Refinement of Surface Mesh for Accurate Multi-View Reconstruction

International Workshop on Representation and Modeling of Large-scale 3D Environments
Asian Conference on Computer Vision
Xian, China, September 2009

Radim Tyleček, Radim Šára
{tylecr1|sara}@cmp.felk.cvut.cz

Center for Machine Perception, Czech Technical University, Prague

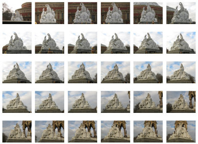
Motivation

- High resolution images available
- State-of-the-art MVS results still below accuracy of laser scanners
- Goal: elimination of sources of inaccuracy
  - imprecise camera calibration
  - variable capture conditions
  - suboptimal representation

3D Photography

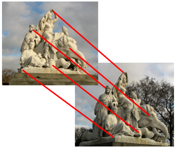
Surface Reconstruction Pipeline

Input images



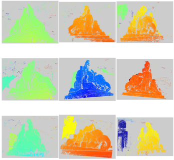
$\Rightarrow$

Corresponding regions

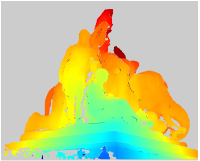


$\Rightarrow$

Pair-wise disparity maps



$\Downarrow$

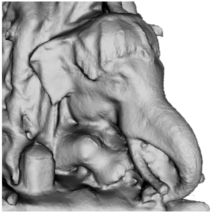


Fused depth maps

$\Leftarrow$

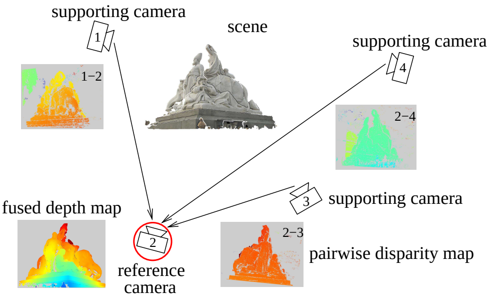
Surface mesh

$\Leftarrow$



Refined mesh

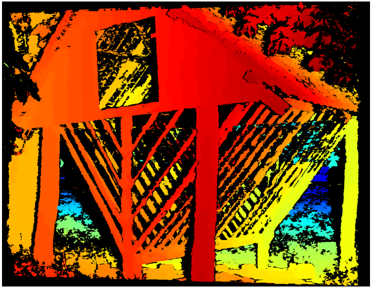
Depth Map Fusion [Tyl09]



- Image-based representation with a set of reference cameras
- Global problem of joint estimation of depths and cameras

[Tyl09] R.Tylecek, R.Sara: Depth Map Fusion with Camera Calibration Refinement, CVWW 2009

Why Pair-wise Stereo?

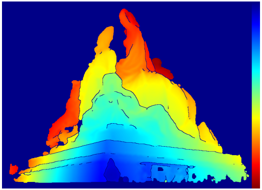


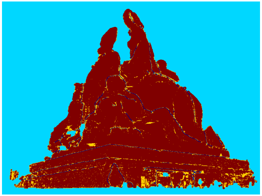
- Mature methods developed and available [Cech07]
- Less vulnerable to calibration errors than traditional MVS

[Cech07] J.Cech, R.Sara: Efficient sampling of disparity space for fast and accurate matching. BenCOS CVPR 2007

Depth Map Representation

Depth maps

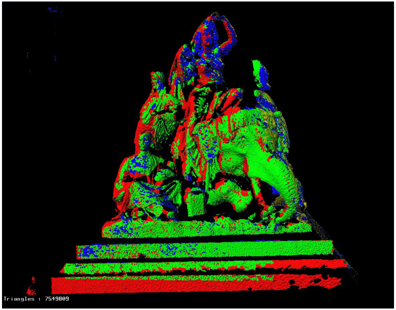




Visibility and discontinuity maps

⇒ Back-projection

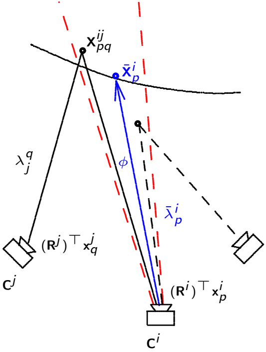
Registered depth maps



- Effective representation natural to input data
- Complexity linear in the number of reference cameras

