A convex Solution to Disparity Estimation from Light Fields via the Primal-Dual Method

Mahdad Hosseini Kamal[*], Paolo Favaro[**], and Pierre Vandergheynst[*]

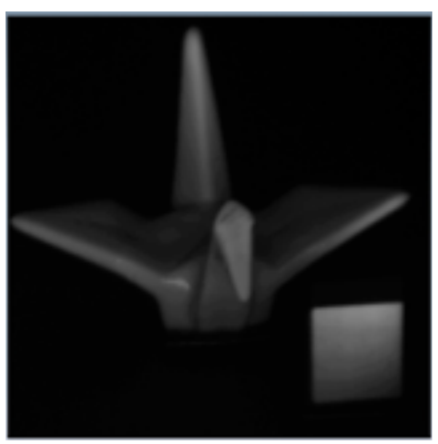
[*] Ecole Fédérale Polytechnique de Lausanne (EPFL)
[**] University of Bern

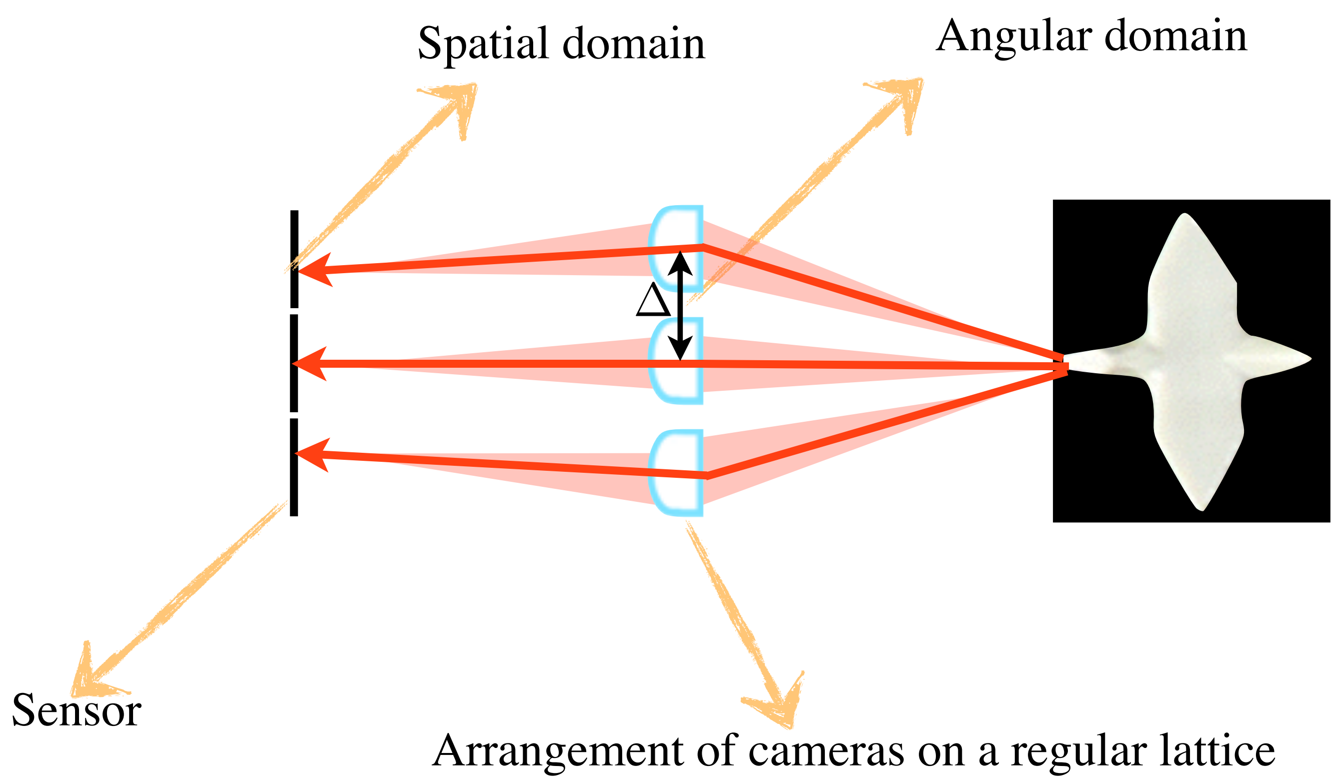
Light Fields

- Definition:
  - The amount of light traveling in every direction through every point in space

Light Fields

- Definition:
  - The amount of light traveling in every direction through every point in space





Light Fields

- Definition:
  - The amount of light traveling in every direction through every point in space
- Applications:
  - Generating a new view [1]

[1] S.J. Gortler et al., SIGGRAPH 1996, M. Levoy et al. SIGGRAPH 1996

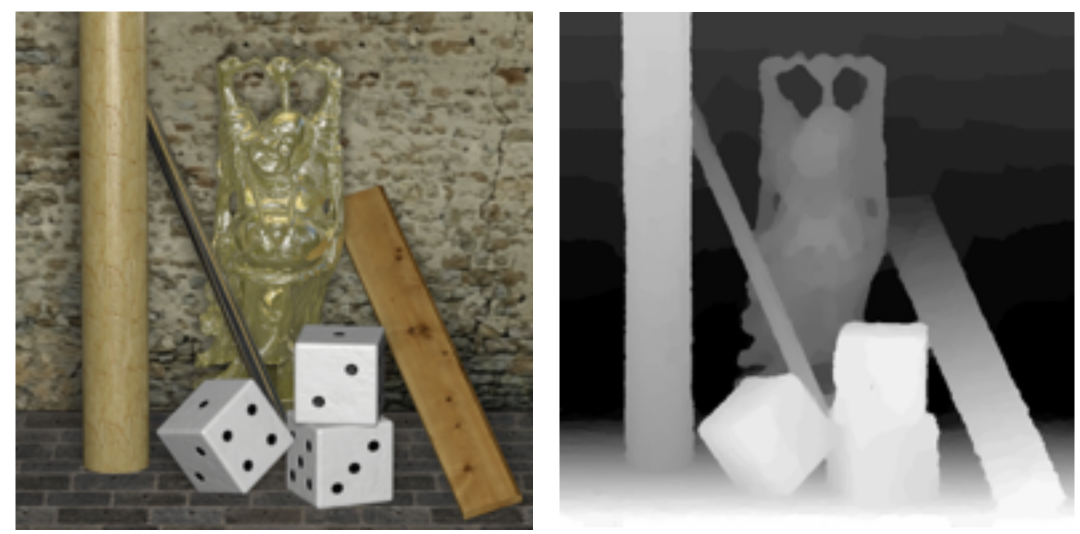
Light Fields

- Definition:
  - The amount of light traveling in every direction through every point in space
- Applications:
  - Generating a new view [1]
  - Digital refocusing [2]

[1] S.J. Gortler et al., SIGGRAPH 1996, M. Levoy et al. SIGGRAPH 1996
[2] R. Ng et al., CSTR 05

Light Fields

- Definition:
  - The amount of light traveling in every direction through every point in space
- Applications:
  - Generating a new view [1]
  - Digital refocusing [2]
  - Disparity estimation [3]



[1] S.J. Gortler et al., SIGGRAPH 1996, M. Levoy et al. SIGGRAPH 1996
[2] R. Ng et al., CSTR 05
[3] S. Heber et al. EMMCVPR 13, ECCV 14, S. Wanner et al. CVPR 12

Light Fields

- Definition:
  - The amount of light traveling in every direction through every point in space
- Applications:
  - Generating a new view [1]
  - Digital refocusing [2]
  - Disparity estimation [3]

Leverage low-dimensional structures in high-dimensional data

[1] S.J. Gortler et al., SIGGRAPH 1996, M. Levoy et al. SIGGRAPH 1996
[2] R. Ng et al., CSTR 05
[3] S. Heber et al. EMMCVPR 13, ECCV 14, S. Wanner et al. CVPR 12

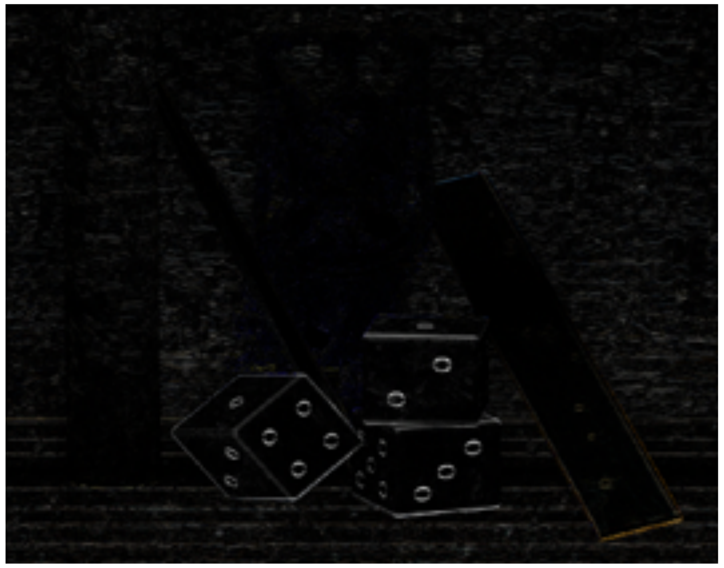
Multi-views and Light Fields

- Light field views are related by ***disparity*** $\rho : \Omega \mapsto [0, +\infty)$

Angular distance

$$L_{i,j}(x, y) = L_{p,q}(x - \rho\Delta_x(p - i), y - \rho\Delta_y(q - j))$$

Disparity



Displacement

Center view: $L_{i,j}$

Neighbor view: $L_{p,q}$

Multi-views and Light Fields

- Disparity estimation follows
  - Convex formulation in discrete domain → no initial estimate
  - High-dimensional representation

$$\min_{\rho} \sum_{\substack{i,j,p>i \\ q>j,x,y}} \Phi(L_{i,j}(x,y) - L_{p,q}(x - \rho(p-i)\Delta_x, y - \rho(q-j)\Delta_y)) + \Gamma(\rho)$$

Robust penalty term

Prior knowledge

Generating the Patch Matrix

- Extracting patches from each light field
- Collect all patches in a matrix
- True disparity → columns of the patch matrix are identical

Center view

View 1

View 2

Generating the Patch Matrix

- Extracting patches from each light field
- Collect all patches in a matrix
- True disparity → columns of the patch matrix are identical

