Light-field Image Compression Based on Variational Disparity Estimation and Motion-Compensated Wavelet Decomposition

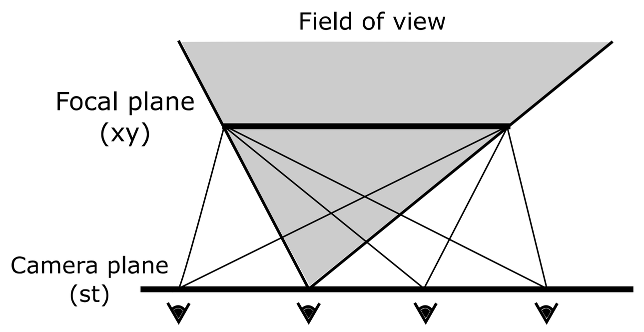
Trung-Hieu Tran*[1], Yousef Baroud[1], Zhe Wang[1], Sven Simon[1], David Taubman[2]
[1] The University of Stuttgart, Germany   [2] The University of New South Wales, Australia
*Email: trung.hieu.tran@ipvs.uni-stuttgart.de

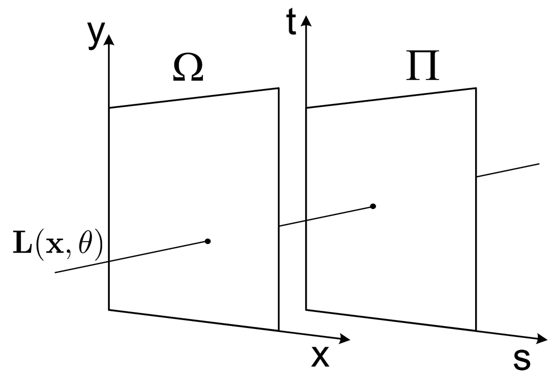
1. Light-field Imaging

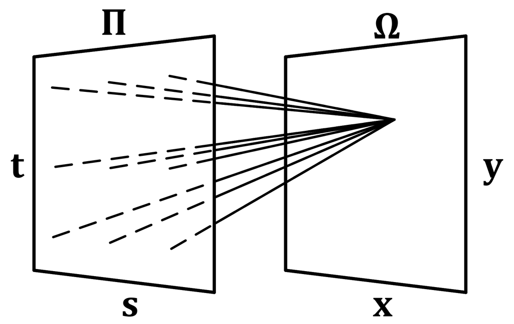
- A collection of pin hole views from view points on plane parallel to the image plane.
- A collection of light rays passing two parallel planes.

$$\mathbf{L} : \Omega \times \Pi \to \mathbb{R}, \quad (\mathbf{x}, \theta) \to \mathbf{L}(\mathbf{x}, \theta)$$

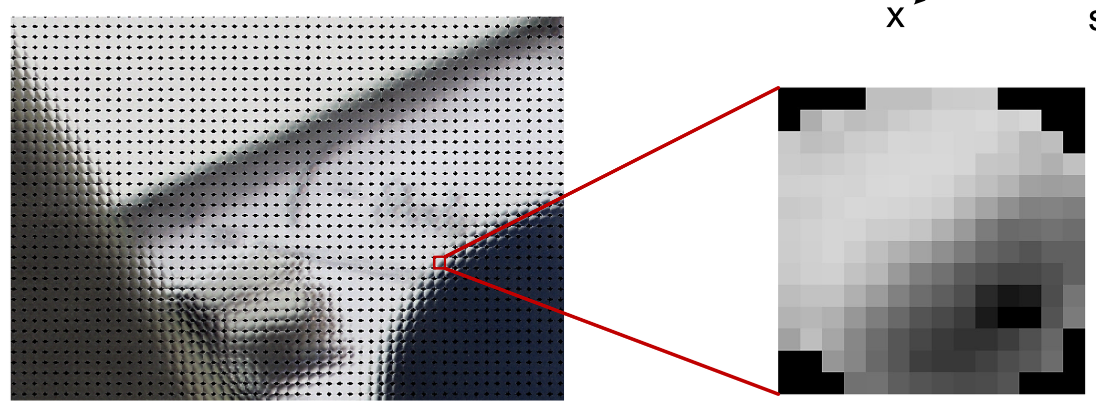
- $\mathbf{x} = (x, y)^T \in \Omega$
- $\theta = (s, t)^T \in \Pi$
- $\Omega, \Pi \subset \mathbb{R}^2$