Model Refinement

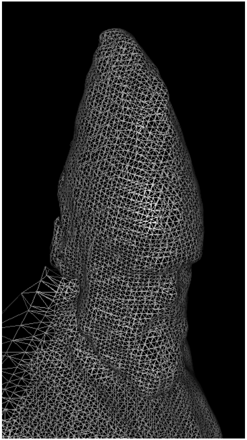
Model = depth, visibility and discontinuity maps + cameras



$$\mathbf{R}^j \mathbf{C}^i - \mathbf{R}^j \mathbf{C}^j + \bar{\lambda}^i_p \mathbf{R}^j (\mathbf{R}^i)^\top (\mathbf{K}^i)^{-1} \mathbf{x}^i_p = \lambda^j_q$$

- Camera-depth constraint for each correspondence (K-means like)
- Second-order surface model (depth continuity assumption)
- One global optimization problem with **depths** $\bar{\lambda}$ and **camera translations C** as free parameters

Surface Reconstruction and Refinement



- Change of representation to triangular mesh
- Depth maps merged with PSR [Kaz06]
- Good initial estimate of surface
- Use of camera calibration refined in previous step
- Refinement by combined stereo and contour matching for photo-consistency

[Kaz06] M. Kazhdan, M. Bolitho and H. Hoppe:
Poisson surface reconstruction. Eurographics 2006.

Photo-consistency Measure

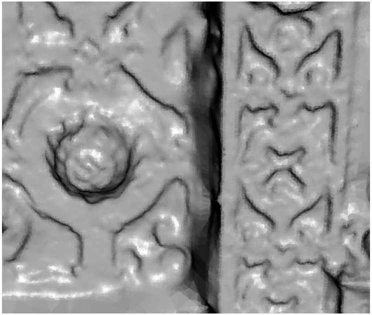
We define a stereo photo-consistency function (Normalized SSD)

$$\phi_I(\mathbf{X}) = \sum_{i,j \in V(\mathbf{X}),\, i \neq j} \frac{2\| I_i(\pi_i(\mathbf{X})) - I_j(\pi_j(\mathbf{X}))\|^2}{\sigma_i^2(\pi_i(\mathbf{X})) + \sigma_j^2(\pi_j(\mathbf{X}))} \tag{1}$$

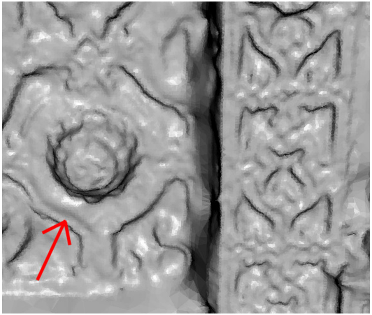
- Given world point $\mathbf{X}$, set of images I_i, $i = 1, \ldots, N$
- Images $I_i = I_i^0 - C_i$ are offset-corrected for overall color balance estimated from projections on current surface
- $V(\mathbf{X})$ is a set of images in which point $\mathbf{X}$ is visible
- $\pi_i(\mathbf{X}) \simeq \mathbf{P}_i\mathbf{X}$ is perspective projection function
- $\sigma_{i,j}$ independently pre-computed image variances (normalizing factors)

Photo-consistency Measure

What is the effect of offset correction?

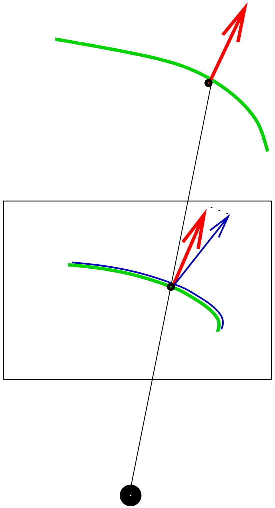


with offset correction



without offset correction

Contour Matching

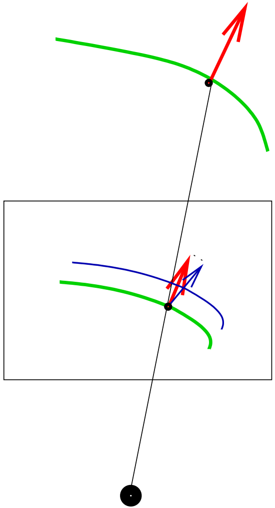


Projection of contour generators on a smooth surface should match local maxima of image gradient ∇I (apparent contours)

$$\phi_C(\mathbf{X}) = \frac{1}{|\Omega(\mathbf{X})|} \sum_{k \in \Omega(\mathbf{X})} \left| \left\langle \nabla I(\pi_k(\mathbf{X})), \varpi_k(\mathbf{N}(\mathbf{X})) \right\rangle \right|$$