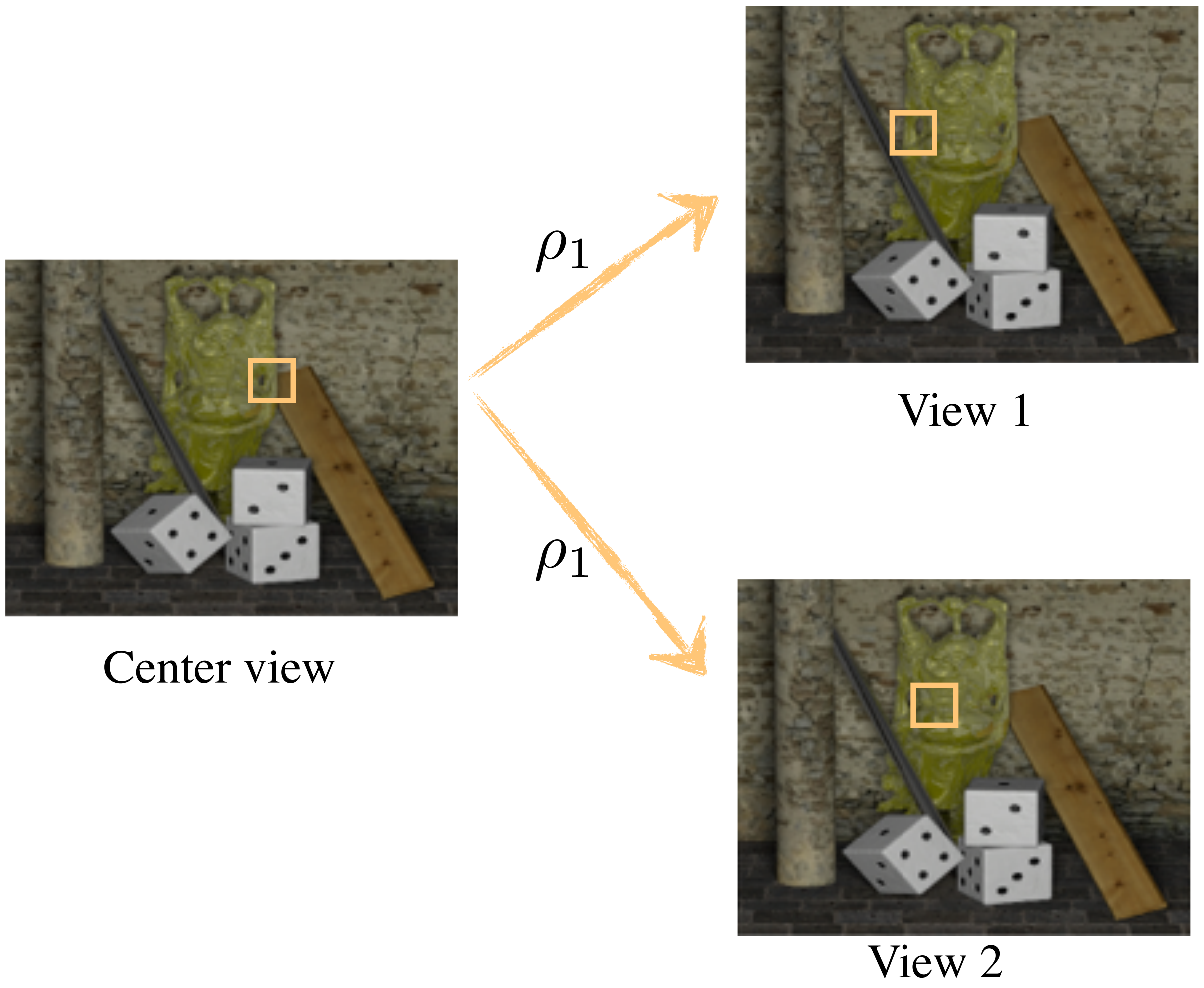
Center view

View 1

View 2

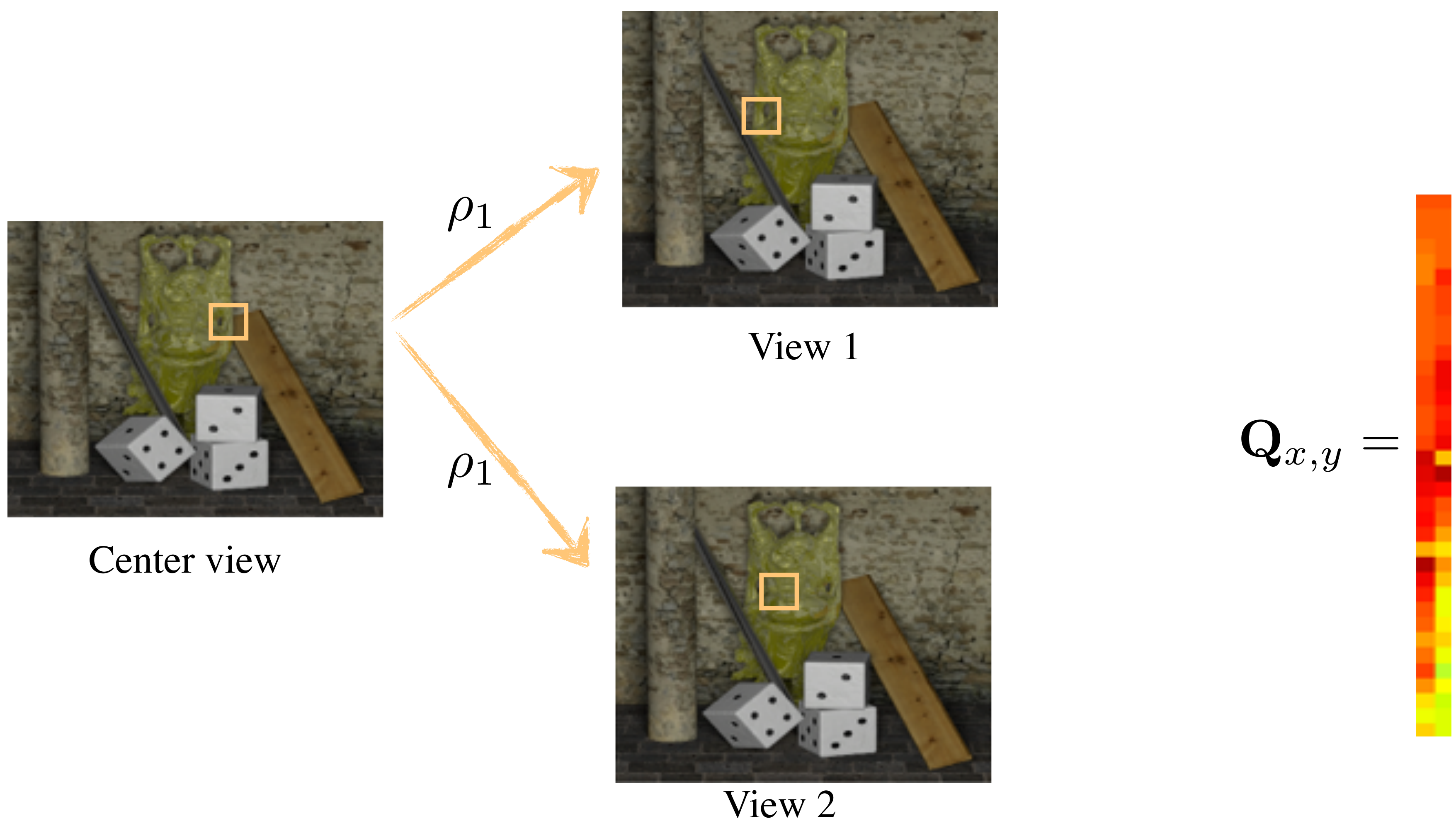
Generating the Patch Matrix

- Extracting patches from each light field
- Collect all patches in a matrix
- True disparity → columns of the patch matrix are identical



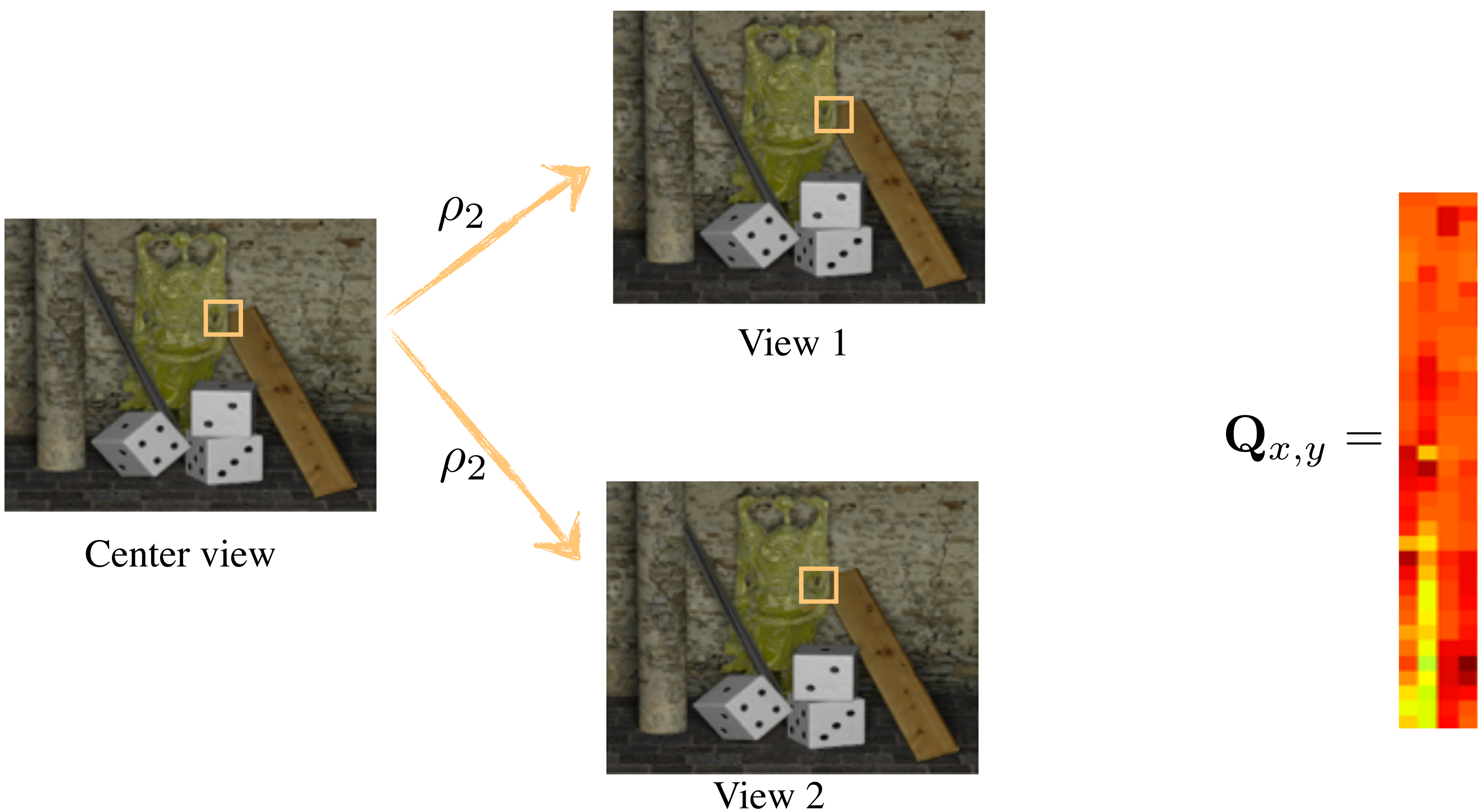
Generating the Patch Matrix

- Extracting patches from each light field
- Collect all patches in a matrix
- True disparity → columns of the patch matrix are identical