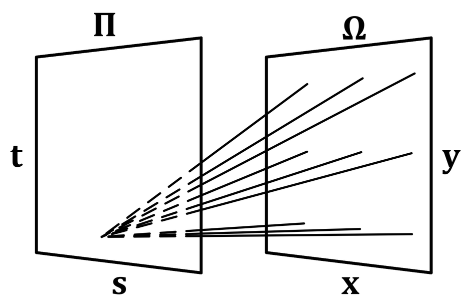




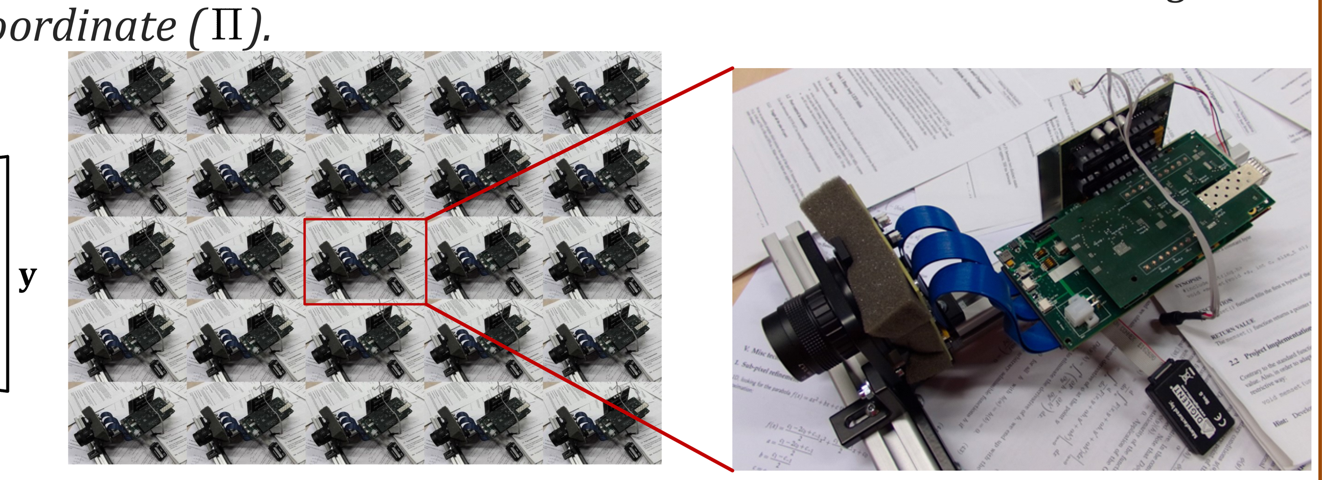
Fixing spatial coordinate (Ω), varying directional coordinate (Π).



Microlens image

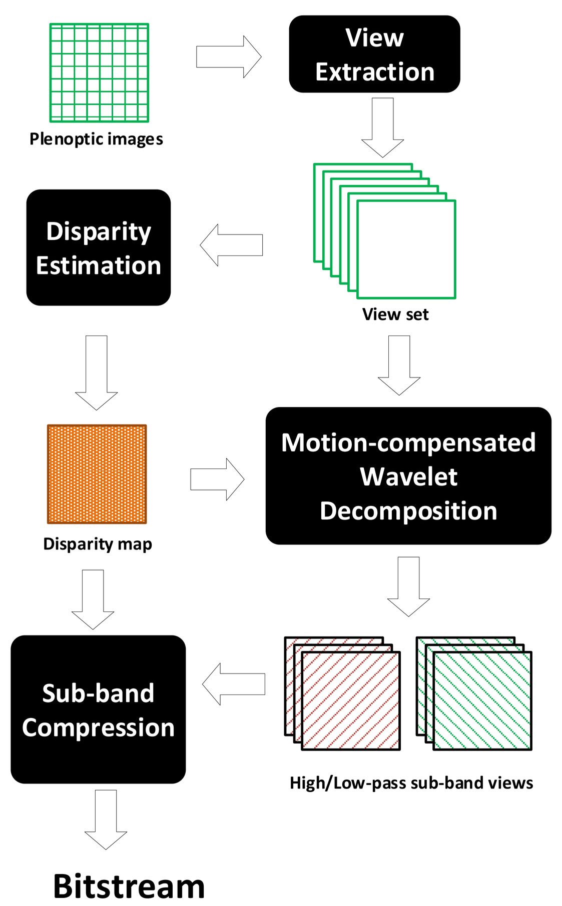


Varying spatial coordinate (Ω), fixing directional coordinate (Π).



