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

1 Introduction
   • Reconstruction Pipeline

2 Method
   • Idea
   • Algorithm design
   • Depth task
   • Visibility task

3 Experiments
   • Results
   • Evaluation
   • Future work

4 Conclusion
3D Reconstruction Pipeline

Input images $\Rightarrow$ Corresponding regions $\Rightarrow$ Disparity maps $\Rightarrow$ Point cloud $\Rightarrow$ 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

Radim Tyleček and Radim Šára, CMP Prague
Idea of Depth Map Fusion

- Representation with a set of reference cameras
Idea of Depth Map Fusion

3D neighbourhood \( \mathbf{X}_{pq}^j \in \mathcal{N}_3(\mathbf{x}_p^i, \mathbf{C}^i) \)

camera center \( \mathbf{C}^i \)

image plane in image \( i \)

\[
\bar{x}_i \\
\mathbf{x}_p \in \mathcal{N}_3(\mathbf{x}_p^i, \mathbf{C}^i)
\]

pixel area

radim tyleček and radim šára, cmp prague

depth map fusion, cvww 2009
Visibility estimation

Visibility map

- Visibility labels
  - $v = 0$ surface not visible
  - $v = 1$ visible, no data → interpolation
  - $v = 2$ visible, data present
- Discontinuity – line of pixels with $v = 0$
Global Structure and Motion optimisation problem

\[(\mathcal{X}^*, \Lambda^*, V^*, C^*) = \underset{\mathcal{X}, \Lambda, V, C}{\arg \max} P(\mathcal{X}, \Lambda, V, C | I)\]

Cameras \(C\) \rightarrow Points \(\mathcal{X}\)

Depths \(\Lambda\)
- surface model
- consistency

\[(\Lambda^*, C^*) = \underset{\Lambda, C}{\arg \max} P(\Lambda, C | \mathcal{X}, V)\]

Images \(I\)

Visibility \(V\)
- outlier filtering
- discontinuities

\[V^* = \underset{V}{\arg \max} P(V | I, \Lambda, \mathcal{X})\]
Depth estimation

Geometric constraints

\[ \bar{\lambda}_p = \arg \min_{\bar{\lambda}} \sum_{(j,q)} \| \bar{X}_p - \bar{X}_{pq} \|^2 \]  

(1)

\[ \bar{X}_p = \Phi(\bar{\lambda}_p) \ldots \text{backprojection} \]  

(2)
Depth estimation

Depth task:
geometric constraints → system of linear equations:

\[ R^{j(3)} C^i + R^{j(3)} R^{i\top} K^{i-1} x_p \bar{\lambda}_p - \lambda_q = R^{j(3)} C^j \]  \hspace{1cm} (3)

surface model → system of linear equations:

\[ \frac{1}{\sigma^2_\lambda} (\bar{\lambda}_p - \lambda^i_p) + \sum_{\bar{p} \in N_p} \frac{1}{(\sigma_{c,\bar{p}})^2} (\lambda^i_p - \lambda^i_{\bar{p}}) = 0 \]  \hspace{1cm} (4)
Camera position refinement

Original camera position
Camera position refinement

Refined camera position
Visibility estimation

Visibility task: optimal labelling $\rightarrow$ minimum graph cut:

$$E(V^i) = \sum_{p=1}^{n} E(v^i_p) + \frac{1}{2\sigma_v^2} \sum_{(p,\bar{p}) \in N_2(i)} (v^i_p - v^i_{\bar{p}})^2$$

$$E(v^i_p) = \sum_{(q,j) \in \chi^i_p; v^j_q \geq 1} E(v^i_p, v^j_q) + \sum_{(p,\bar{p}) \in N_2(i|V)} \frac{(\lambda^i_p - \lambda^i_{\bar{p}})^2}{2(\sigma^i_{\lambda,p})^2}$$

$$E(v^i_p, v^j_q) = \begin{cases} 
\frac{(I^i_p - I^j_q)^2}{2\sigma_i^2} & \text{pro } v^i_p = v^j_q = 2 \\
- \log h(I^i_p) & \text{otherwise.}
\end{cases}$$
Visibility estimation

Visibility map

Initial

After first iteration

Radim Tyleček and Radim Šára, CMP Prague

Depth Map Fusion, CVWW 2009
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

![Different objects and scenes](image1.jpg) ![Different objects and scenes](image2.jpg) ![Different objects and scenes](image3.jpg)
London dataset
Castle dataset
Daliborka dataset

[Series of images showing different views of a 3D model of a building]

Introduction
Method
Experiments
Conclusion

Results
Evaluation
Future work

Radim Tyleček and Radim Šára, CMP Prague
Depth Map Fusion, CVWW 2009
Fountain dataset
Evaluation

Image

Result rendering

Ground truth

Depth error
Evaluation

fountain−P11

Ground truth
Surface projected to cameras
Depth measurement error $\sigma$

### Graph Details
- **X-axis**: \( \sigma \)
- **Y-axis**: cumulative
- **Legend**:
  - FUR
  - ST6
  - ST4
  - ZAH
  - TYL
  - JAN

---

*Radim Tyleček and Radim Šára, CMP Prague*  
*Depth Map Fusion, CVWW 2009*
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.