MeshStereo: A Global Stereo Model with Mesh Alignment Regularization for View Interpolation

Chi Zhang[1], Zhiwei Li[2], Rui Cai[2], Yanhua Cheng[2], Hongyang Chao[1], Yong Rui[2]

1 Sun Yat-Sen University, Guangzhou, P.R. China

2 Microsoft Research, Beijing, P.R. China

DECEMBER 11-18, 2015 ICCV15 International Conference on Computer Vision

Motivation

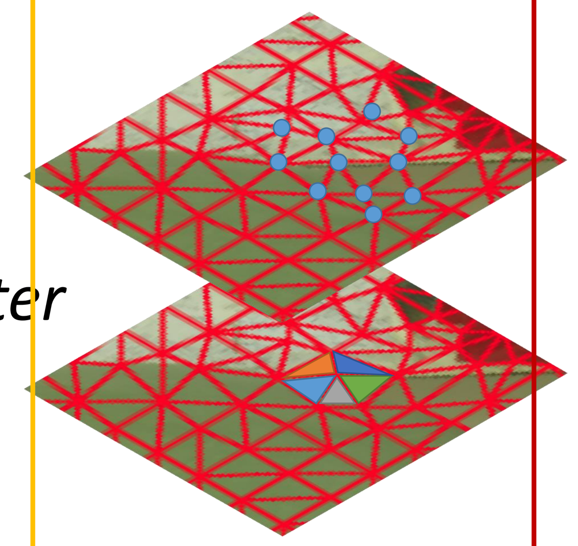
- *Output high-quality meshes for view interpolation*
- *Unify depth map estimation and mesh generation*

Variables

$\boldsymbol{\alpha}$ *A splitting probability for each 2D vertex*

$\mathbf{D}$ *A depth value for each barycenter*

$\mathbf{N}$ *A normal for each triangle*



Optimization

$$E_{\text{UPPER}} = E_{\text{Alignment}} + E_{\text{SplitPenalty}} + E_{\text{SplitSmooth}}$$

Optimize alternatively between the two layers until convergence

$$E_{\text{LOWER}} = E_{\text{Alignment}} + E_{\text{MatchingCost}} + E_{\text{NormalSmooth}}$$

$\min_{\boldsymbol{\alpha}} E_{\text{UPPER}}$ *Quadratic in* $\boldsymbol{\alpha}$*, has closed-form solution*

$\min_{\mathbf{N},\mathbf{D}} E_{\text{LOWER}}$ *Non-convex, difficult, Relax it and optimize in another loop*

$$\min_{\mathbf{N},\mathbf{D}} E_{\text{Lower}} \overset{\theta\to\infty}{\equiv} \min_{\mathbf{N},\mathbf{D},\tilde{\mathbf{N}},\tilde{\mathbf{D}}} \left\{ E_{\text{MatchingCost}}(\tilde{\mathbf{N}},\tilde{\mathbf{D}}) + \frac{\theta}{2} E_{\text{Couple}}(\mathbf{N},\mathbf{D},\tilde{\mathbf{N}},\tilde{\mathbf{D}}) + E_{\text{NormalSmooth}}(\mathbf{N},\mathbf{D}) + E_{\text{Alignment}}(\mathbf{N},\mathbf{D}) \right\}$$

where $E_{\text{Couple}}(\mathbf{N},\mathbf{D},\tilde{\mathbf{N}},\tilde{\mathbf{D}}) = \sum_i (\boldsymbol{\Pi}_i - \tilde{\boldsymbol{\Pi}}_i)^\top \Sigma (\boldsymbol{\Pi}_i - \tilde{\boldsymbol{\Pi}}_i)$

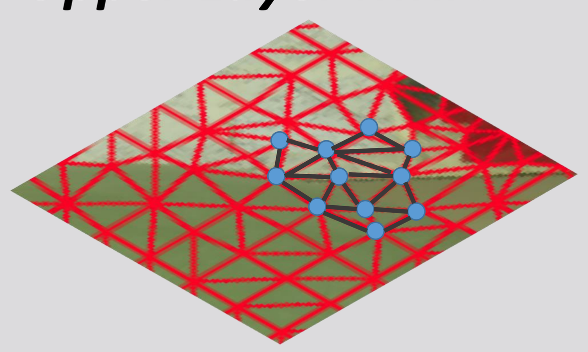
where $\boldsymbol{\Pi}_i = [\mathbf{n}_i^\top, d_i]^\top$, and $\sigma = \text{diag}(\sigma_n, \sigma_n, \sigma_n, \sigma_d)$

- *Increase* θ *from 0 to* ∞*, optimize alternatively between blue and green*
- *Optimize blue part by PatchMatch*
- *Optimize green part in closed-form*

Formulation

$$E_{\text{All}}(\mathbf{N},\mathbf{D},\boldsymbol{\alpha}) = E_{\text{MatchingCost}} + E_{\text{NormalSmooth}} + E_{\text{Alignment}} + E_{\text{SplitPenalty}} + E_{\text{SplitSmooth}}$$