Sub-aperture image

2. Proposed Coding Framework



- Compression of sub-aperture images.
- Estimate disparity map and use it in a motion-compensated wavelet decomposition scheme.
- Support both Lossless and Lossy compression.

View Extraction
- Decodes lenslet images into sub-aperture images.

Disparity Estimation
- Estimate disparity map using variational framework.

Motion-compensated Wavelet Decomposition
- 2-step lifting scheme with reversible wavelet transform.

Sub-band Compression
- JPEG2000 coding of high/low-pass sub-band view.

3. Disparity Estimation Framework

Plenoptic Variational Optimization Framework [1]

$$\underset{\omega}{\operatorname{argmin}}\, E(\omega) = \int_{\Omega} \mathbf{D}(\mathbf{x}, \omega) + \alpha \mathbf{S}(\mathbf{x}, \omega) d\mathbf{x}$$

Data Term

- Intensity Constancy Assumption
  $F_{g,i}(\mathbf{x}, \omega) = \mathbf{L}(\mathbf{x}, \theta_i) - \mathbf{L}(\mathbf{x} + \theta_i \omega, \theta_0) = 0$
- Gradient Constancy Assumption
  $F_{G,i}(\mathbf{x}, \omega) = \nabla \mathbf{L}(\mathbf{x}, \theta_i) - \nabla \mathbf{L}(\mathbf{x} + \theta_i \omega, \theta_0) = 0$
- RGB Color space

$$D(\mathbf{x}, \omega) = \sum_{\theta_i \in \Pi} \Psi\Big(\sum_{c \in \{RBG\}} F_{g,i}(\mathbf{x}, \omega)^2 \Big) + \gamma \sum_{\theta_i \in \Pi} \Psi\Big(\sum_{c \in \{RBG\}} F_{G,i}(\mathbf{x}, \omega)^2 \Big)$$

- HSV Color space

$$D(\mathbf{x}, \omega) = \sum_{\theta_i \in \Pi} \sum_{c \in \{HSV\}} \Psi(F_{g,i}(\mathbf{x}, \omega)^2) + \gamma \sum_{\theta_i \in \Pi} \sum_{c \in \{HSV\}} \Psi(F_{G,i}(\mathbf{x}, \omega)^2)$$

- Sub-quadratic function $\Psi(s^2) = \sqrt{s^2 + \epsilon^2}$ for better handling outliers.

Flow-driven Smoothness Term

$$S(\mathbf{x}, \omega) = \Psi\left(|\nabla \omega|^2\right)$$

- Piecewise smoothness and better handle outliers.

References

[1] Trung-Hieu Tran, Zhe Wang, and Sven Simon, "Variational Disparity Estimation Framework for Plenoptic Images," in 2017 IEEE International Conference on Multimedia & Expo Workshops (ICMEW). Jul 2017, pp. 1–6, IEEE.

[2] Wanner, S., Meister, S., & Goldluecke, B. (2013). Datasets and Benchmarks for Densely Sampled 4D Light Fields. Vision, Modeling & Visualization, 225–226.

[3] "JPEG Pleno Database: EPFL Light-field data set," http://jpeg.org/plenodb/lf/epfl/

4. Motion-compensated Wavelet Decomposition

- 2D wavelet transform and motion compensation are jointly performed.
- Reversible integer-to-integer wavelet transforms for lossless and lossy compression.
- Deploy a two-step lifting scheme based on Haar and bi-orthogonal 5/3 wavelet kernel.