- $\Omega(\mathbf{X})$ – set of cameras that see $\mathbf{X}$ as a contour point
- Avoids explicit detection of contours in images
- Takes direction into account
- Requires robust detection of contour vertices (paths)
- Smooth vs. sharp contour generators

Contour Matching

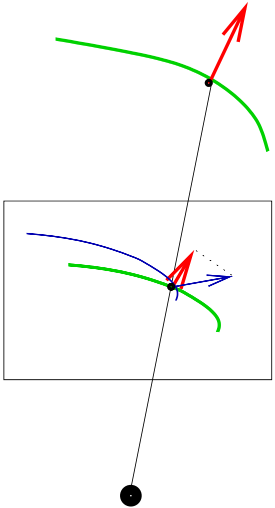


Projection of contour generators on a smooth surface should match local maxima of image gradient ∇I (apparent contours)

$$\phi_C(\mathbf{X}) = \frac{1}{|\Omega(\mathbf{X})|} \sum_{k \in \Omega(\mathbf{X})} \left| \left\langle \nabla I(\pi_k(\mathbf{X})), \varpi_k(\mathbf{N}(\mathbf{X})) \right\rangle \right|$$

- $\Omega(\mathbf{X})$ – set of cameras that see **X** as a contour point
- Avoids explicit detection of contours in images
- Takes direction into account
- Requires robust detection of contour vertices (paths)
- Smooth vs. sharp contour generators

Contour Matching



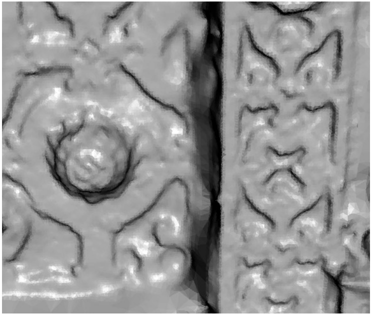
Projection of contour generators on a smooth surface should match local maxima of image gradient ∇I (apparent contours)

$$\phi_C(\mathbf{X}) = \frac{1}{|\Omega(\mathbf{X})|} \sum_{k \in \Omega(\mathbf{X})} \left| \left\langle \nabla I(\pi_k(\mathbf{X})), \varpi_k(\mathbf{N}(\mathbf{X})) \right\rangle \right|$$

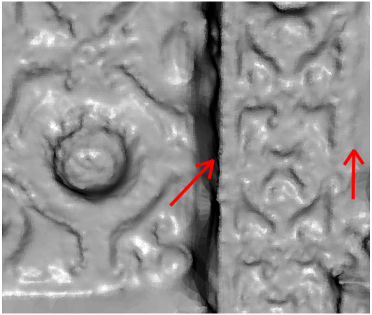
- $\Omega(\mathbf{X})$ – set of cameras that see **X** as a contour point
- Avoids explicit detection of contours in images
- Takes direction into account
- Requires robust detection of contour vertices (paths)
- Smooth vs. sharp contour generators

Photo-consistency Measure

What is the effect of contour matching?



with contours



without contours

Surface evolution

We define a surface energy

$$E_\Omega(S) = \int_S \Big(\phi_I(\mathbf{X}) - \alpha\phi_C(\mathbf{X})\Big)\, dA = \int_S \phi(\mathbf{X})\, dA \qquad (2)$$

combining stereo and contour matching and minimize it by iterative surface flow [2]

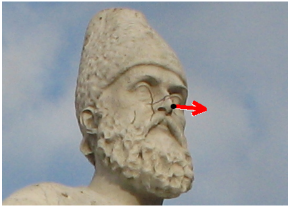
$$\frac{\partial S}{\partial t}(\mathbf{X}) = \Big(H(\mathbf{X})\phi(\mathbf{X}) - \langle \nabla\phi(\mathbf{X}), \mathbf{N}\rangle\Big)\mathbf{N}, \qquad (3)$$

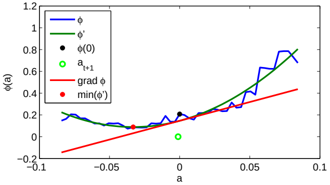
- $H(\mathbf{X})$ is the mean curvature of surface at point $\mathbf{X}$
- implicit regularization

[2] H.Jin: Variational methods for shape reconstruction in computer vision. PhD thesis, Washington Univ. (2003)