Upper Layer MRF:



$$E_{\text{SplitPenalty}}(\boldsymbol{\alpha}) = \sum_s \alpha_s \cdot \tau_s$$

$$\tau_s = \exp(-|\nabla I^3(x_s,y_s)|/\gamma_1)$$

- *Penalize splitting at homogeneous regions*

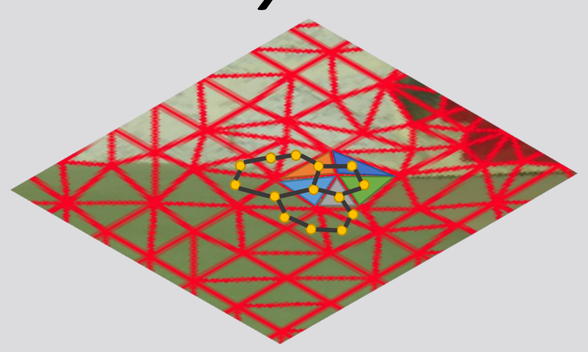
$$E_{\text{SplitSmooth}}(\boldsymbol{\alpha}) \sum_{s,t\in\mathcal{N}} w_{st}(\alpha_s - \alpha_t)^2$$

$$w_{st} = \exp(-|k(\mathbf{x}_s) - k(\mathbf{x}_t)|/\gamma_2)$$

$$k(\mathbf{x}) = \arg\max_j \{|I^l(\mathbf{x}) - I(\mathbf{x})| < 10, \forall l \le j\}$$

- *Encourage similar splitting properties when neighboring vertices have similar 'visual complexity'*

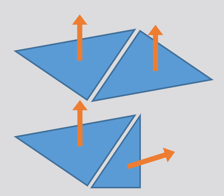
Lower Layer MRF:



$$E_{\text{MatchingCost}}(\mathbf{N},\mathbf{D}) = \sum_i \sum_{p\in\text{Tri}_i} \rho(\mathbf{n}_i, d_i, p)$$

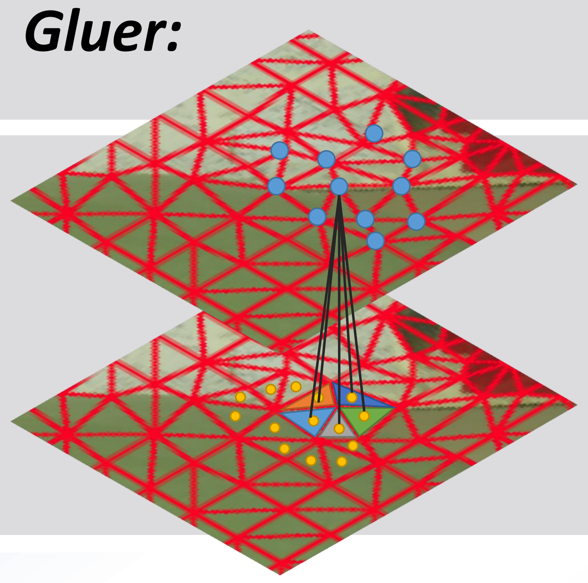
- *Each depth-normal pair induce a disparity plane over that triangle*
- *Matching cost combines census and gradient features*

$$E_{\text{NormalSmooth}}(\mathbf{N}) = \sum_{i,j\in\mathcal{N}} w_{ij}(\mathbf{n}_i - \mathbf{n}_j)^\top(\mathbf{n}_i - \mathbf{n}_j)$$



- *Encouraged*
- *Discouraged*

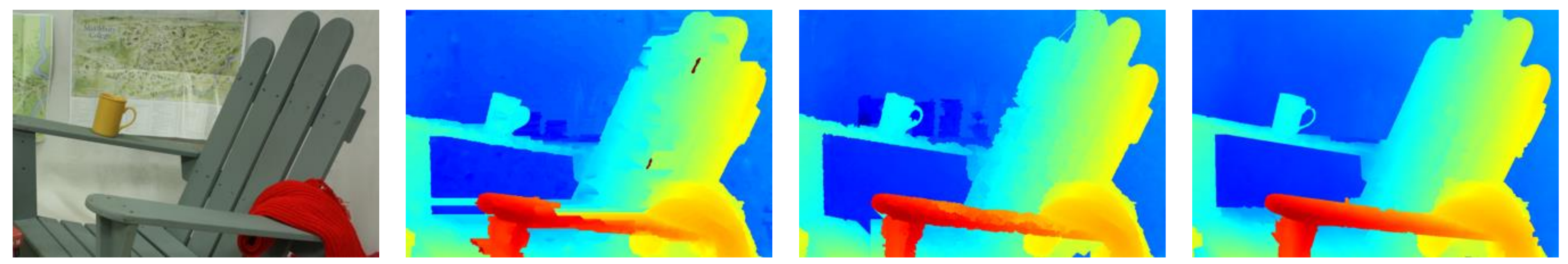
Gluer:



$$E_{\text{Alignment}}(\mathbf{N},\mathbf{D},\boldsymbol{\alpha}) = \sum_s (1-\alpha_s) \cdot \sum_{i,j\in G_s} w_{ij}\left(\mathcal{D}_i(\mathbf{x}_s) - \mathcal{D}_j(\mathbf{x}_s)\right)^2$$

- *Require tight alignment at non-splitting vertices*
- *Not active at splitting vertices*

Stereo Results



Color | ELAS | SGM | Ours

- *Preserve fine structures*
- *First place on Middlebury 3.0 at submission time*

Generated Meshes