1D Forward transform

$$\mathrm{h}_k(\mathbf{x}) = \mathrm{V}_{2k+1}(\mathbf{x}) + \alpha_1 \mathcal{W}_{2k,2k+1}[\mathrm{V}_{2k}(\mathbf{x})] + \alpha_2 \mathcal{W}_{2k+2,2k+1}[\mathrm{V}_{2k+2}(\mathbf{x})]$$

$$\mathrm{l}_k(\mathbf{x}) = \mathrm{V}_{2k}(\mathbf{x}) + \beta_1 \mathcal{W}_{2k+1,2k}[\mathrm{h}_k(\mathbf{x})] + \beta_2 \mathcal{W}_{2k-1,2k}[\mathrm{h}_{k-1}(\mathbf{x})]$$

1D Inverse transform

$$\mathrm{V}_{2k+1}(\mathbf{x}) = \mathrm{h}_k(\mathbf{x}) - \alpha_1 \mathcal{W}_{2k,2k+1}[\mathrm{V}_{2k}(\mathbf{x})] - \alpha_2 \mathcal{W}_{2k+2,2k+1}[\mathrm{V}_{2k+2}(\mathbf{x})]$$

$$\mathrm{V}_{2k}(\mathbf{x}) = \mathrm{l}_k(\mathbf{x}) - \beta_1 \mathcal{W}_{2k+1,2k}[\mathrm{h}_k(\mathbf{x})] - \beta_2 \mathcal{W}_{2k-1,2k}[\mathrm{h}_{k-1}(\mathbf{x})]$$

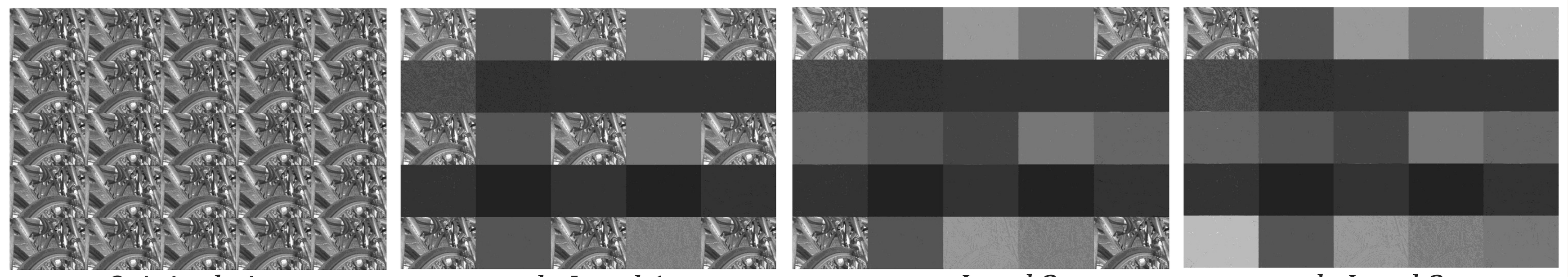
2D transform

Apply 1D transform horizontally and then vertically.

- $\mathrm{h}_k(\mathbf{x})$ *High-pass sub-band view*
- $\mathrm{l}_k(\mathbf{x})$ *Low-pass sub-band view*
- $\mathrm{V}_i(\mathbf{x})$ *A view indexed either horizontally or vertically*
- $\mathcal{W}_{i,j}[*]$ *Wrapping from the coordinate of view i to coordinate of view j*

	Haar	5/3
α_1	-1	- ½
α_2	0	- ½
β_1	½	¼
β_2	0	¼

Wavelet decomposition parameters



a. Original view set *b. Level 1* *c. Level 2* *d. Level 3*

Demonstration of 3-level MWD for a 5x5 view set of "bikes" lenslet data [3]

