Inline Computational Imaging meets Convex Optimization

Thomas Pock, Svorad Štolc et al.
pock@icg.tugraz.at
svorad.stolc@ait.ac.at

AIT Austrian Institute of Technology GmbH
Center for Vision, Automation & Control
Vienna, Austria

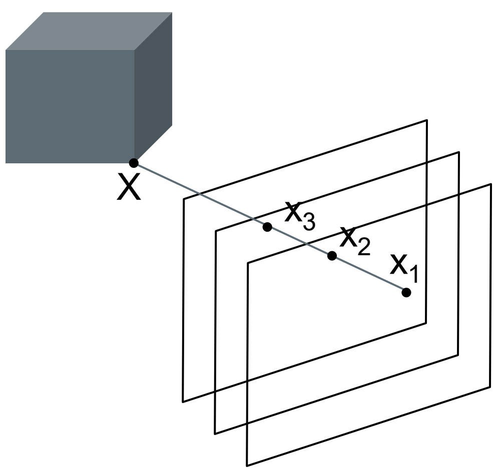
www.ait.ac.at/hpv

3D Reconstruction - Overview

- Acquired data combines multi-view stereo and photometry information
- Our 3D reconstruction is based on combining two complementary techniques:
  - Multi-view stereo
  - Photometric stereo
- Combining stereo and photometry yields unique depth solutions [Chambolle 1992]

Robust feature computation

- For computing 3D information we need to find corresponding points in the different views
- We apply a high-pass feature transform that is robust with respect to different lighting angles