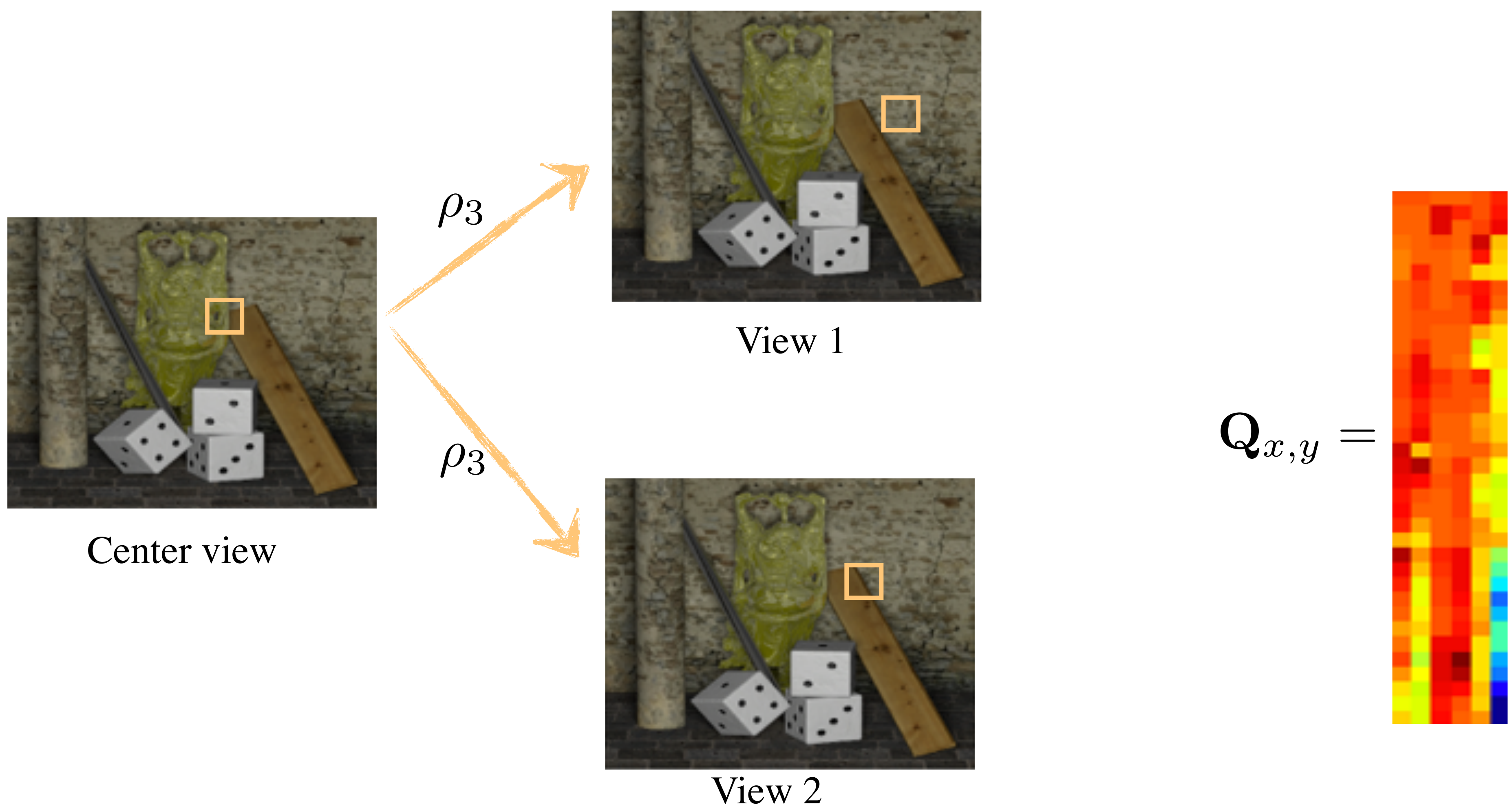
Generating the Patch Matrix

- Extracting patches from each light field
- Collect all patches in a matrix
- True disparity → columns of the patch matrix are identical



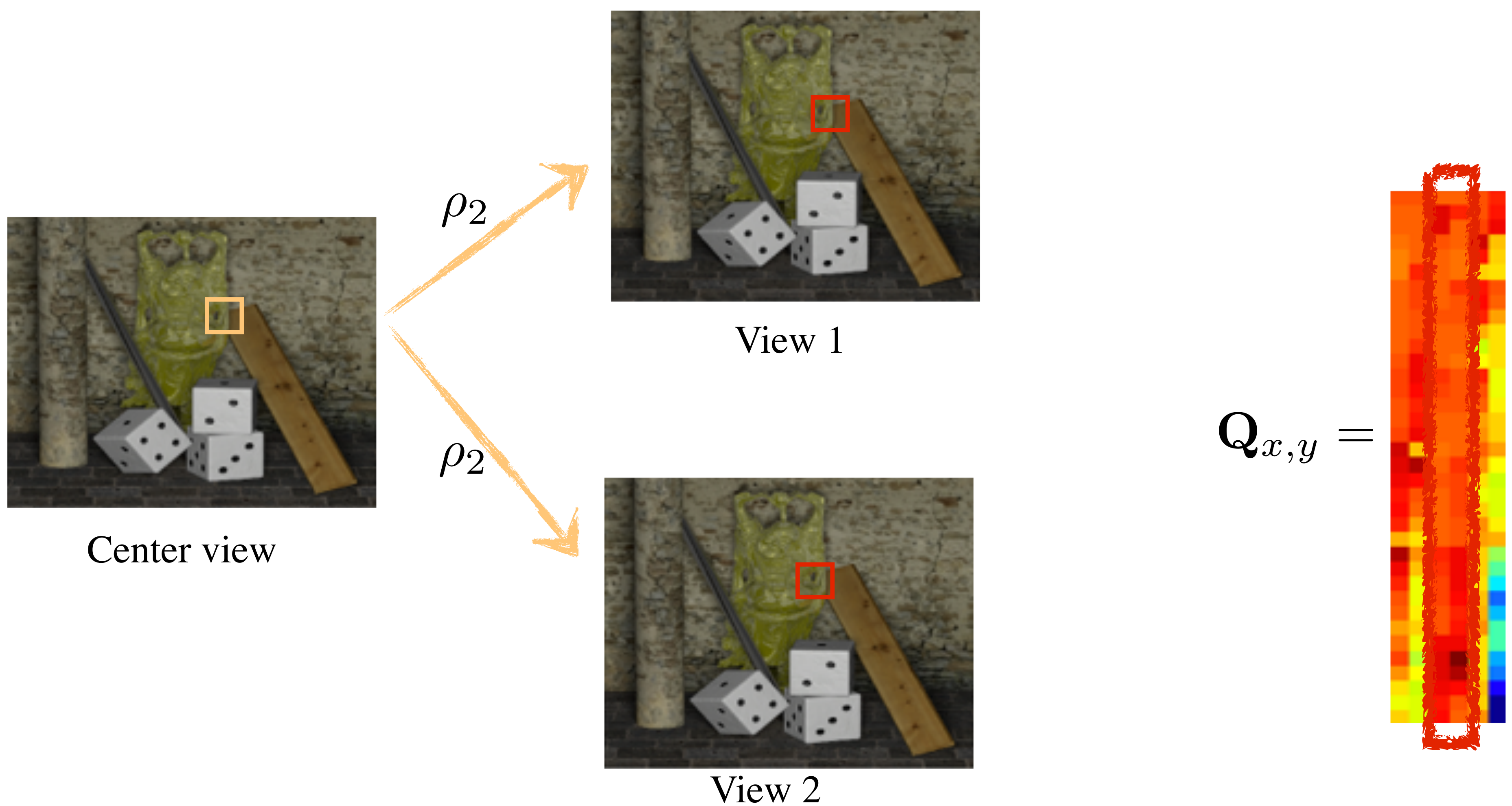
Generating the Patch Matrix

- Extracting patches from each light field
- Collect all patches in a matrix
- True disparity → columns of the patch matrix are identical



Generating the Patch Matrix

- Extracting patches from each light field
- Collect all patches in a matrix
- True disparity → columns of the patch matrix are identical



Generating the Patch Matrix

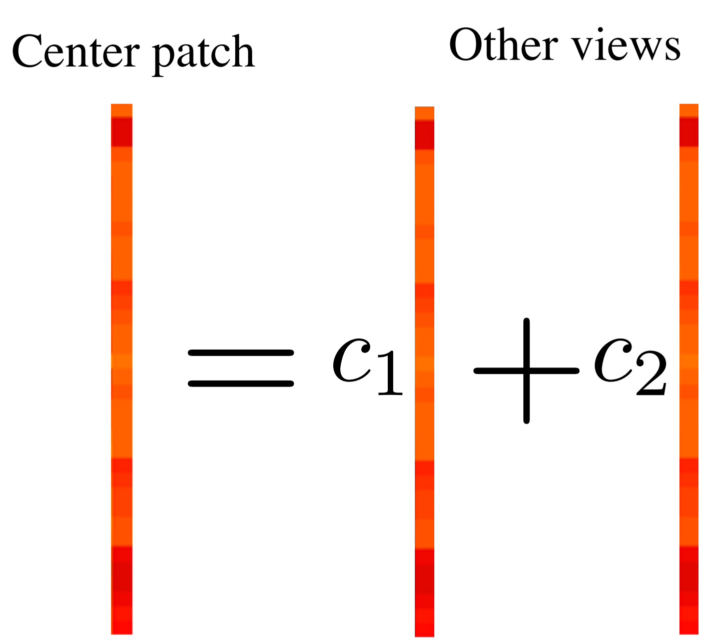
- Most of the time Lambertian approximation holds (low-rank property of light field)

