{\lambda}^{2}}\,(\bar{\lambda}_{p}^{i} - \lambda_{p}^{i}) + \sum_{\bar{p}\in N_{p}} \frac{1}{(\sigma_{c,\bar{p}}^{\;i})^{2}}\,(\lambda_{p}^{i} - \lambda_{\bar{p}}^{i}) = 0 \tag{4}$$

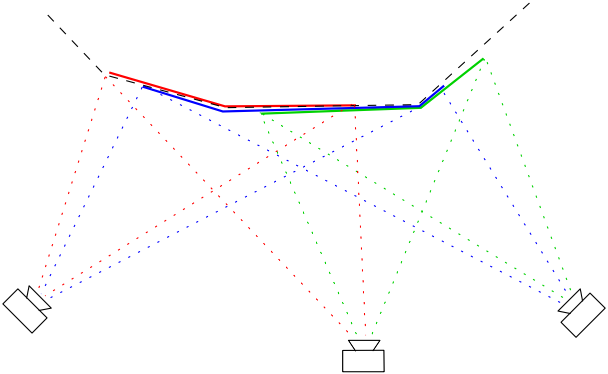
Camera position refinement

Original camera position



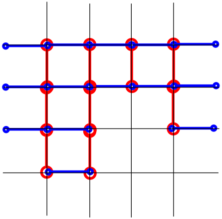
Camera position refinement

Refined camera position



Visibility estimation

Visibility task: optimal labelling → minimum graph cut:



$$E(V^i) = \sum_{p=1}^{n} E(v_p^i) + \frac{1}{2\sigma_v^2} \sum_{(p,\bar{p})\in\mathcal{N}_2(i)} (v_p^i - v_{\bar{p}}^i)^2$$

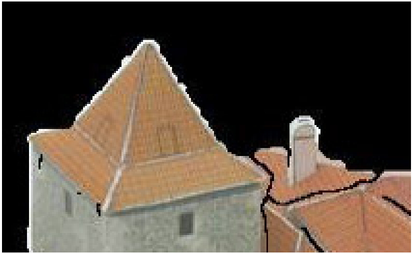
$$E(v_p^i) = \sum_{(q,j)\in\chi_p^i;\; v_q^j\geq 1} E(v_p^i, v_q^j) + \sum_{(p,\bar{p})\in\mathcal{N}_2(i|V)} \frac{(\lambda_p^i - \lambda_{\bar{p}}^i)^2}{2(\sigma_{\lambda,p}^{\;i})^2}$$

$$E(v_p^i, v_q^j) = \begin{cases} \frac{(\mathbf{I}_p^i - \mathbf{I}_q^j)^2}{2\sigma_I^2} & \text{pro } v_p^i = v_q^j = 2 \\ -\log h(\mathbf{I}_p^i) & \text{otherwise.} \end{cases}$$

Visibility estimation

Visibility map

Initial



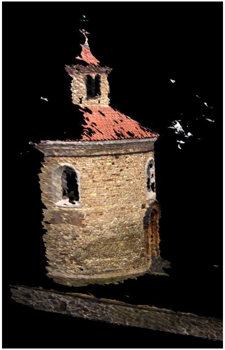
After first iteration

Experiments

- High accuracy comparable with state-of the art methods
- Error suppression (outlier removal, smoothing out noise)
- Camera calibration refinement

Experiments

- Different objects and scenes



London dataset

Castle dataset

Daliborka dataset

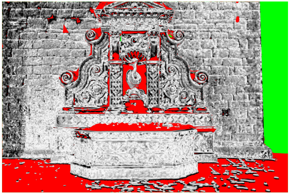
Fountain dataset

Evaluation

Image

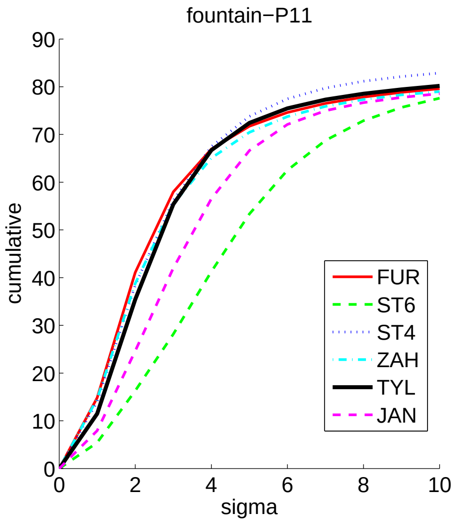
Result rendering

Ground truth



Depth error

Evaluation



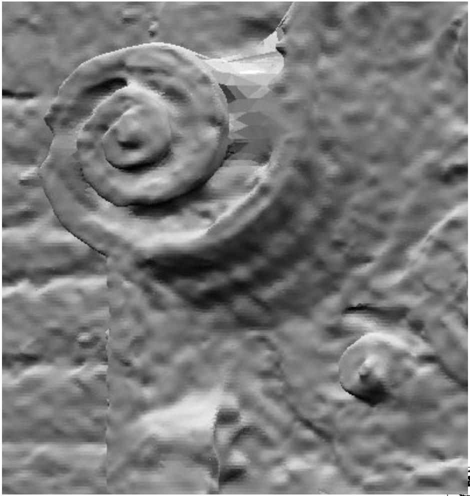
- Ground truth
- Surface projected to cameras
- Depth measurement error σ

Future work

- High accuracy images available
- Photometric mesh refinement
- Second-order surface model

Future work

- High accuracy images available
- Photometric mesh refinement
- Second-order surface model



Future work

- High accuracy images available
- Photometric mesh refinement
- Second-order surface model

Summary

- Surface reconstruction with Depth Map Fusion
- Camera calibration refinement
- To do: Photometric mesh refinement

Thank you.