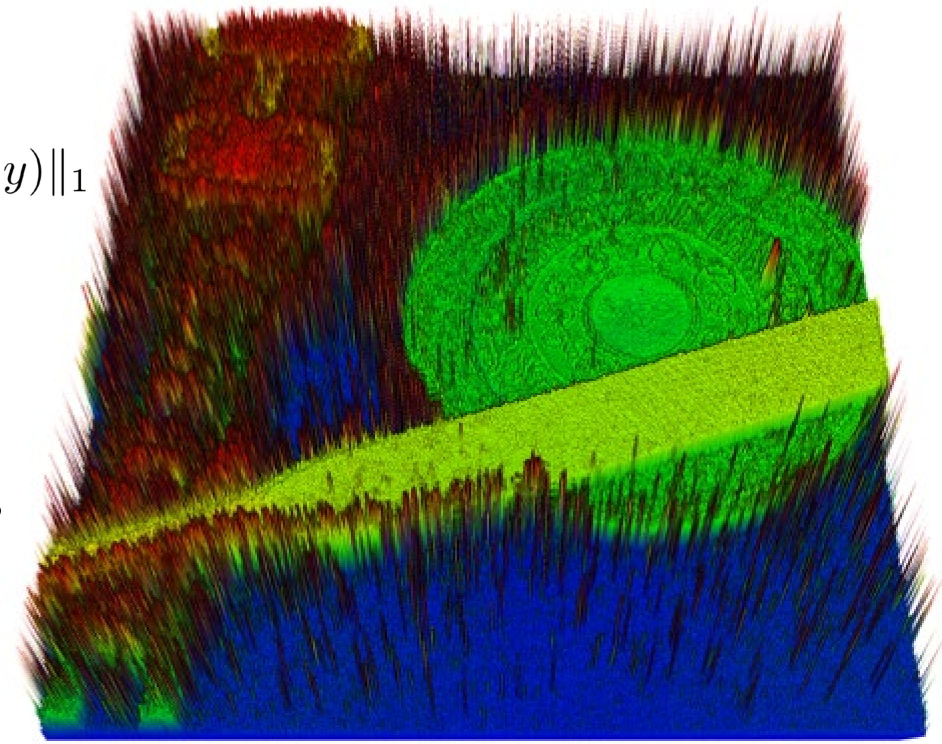


Stereo Matching

- Features matching for all possible depth values yields a cost volume

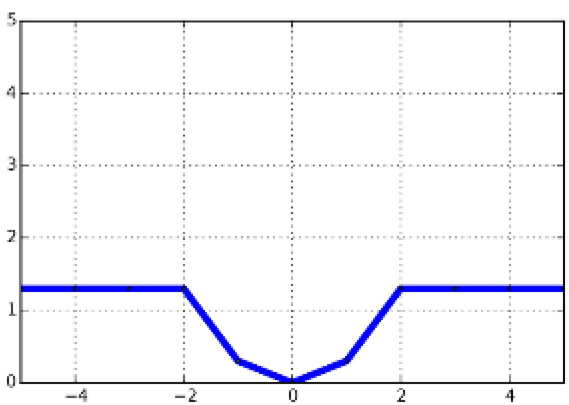
$$C(x,y,d) = \sum_{i=1}^{N} \|f_0(x,y) - f_i(w(x,d),y)\|_1$$

- Computing the best match in the cost volume yields a very noisy 3D model
- A global smoothness assumption is needed
- We use an energy minimization approach based on a Markov Random (MRF) field formulation

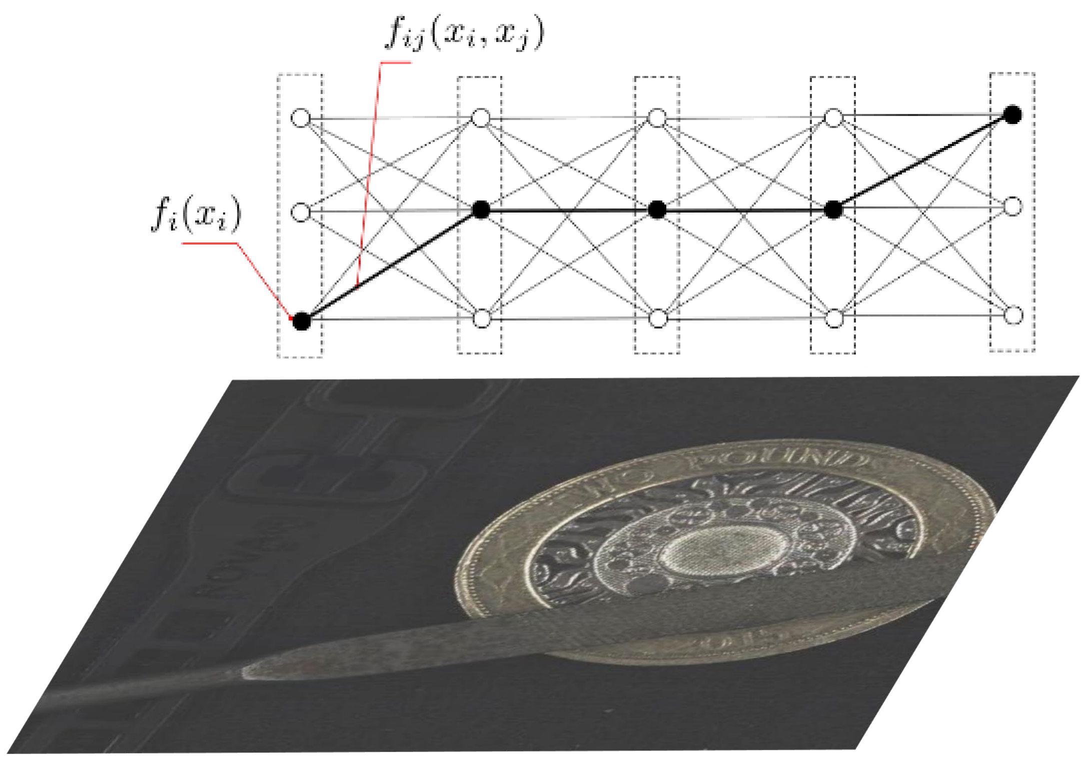


Markov Random Field (MRF)

$$\min_{x \in \mathcal{L}^{\mathcal{V}}} f(x) := \sum_{i \in \mathcal{V}} f_i^T x_i + \sum_{ij \in \mathcal{E}} x_i^T f_{ij} x_j, \quad \mathcal{L} = \{x \in \{0,1\}^D : 1^T x = 1\}$$