Patch matrix

Patch weights

Center patch

$$\mathbf{y}_{x,y} = \mathbf{Q}^{\rho}_{x,y}\mathbf{c}^{\rho}_{x,y}$$

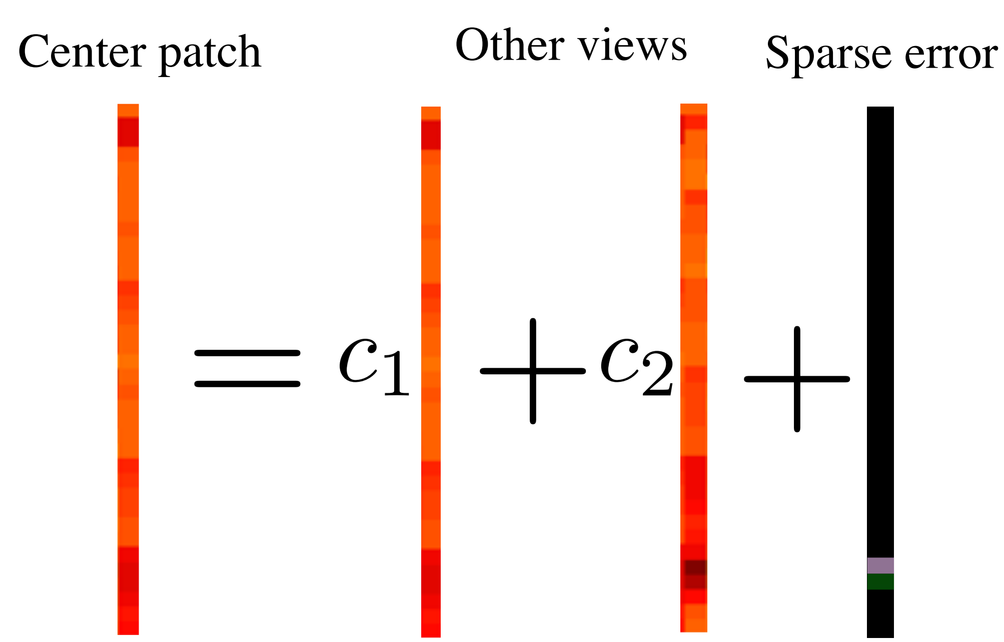


Generating the Patch Matrix

- Most of the time Lambertian approximation holds (low-rank property of light field)
- Non-lambertianity, occlusion, specularities → sparse representation

Patch matrix ↖ Patch weights ↗

Center patch ← $\mathbf{y}_{x,y} = \mathbf{Q}^{\rho}_{x,y}\mathbf{c}^{\rho}_{x,y} + \mathbf{e}_{x,y}$ → Sparse noise



Generating the Patch Matrix

- Most of the time Lambertian approximation holds (low-rank property of light field)
- Non-lambertianity, occlusion, specularities → sparse representation

Patch matrix

Patch weights

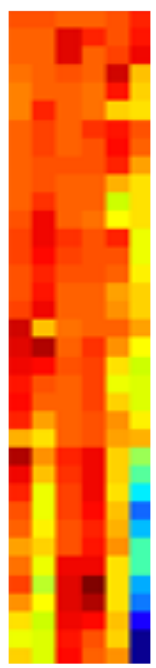
Center patch ← $\mathbf{y}_{x,y} = \mathbf{Q}^{\rho}_{x,y}\mathbf{c}^{\rho}_{x,y} + \mathbf{e}_{x,y}$ → Sparse noise

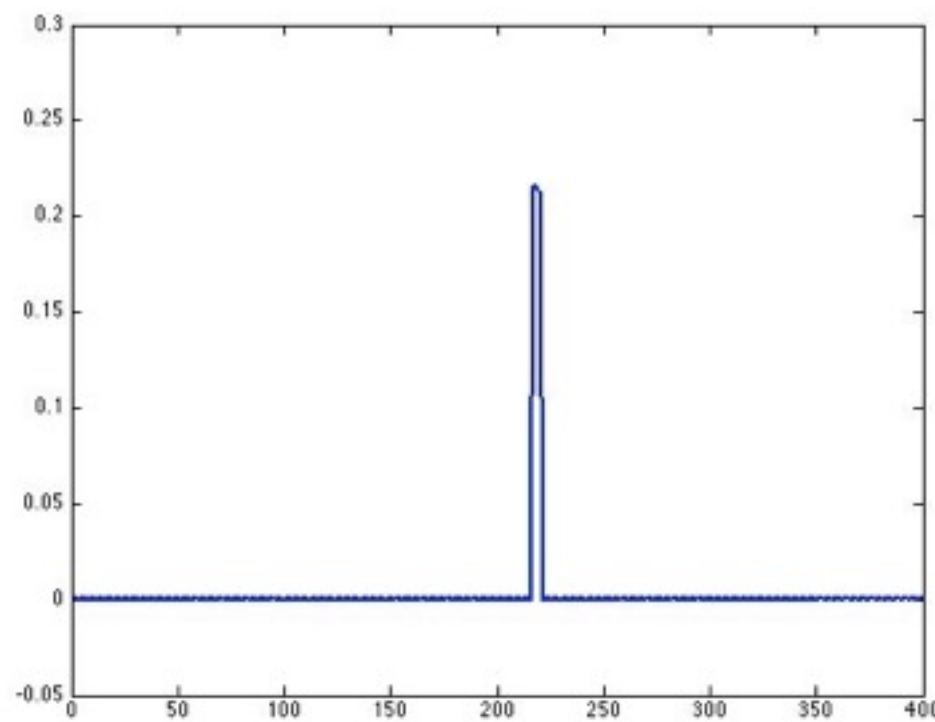
$$\mathbf{y}_{x,y} = [\mathbf{Q}^{\rho_1}_{x,y}\ \mathbf{Q}^{\rho_2}_{x,y}\ \ldots\ \mathbf{Q}^{\rho_D}_{x,y}][\mathbf{c}^{\rho_1}_{x,y}\mathbf{c}^{\rho_2}_{x,y}\cdots\mathbf{c}^{\rho_D}_{x,y}]^T + \mathbf{e}_{x,y} \doteq \mathbf{Q}_{x,y}\mathbf{c}_{x,y} + \mathbf{e}_{x,y}$$

Discrete disparity range

Disparity Estimation

- Group members → weights for each disparity hypothesis
- Group Lasso → non-zero entries in patch weights $\mathbf{c}_{x,y}$

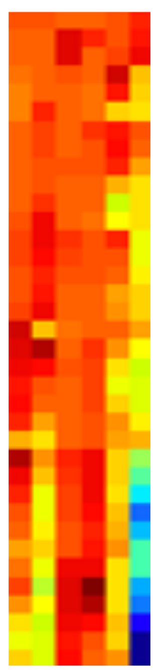
$\mathbf{Q}_{x,y} =$ 

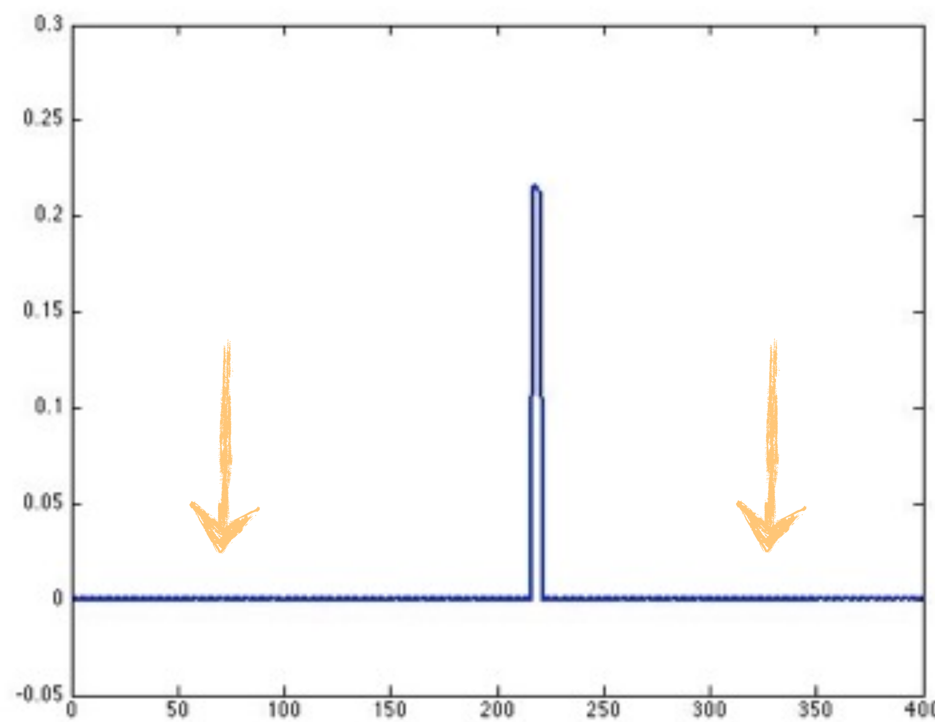


$\mathbf{c}_{x,y}$ Patch weights

Disparity Estimation

- Group members → weights for each disparity hypothesis
- Group Lasso → non-zero entries in patch weights $\mathbf{c}_{x,y}$

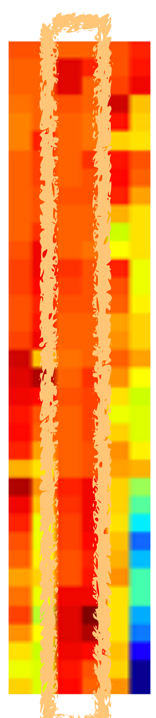
$\mathbf{Q}_{x,y} =$ 

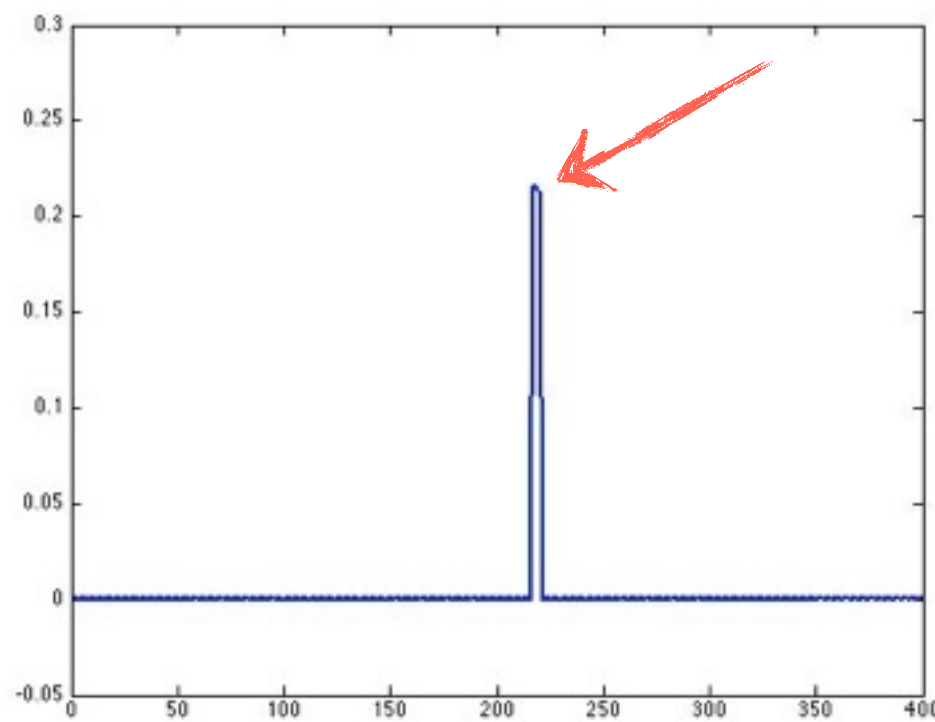


$\mathbf{c}_{x,y}$ Patch weights

Disparity Estimation

- Group members → weights for each disparity hypothesis
- Group Lasso → non-zero entries in patch weights $\mathbf{c}_{x,y}$