Computation of the gradient $\nabla\phi$

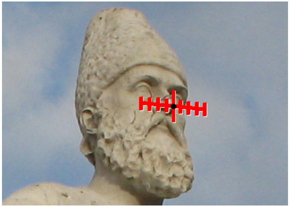
- Sampling of image points on projection of surface normal
- Second-order curve fitting for filtering
- Pixel-wide image sampling (needs adequate mesh resolution)
- Coarse-to-fine strategy (scale-space approach)
- Decreasing window size (by 5% in iteration down to 0.1 of original size)

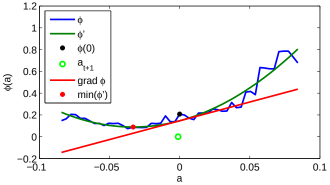




Computation of the gradient $\nabla\phi$

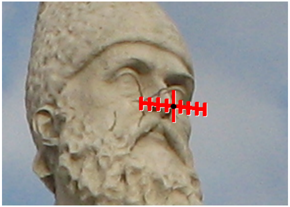
- Sampling of image points on projection of surface normal
- Second-order curve fitting for filtering
- Pixel-wide image sampling (needs adequate mesh resolution)
- Coarse-to-fine strategy (scale-space approach)
- Decreasing window size (by 5% in iteration down to 0.1 of original size)

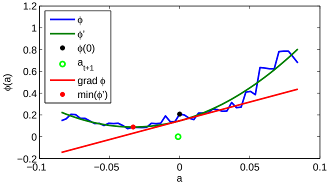




Computation of the gradient $\nabla \phi$

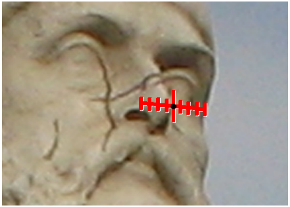
- Sampling of image points on projection of surface normal
- Second-order curve fitting for filtering
- Pixel-wide image sampling (needs adequate mesh resolution)
- Coarse-to-fine strategy (scale-space approach)
- Decreasing window size (by 5% in iteration down to 0.1 of original size)

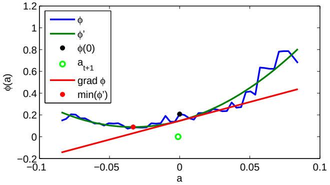




Computation of the gradient $\nabla\phi$

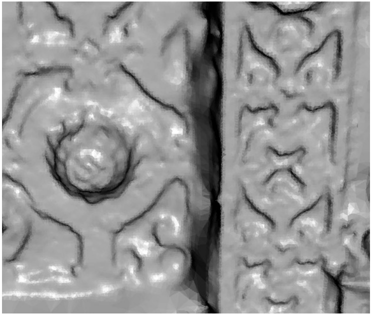
- Sampling of image points on projection of surface normal
- Second-order curve fitting for filtering
- Pixel-wide image sampling (needs adequate mesh resolution)
- Coarse-to-fine strategy (scale-space approach)
- Decreasing window size (by 5% in iteration down to 0.1 of original size)



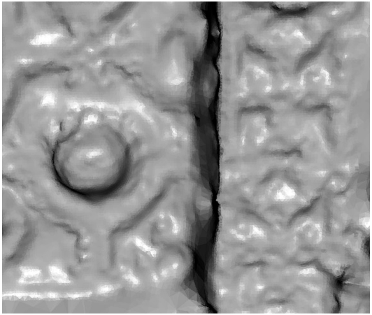


Computation of the gradient $\nabla\phi$

What is the effect of scale-space approach?

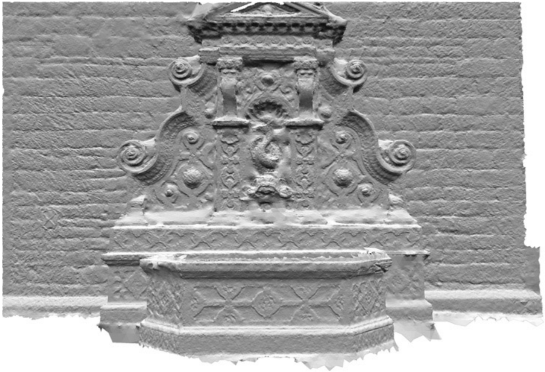


variable scale



fixed scale

Experiments on Standard datasets



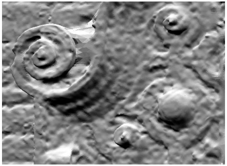
fountain-P11

- Increase of accuracy
- Edges emphasized
- Higher surface quality
- Flat surfaces smooth

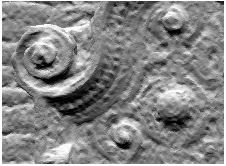
Experiments on Standard datasets

Fountain-P11 dataset detailed rendering.