Smoothness potential function



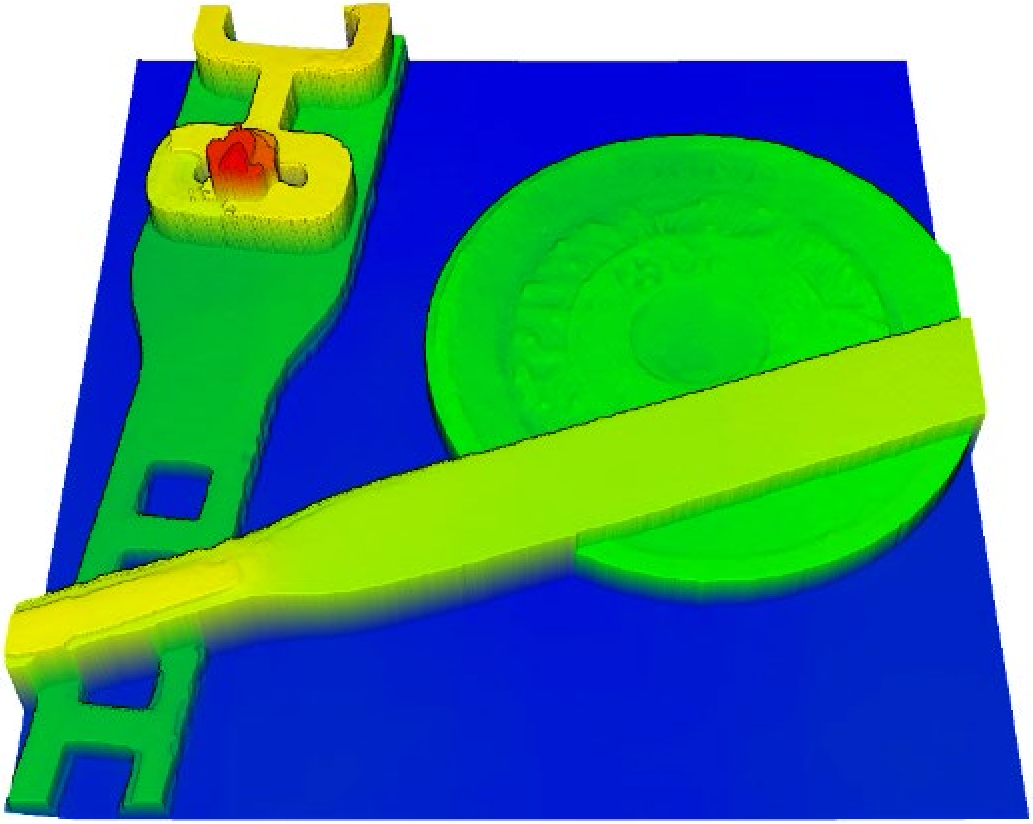
Solving the MRF

- Solving the MRF model is NP-hard
- We instead solve the following Linear Programming (LP) relaxation [Schlesinger 1976]

$$\begin{aligned}
\min_{v,w} \quad & \sum_{i \in \mathcal{V}} \sum_{l} f_i(l) v_i^l + \sum_{\{i,j\} \in \mathcal{E}} \sum_{l,m} f_{i,j}(l,m) w_{i,j}^{l,m}, \\
\text{s.t.} \quad & \sum_{l} v_i^l = 1,\ i \in \mathcal{V}, \quad v_i^l \geq 0,\ i \in \mathcal{V},\ l \in \mathcal{L}, \\
& \sum_{l} w_{i,j}^{l,m} = v_j^m,\ \{i,j\} \in \mathcal{E},\ m \in \mathcal{L}, \\
& w_{i,j}^{l,m} \geq 0,\ \{i,j\} \in \mathcal{E},\ l,m \in \mathcal{L}.
\end{aligned}$$