$\mathbf{Q}_{x,y} =$ 



$\mathbf{c}_{x,y}$ Patch weights

Disparity Estimation

- Group members → weights for each disparity hypothesis
- Group Lasso → non-zero entries in patch weights $\mathbf{c}_{x,y}$

$$||\mathbf{c}||_{2,1} = \sum_{x,y} \sum_{k=1,\ldots,D} ||\mathbf{c}_{x,y}^{\rho_k}||_2$$

Disparity Estimation

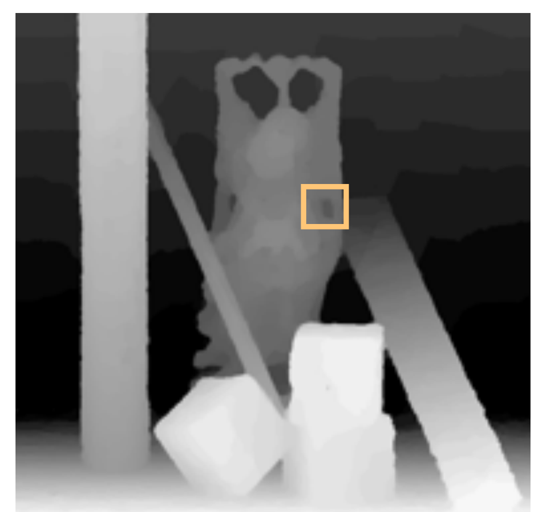
- Neighbor disparity patches are smooth
- Vector-valued isotropic total variation for smooth disparity

$$\|\nabla \mathbf{c}\|_{2,1} \doteq \sum_{x,y} \sqrt{\|\mathbf{c}_{x,y} - \mathbf{c}_{x+1,y}\|_2^2 + \|\mathbf{c}_{x,y} - \mathbf{c}_{x,y+1}\|_2^2}$$

Disparity Estimation

- Neighbor disparity patches are smooth
- Vector-valued isotropic total variation for smooth disparity

$$\|\nabla \mathbf{c}\|_{2,1} \doteq \sum_{x,y} \sqrt{\|\mathbf{c}_{x,y} - \mathbf{c}_{x+1,y}\|_2^2 + \|\mathbf{c}_{x,y} - \mathbf{c}_{x,y+1}\|_2^2}$$



Disparity Estimation

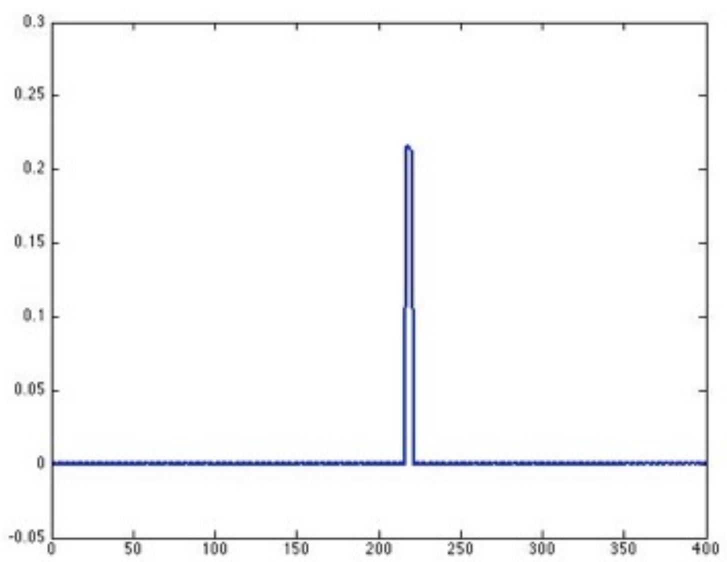
- Neighbor disparity patches are smooth
- Vector-valued isotropic total variation for smooth disparity

$$\|\nabla \mathbf{c}\|_{2,1} \doteq \sum_{x,y} \sqrt{\|\mathbf{c}_{x,y} - \mathbf{c}_{x+1,y}\|_2^2 + \|\mathbf{c}_{x,y} - \mathbf{c}_{x,y+1}\|_2^2}$$

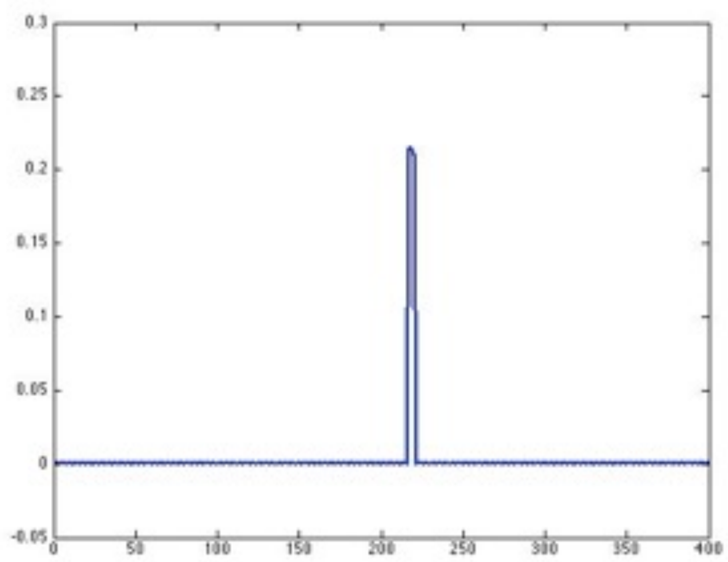
Disparity Estimation

- Neighbor disparity patches are smooth
- Vector-valued isotropic total variation for smooth disparity

$$\|\nabla \mathbf{c}\|_{2,1} \doteq \sum_{x,y} \sqrt{\|\mathbf{c}_{x,y} - \mathbf{c}_{x+1,y}\|_2^2 + \|\mathbf{c}_{x,y} - \mathbf{c}_{x,y+1}\|_2^2}$$



$\mathbf{c}_{x,y+1}$



$\mathbf{c}_{x,y}$

Disparity Estimation

- Disparity estimation reads

$$\min_{\mathbf{c},\mathbf{e}} \frac{1}{2}\|\mathbf{y} - \mathbf{Q}\mathbf{c} - \mathbf{e}\|_2^2 + \mu\|\mathbf{e}\|_1 + \lambda\|\nabla \mathbf{c}\|_{2,1} + \gamma\|\mathbf{c}\|_{2,1}$$

Sparse noise

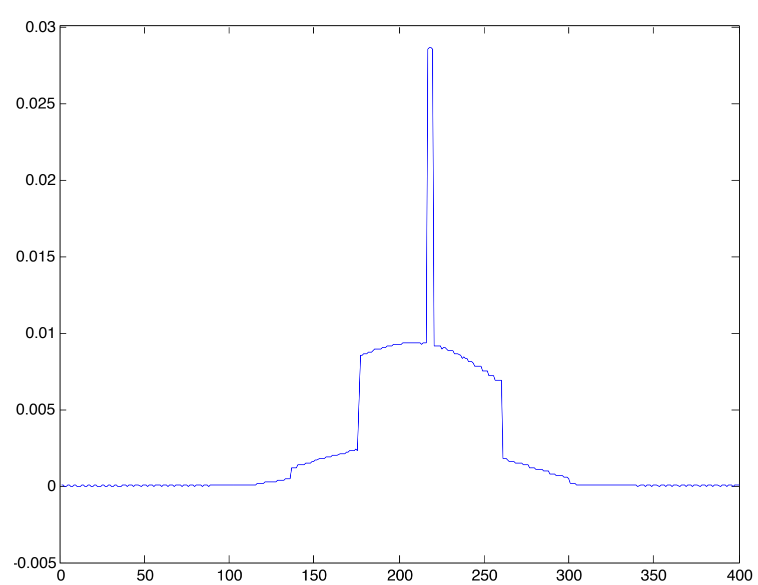
Smooth disparity

Single disparity block

Disparity Estimation

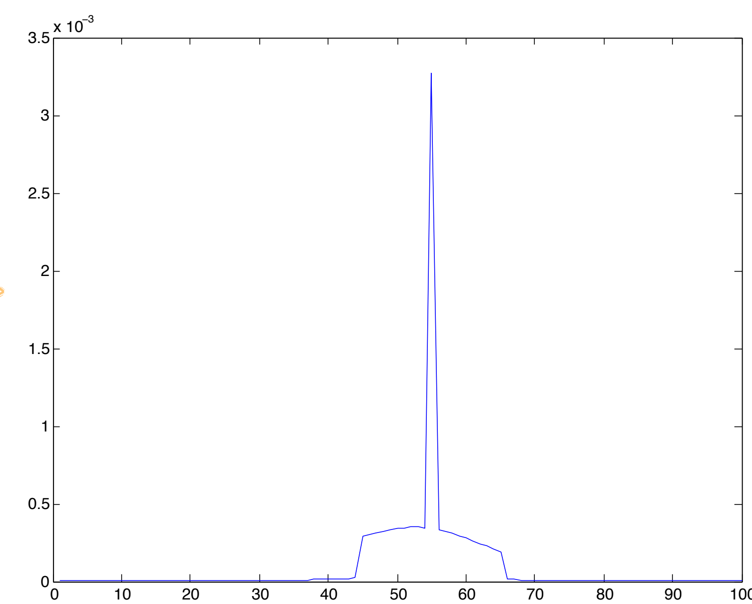
- Disparity of each patch:
  - Estimate the patch unknown weights: $\mathbf{c}_{x,y}$
  - Disparity → hypothesis with the highest norm

$$\hat{\rho} = \operatorname{argmin}_{\rho \in \{\rho_1, \dots, \rho_D\}} \; ||\mathbf{c}_{x,y}||_2$$