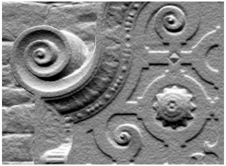
a) input image



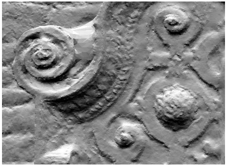
c) depth map fusion



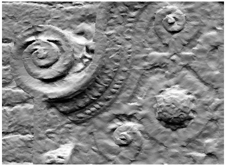
e) result of FUR



b) ground truth

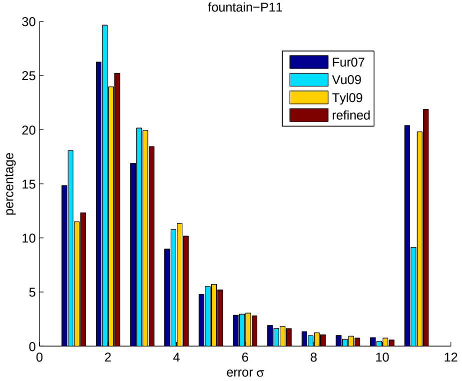


d) mesh refinement



f) result of VU

Evaluation on Standard datasets



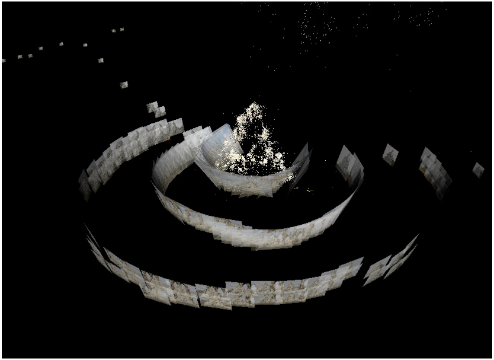
- Ground truth from laser scanners
- Surface projected to cameras
- Depth measurement error σ
- Most details below ground truth error σ
- Completeness vs. Accuracy

`http://cvlab.epfl.ch/~strecha/multiview/denseMVS.html`

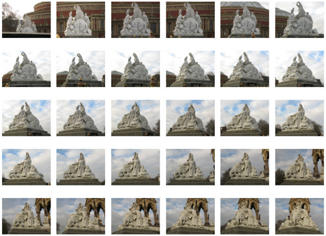
[Fur07] Y.Furukawa, J.Ponce: Accurate, dense, and robust multi-view stereopsis. CVPR 2007.

[Vu09] H.Vu, R.Keriven, P.Labatut, J.P.Pons: Towards high-resolution large-scale multi-view stereo. CVPR 2009.

Experiments on large outdoor dataset

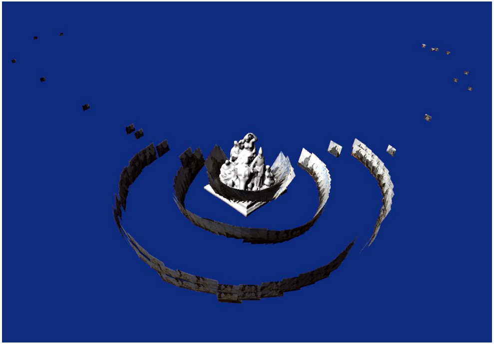


Asia scene



input images (238)

Experiments on large outdoor dataset

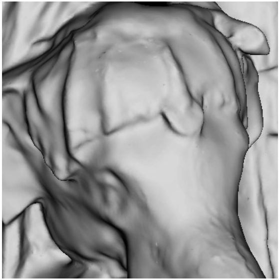


Asia refined result

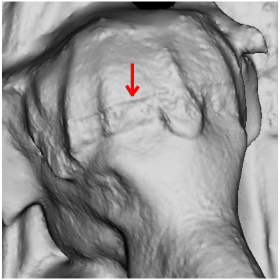
textured

Results on the Asia dataset.

a) depth map fusion



b) elephant's head



c) refined detail

Summary

Refinement of Surface Mesh
for Accurate Multi-View Reconstruction

- Pipeline for accurate 3D reconstruction
- Surface reconstruction with Depth Map Fusion
- Camera calibration refinement
- Image offset correction
- Photometric mesh refinement
  - Combining stereo and contour matching
  - Scale-space approach

Thank you.