- The LP is of very large scale (W x H x L^{2}) and hence hard to solve
- We apply an entropic smoothing approach and derive the dual problem
- The dual problem can be solved via block-coordinate descent
- Can be efficiently implemented on graphics processing units (GPUs)

Result of the MRF



Photometric Stereo

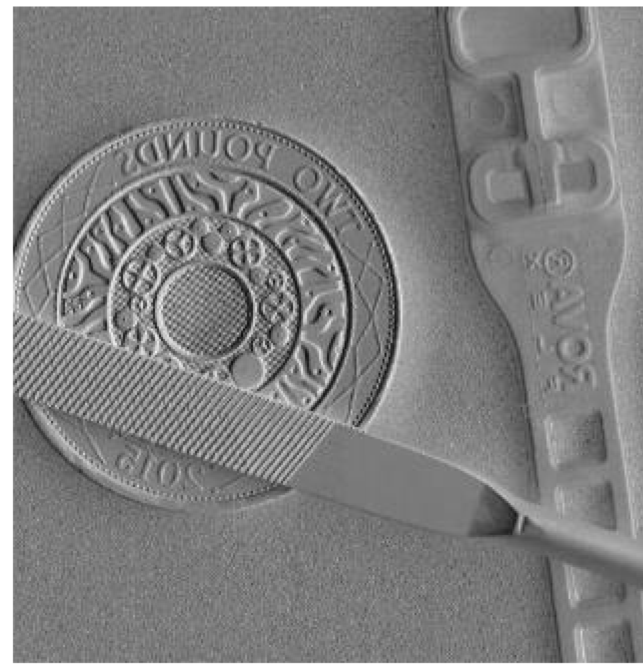
- The light field data also contains photometric information
- The surface normal information is given only in the transport direction
- We pre-compute the surface gradient in the transport direction

$$n(x,y) = \begin{pmatrix} n^x(x,y) \\ n^y(x,y) \\ n^z(x,y) \end{pmatrix} \qquad g(x,y) = \begin{pmatrix} \nabla_x d(x,y) \\ \nabla_y d(x,y) \\ 1 \end{pmatrix}$$

$$n = \frac{g(x,y)}{\sqrt{(\nabla_x d(x,y))^2 + (\nabla_y d(x,y))^2 + 1}}$$

$$g^x(x,y) = \frac{n^x(x,y)}{n^z(x,y)}$$

$$g^y(x,y) = \frac{n^y(x,y)}{n^z(x,y)}$$



$g^x(x,y)$

Total Generalized Variation (TGV) regularization

- Having an estimate for the gradient of the depth map in transport direction, we can compute a refined depth map using the total generalized variation [Bredies, Kunisch, Pock 2007]

$$\min_{u,v} \underbrace{\alpha_0 \|\nabla u - v\|_{2,1}}_{\text{first order smoothness}} + \underbrace{\alpha_1 \|\nabla v\|_{2,1}}_{\text{second order smoothness}} + \underbrace{\frac{\mu}{2}\|v^x - g^x\|^2}_{\text{initial gradient map}} + \underbrace{\frac{1}{2}\|u - d\|^2}_{\text{initial depth map}}$$

- The missing y-gradient is recovered from the regularization term
- Convex but non-smooth functional
- Can be minimized using first-order primal-dual algorithms [Chambolle, Pock, 2011]

$$\begin{cases} x^{n+1} &= \operatorname{prox}_{\tau G}(x^n - \tau K^* y^n) \\ y^{n+1} &= \operatorname{prox}_{\sigma F^*}(y^n + \sigma K(x^{n+1} + \theta(x^{n+1} - x^n))) \end{cases}$$

Result of TGV regularization

AIT Inline Computational Imaging: Industrial Use Cases