$\mathbf{c}_{x,y}$

Norm of patch weights per group →



Group norm

Primal-dual Formulation

- Combining all unknowns $\mathbf{c}$ and $\mathbf{e}$ into a single variable $\mathbf{x}$

$$F_1(\mathbf{Ax}-\mathbf{y}) \doteq \frac{1}{2}\|\mathbf{y}-\mathbf{Qc}-\mathbf{e}\|_2^2$$

$$F_2(\Pi_{\mathbf{e}}\mathbf{x}) \doteq \|\mathbf{e}\|_1$$

$$F_3(\mathbf{Bx}) \doteq \|\nabla\mathbf{c}\|_{2,1} + \frac{\gamma}{\lambda}\|\mathbf{c}\|_{2,1}$$

- $\mathbf{I}_d$: Identity matrix
- $\mathbf{A} \doteq [\mathbf{Q}\ \mathbf{I}_d]$
- $\Pi_{\mathbf{e}}\mathbf{x} \doteq \mathbf{e}$
- $\mathbf{B} \doteq [\nabla^\top\ \frac{\gamma}{\lambda}\mathbf{I}_d]^\top \Pi_{\mathbf{c}}$
  $\Pi_{\mathbf{c}}\mathbf{x} \doteq \mathbf{c}$

$$\min_{\mathbf{x}} F_1(\mathbf{Ax}-\mathbf{y}) + \mu F_2(\Pi_{\mathbf{e}}\mathbf{x}) + \lambda F_3(\mathbf{Bx})$$

Primal-dual Formulation

- The gradient of the cost function requires dealing with $\mathbf{A}^\top \mathbf{A}$

$$\frac{\partial}{\partial \mathbf{x}} F_1(\mathbf{A}\mathbf{x} - \mathbf{y}) = \mathbf{A}^\top \mathbf{A}\mathbf{x} - \mathbf{A}^\top \mathbf{y}$$

- $\mathbf{B}$ is not tight
- The conjugate functions are simple → Primal-dual formulation [1]

$$\min_{\mathbf{x}} \max_{\mathbf{z}} < K\mathbf{x}, \mathbf{z} > -\hat{F}(\mathbf{z})$$

$$\mathbf{z} \doteq [\mathbf{z}_1^\top \ \mathbf{z}_2^\top \ \mathbf{z}_3^\top]^\top$$

$$\mathbf{K} \doteq [\mathbf{A}^\top \ \mu\Pi_{\mathbf{e}}^\top \ \lambda\mathbf{B}^\top]^\top$$

$$\hat{F}(\mathbf{z}) \doteq F_1^*(\mathbf{z}_1) + \mu F_2^*(\mathbf{z}_2) + \lambda F_3^*(\mathbf{z}_3)$$

[1] A. Chambolle et al., “A First Order Primal-dual Algorithm for Convex Problems with Applications to Imaging”, J of Math Imaging andVision 13

Primal-dual Algorithm

- Parallel steps and simple calculation per patch

$$\mathbf{z}_1^{n+1} = \text{prox}_{\sigma F_1^*}\left(\mathbf{z}_1^n + \sigma(\mathbf{A}\bar{\mathbf{x}}^n - \mathbf{y})\right) \qquad \left(\mathbf{z}_1^n + \sigma(\mathbf{A}\bar{\mathbf{x}}^n - \mathbf{y})\right)/(\sigma+1)$$

$$\mathbf{z}_2^{n+1} = \text{prox}_{\sigma\mu F_2^*}\left(\mathbf{z}_2^n + \sigma\mu\Pi_{\mathbf{e}}\bar{\mathbf{x}}^n\right) \qquad \mathcal{H}_{\sigma\mu}\left(\{\mathbf{z}_2^n/(\sigma\mu) + \Pi_{\mathbf{e}}\bar{\mathbf{x}}^n\}_s\right)$$

$$\mathbf{z}_3^{n+1} = \text{prox}_{\sigma\lambda F_3^*}\left(\mathbf{z}^n + \sigma\lambda\mathbf{B}\bar{\mathbf{x}}^n\right) \qquad \{\mathbf{z}_3^n + \sigma\lambda\mathbf{B}\bar{\mathbf{x}}^n\}_b\left(1 - \max\left\{0, 1 - \frac{1}{\|\{\mathbf{z}_3^n + \sigma\lambda\mathbf{B}\bar{\mathbf{x}}^n\}_b\|_2}\right\}\right)$$

$$\mathbf{x}^{n+1} = \mathbf{x}^n - \tau\mathbf{K}^\top\mathbf{z}^{n+1}$$

$$\bar{\mathbf{x}}^{n+1} = \mathbf{x}^{n+1} + \theta(\mathbf{x}^{n+1} - \mathbf{x}^n)$$

Primal-dual Algorithm

- Parallel steps and simple calculation per patch

$$\mathbf{z}_1^{n+1} = \mathrm{prox}_{\sigma F_1^*}\left(\mathbf{z}_1^n + \sigma(\mathbf{A}\bar{\mathbf{x}}^n - \mathbf{y})\right) \longrightarrow \left(\mathbf{z}_1^n + \sigma(\mathbf{A}\bar{\mathbf{x}}^n - \mathbf{y})\right)/(\sigma+1)$$

$$\mathbf{z}_2^{n+1} = \mathrm{prox}_{\sigma\mu F_2^*}\left(\mathbf{z}_2^n + \sigma\mu\Pi_{\mathbf{e}}\bar{\mathbf{x}}^n\right) \longrightarrow \mathcal{H}_{\sigma\mu}\left(\{\mathbf{z}_2^n/(\sigma\mu) + \Pi_{\mathbf{e}}\bar{\mathbf{x}}^n\}_s\right)$$

$$\mathbf{z}_3^{n+1} = \mathrm{prox}_{\sigma\lambda F_3^*}\left(\mathbf{z}^n + \sigma\lambda\mathbf{B}\bar{\mathbf{x}}^n\right) \longrightarrow \{\mathbf{z}_3^n + \sigma\lambda\mathbf{B}\bar{\mathbf{x}}^n\}_b\left(1 - \max\left\{0, 1 - \frac{1}{\|\{\mathbf{z}_3^n + \sigma\lambda\mathbf{B}\bar{\mathbf{x}}^n\}_b\|_2}\right\}\right)$$

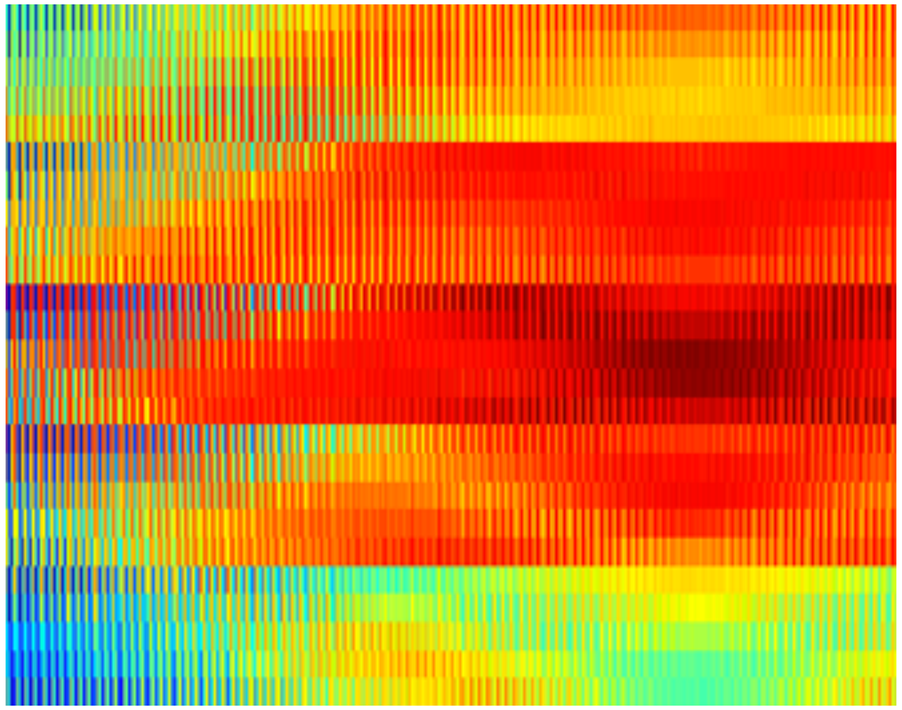
$$\mathbf{x}^{n+1} = \mathbf{x}^n - \tau\mathbf{K}^\top\mathbf{z}^{n+1}$$

$$\bar{\mathbf{x}}^{n+1} = \mathbf{x}^{n+1} + \theta(\mathbf{x}^{n+1} - \mathbf{x}^n)$$

Implementation Detail

- Restrict the possible disparity to a carefully selected subset
- Candidates from $g_{x,y}(\rho)$ and from neighboring pixels

$$g_{x,y}(\rho) = \sum_{i,j} \sum_{p>i,q>j} \Phi(L_{i,j}(x,y) - L_{p,q}(x - \rho\Delta_x(p-i), y - \rho\Delta_y(q-j)))$$

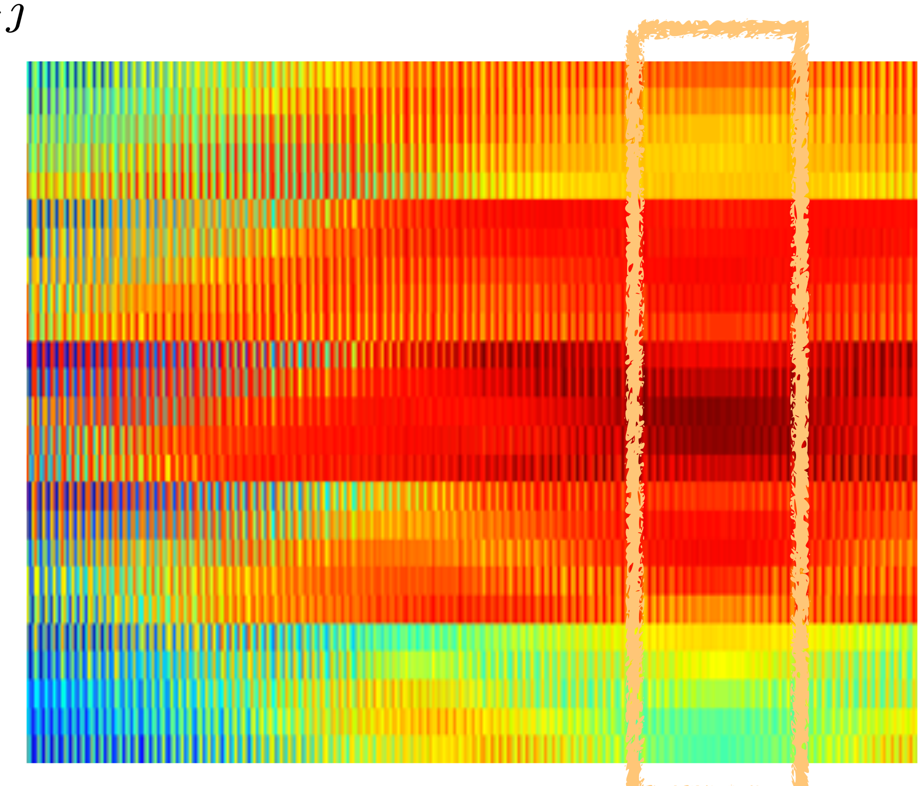
$\mathbf{Q}_{x,y} =$ 

Implementation Detail

- Restrict the possible disparity to a carefully selected subset
- Candidates from $g_{x,y}(\rho)$ and from neighboring pixels

$$g_{x,y}(\rho) = \sum_{i,j} \sum_{p>i,q>j} \Phi(L_{i,j}(x,y) - L_{p,q}(x - \rho\Delta_x(p-i), y - \rho\Delta_y(q-j)))$$

$$\mathbf{Q}_{x,y} =$$

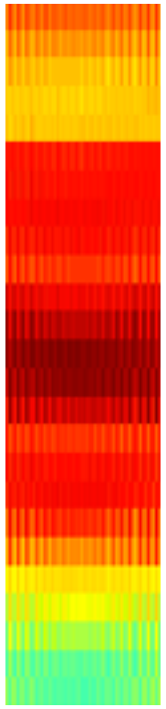


Implementation Detail

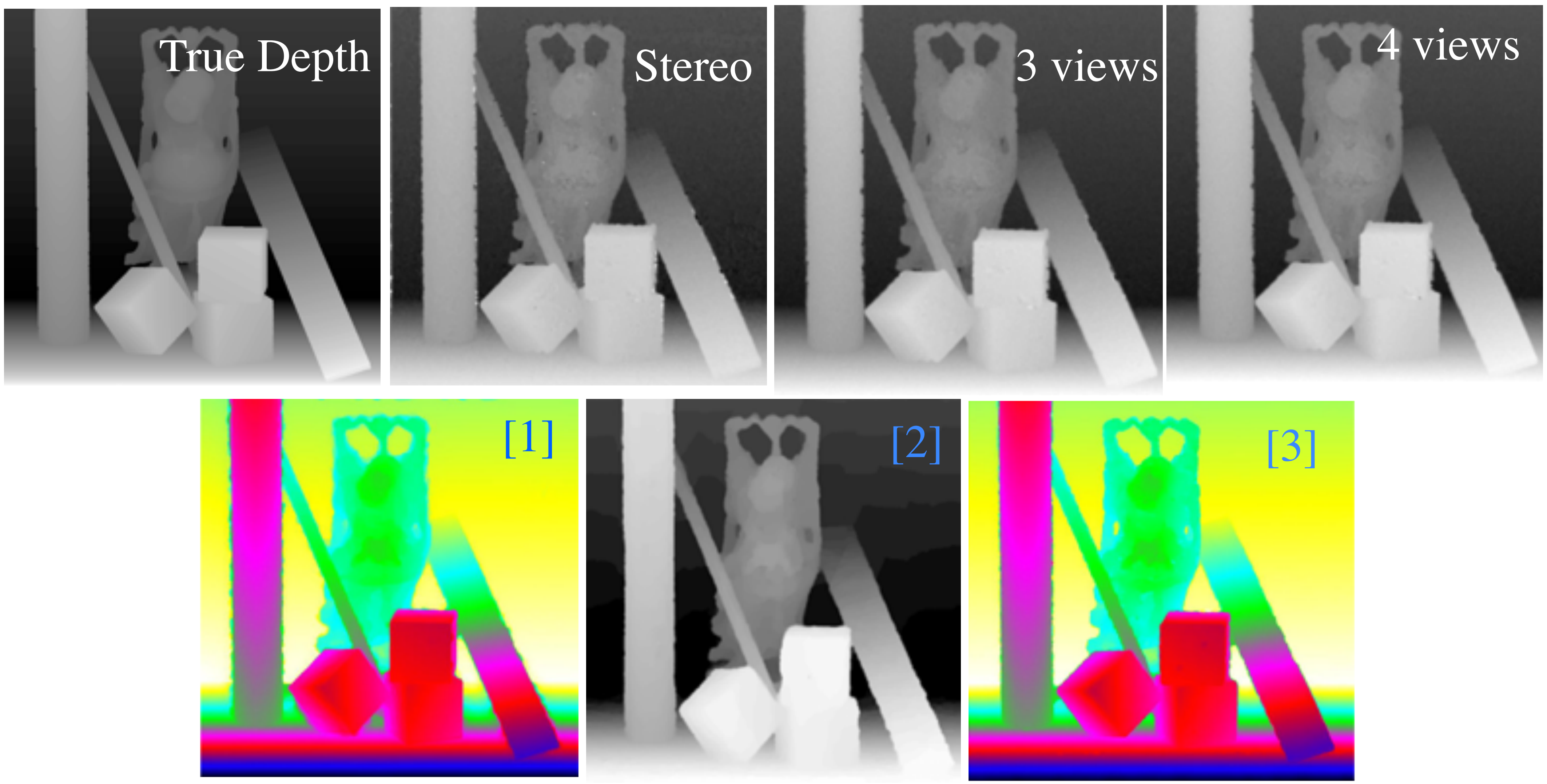
- Restrict the possible disparity to a carefully selected subset
- Candidates from $g_{x,y}(\rho)$ and from neighboring pixels

$$g_{x,y}(\rho) = \sum_{i,j} \sum_{p>i,q>j} \Phi(L_{i,j}(x,y) - L_{p,q}(x - \rho\Delta_x(p-i), y - \rho\Delta_y(q-j)))$$

$$\mathbf{Q}_{x,y} =$$



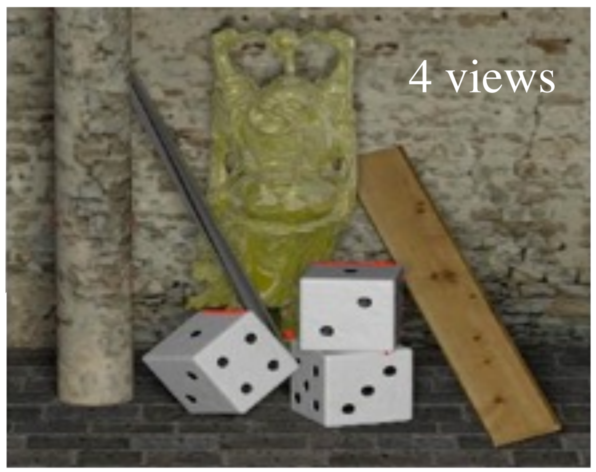
Disparity Results



[1] S. Heber et al., "Variational Shape from Light Field", EMMCVPR 13
[2] S. Wanner et al., "Globally Constant Depth Labeling of 4D Light Fields", CVPR 12
[3] S. Heber et al., "Shape from Light Fields meets Robust PCA", ECCV 14

Disparity Results

- Relative Disparity Error



	4 views			3 views			stereo			[1]			[2]		[3]
Rel error	1%	0.5%	0.2%	1%	0.5%	0.2%	1%	0.5%	0.2%	1%	0.5%	0.2%	1%	0.2%	0.2%
% of pixels	0.13	0.33	1.9	0.139	0.33	1.99	0.42	0.85	3.26	2.12	3.91	8.37	2.9	7.28	5.03

[1] S. Heber et al., "Variational Shape from Light Field", EMMCVPR 13
[2] S. Wanner et al., "Globally Constant Depth Labeling of 4D Light Fields", CVPR 12
[3] S. Heber et al., "Shape from Light Fields meets Robust PCA", ECCV 14

Disparity Results

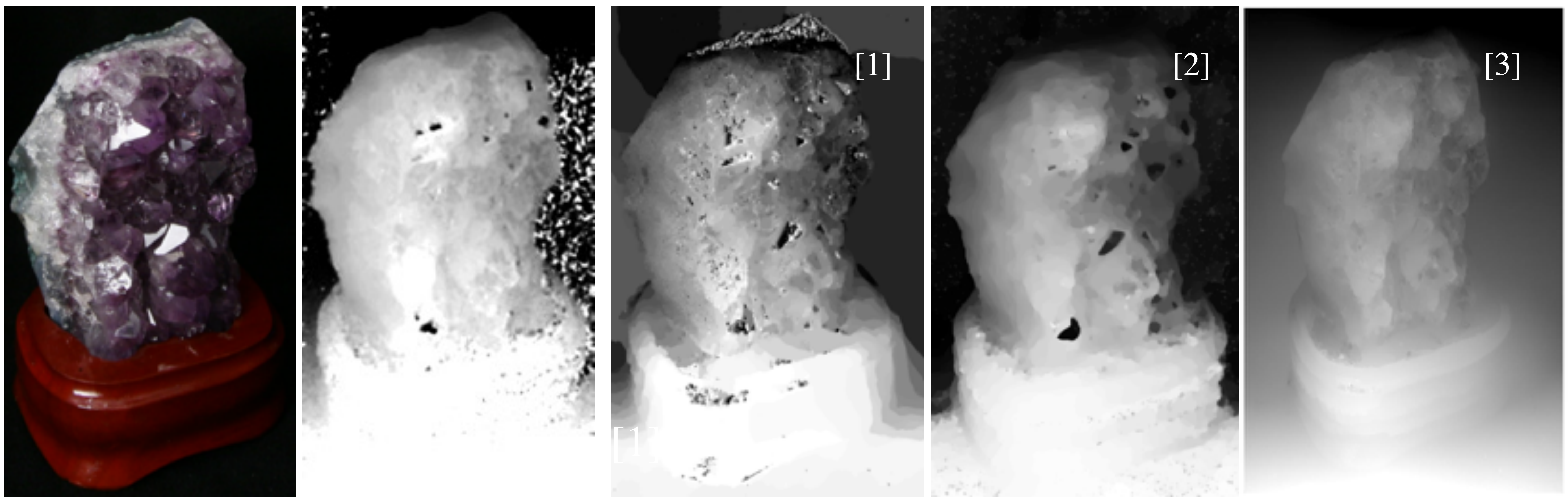
- Relative Disparity Error

	4 views			3 views			stereo			[1]			[2]		[3]
Rel error	1%	0.5%	0.2%	1%	0.5%	0.2%	1%	0.5%	0.2%	1%	0.5%	0.2%	1%	0.2%	0.2%
% of pixels	0.13	0.33	1.9	0.139	0.33	1.99	0.42	0.85	3.26	2.12	3.91	8.37	2.9	7.28	5.03

[1] S. Heber et al., "Variational Shape from Light Field", EMMCVPR 13
[2] S. Wanner et al., "Globally Constant Depth Labeling of 4D Light Fields", CVPR 12
[3] S. Heber et al., "Shape from Light Fields meets Robust PCA", ECCV 14

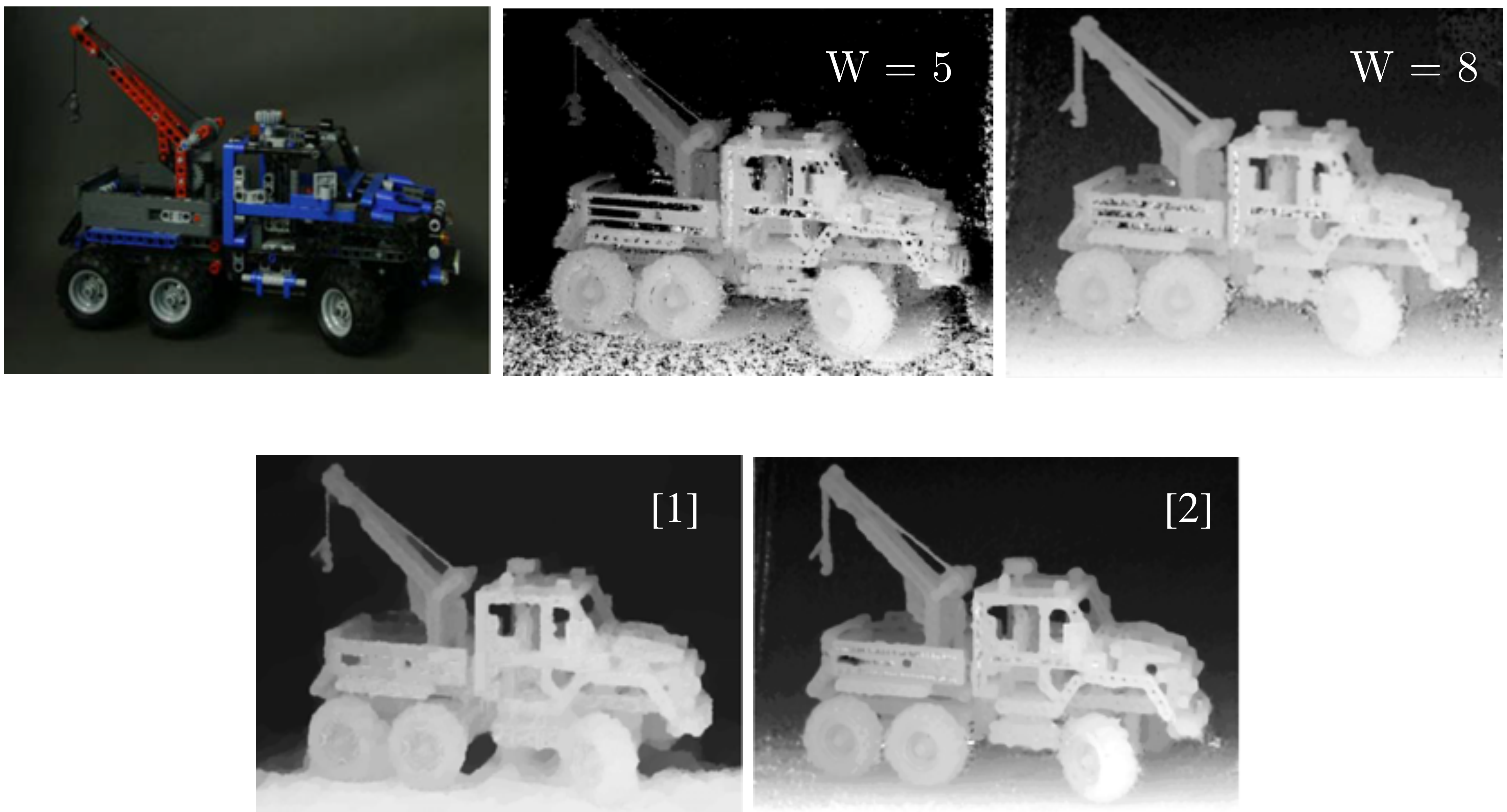
Disparity Results

- Light fields from camera array



[1] S. Heber et al., “Variational Shape from Light Field”, EMMCVPR 13
[2] S. Wanner et al., “Globally Constant Depth Labeling of 4D Light Fields”, CVPR 12
[3] S. Heber et al., “Shape from Light Fields meets Robust PCA”, ECCV 14

Disparity Results



[1] S. Wanner et al., “Globally Constant Depth Labeling of 4D Light Fields”, CVPR 12
[2] C. Kim et al. “Scene Reconstruction from High Spatio-angular Resolution Light Fields”, SIGGRAPH 13

Conclusion

- Leverage redundancy in light fields leads to convex disparity estimation

Conclusion

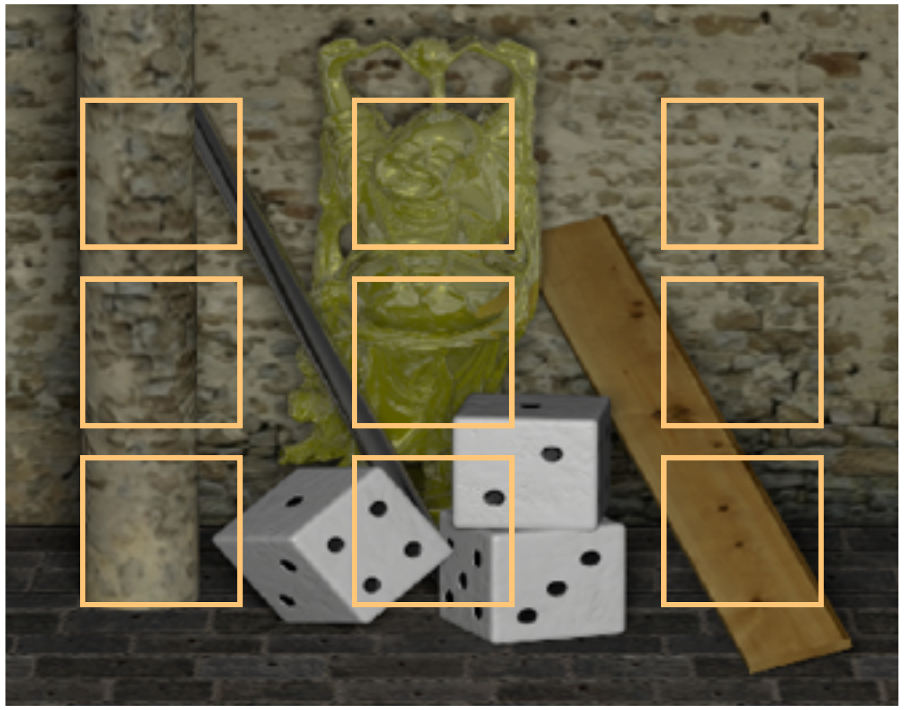
- Leverage redundancy in light fields leads to convex disparity estimation
- Robust occlusions handling

Conclusion

- Leverage redundancy in light fields leads to convex disparity estimation
- Robust occlusions handling
- Disparity estimation using Block Coordinate Descent

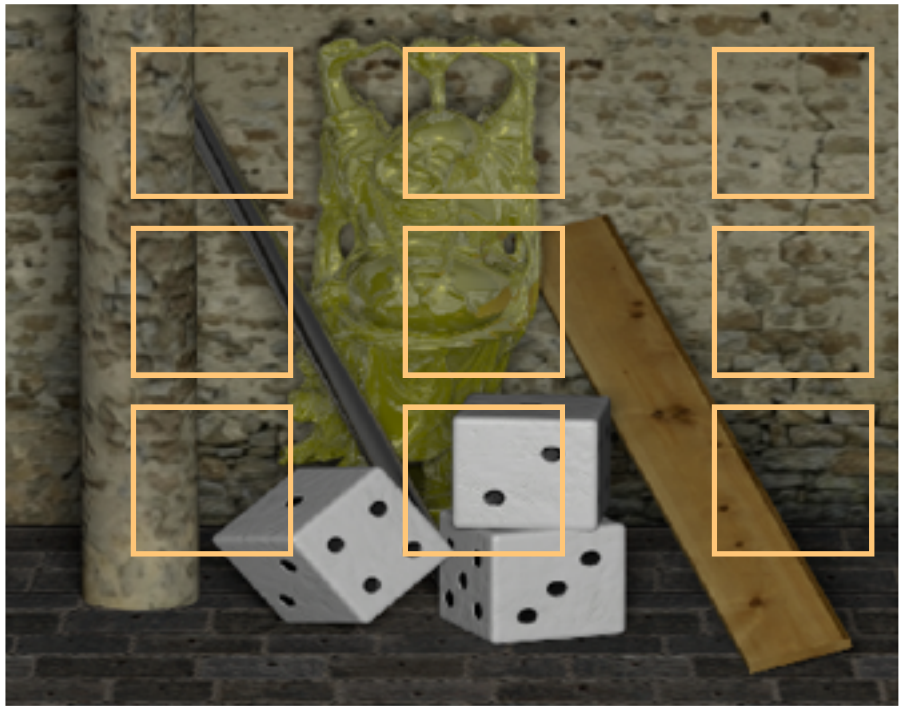
Conclusion

- Leverage redundancy in light fields leads to convex disparity estimation
- Robust occlusions handling
- Disparity estimation using Block Coordinate Descent



Conclusion

- Leverage redundancy in light fields leads to convex disparity estimation
- Robust occlusions handling
- Disparity estimation using Block Coordinate Descent



Conclusion

- Leverage redundancy in light fields leads to convex disparity estimation
- Robust occlusions handling
- Disparity estimation using Block Coordinate Descent

$$\min_{C_{\mathcal{S}_i}, E_{\mathcal{S}_i}} \frac{1}{2}\|Y_{\mathcal{S}_i} - Q_{\mathcal{S}_i}C_{\mathcal{S}_i} - E_{\mathcal{S}_i}\|_2^2 + \mu\|E_{\mathcal{S}_i}\|_1 + \lambda\|\nabla C_{\mathcal{S}_i}\|_{1,2} + \gamma\|C_{\mathcal{S}_i}\|_{1,2}$$

Conclusion

- Leverage redundancy in light fields leads to convex disparity estimation
- Robust occlusions handling
- Disparity estimation using Block Coordinate Descent