5. Experimental Results

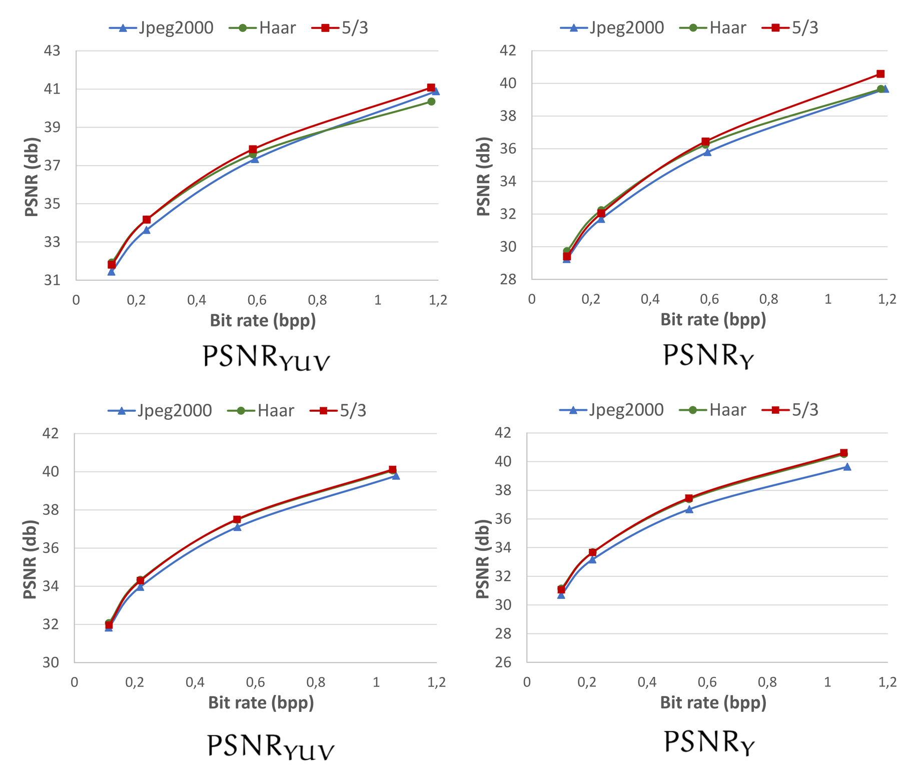
- Running coding framework on both Real-world [3] and Synthetic Light-field [2] data.
- Compare with JPEG-LS and direct application of JPEG2000.

Lossless Compression
- 25% and 5.5 % better than Jpeg-LS and JPEG2000 for synthetic data.
- 11% and 7.3 % better than Jpeg-LS and JPEG2000 for Real-world data.

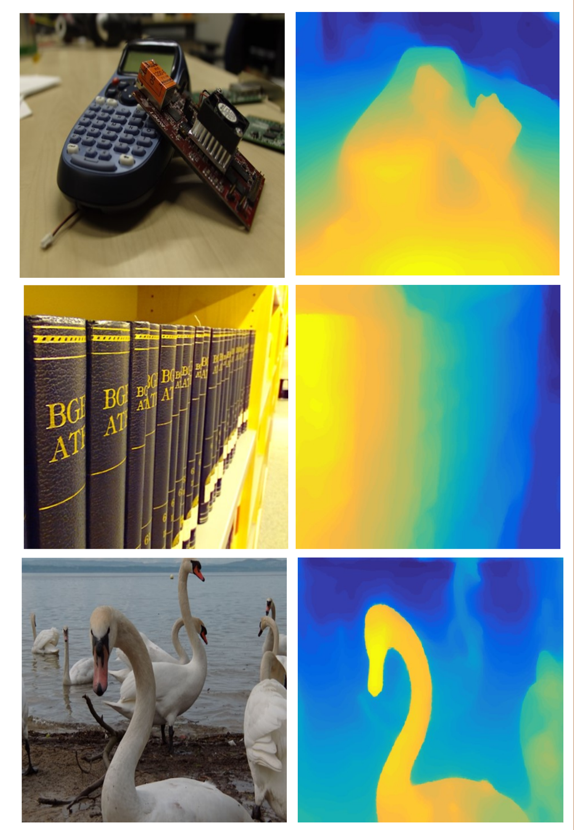
Lossy Compression
- 5/3 wavelet kernel consistently provides a better compression quality compared to JPEG2000 in both synthetic and real-world dataset.

	Image	Raw	JpegLS	Jpeg2000	Haar	5/3
synthetic	buddha2	143,33	90,38	62,42	58,04	57,18
	mona	143,33	77,44	52,16	40,99	40,82
	papillon	143,33	70,17	51,57	44,47	44
	stillLife	143,33	113,86	79,59	67,79	66,5
	horses	143,33	96,41	60,8	59,93	58,65
real-world	books	183,09	93,84	103,74	89,18	86,19
	bikes	183,09	106,73	92,14	84,72	78,73
	danger	183,09	105,42	90,41	80,83	74,68
	pillars	183,09	91,48	76,87	71,21	66,74
	swans1	183,09	95,1	88,45	80,39	75,02

Lossless compressed file-size in mega bytes (MB)



Average PSNR values for different compression ratios on Synthetic data [2] (top) and Real-world data [3] (bottom)



Estimated disparity maps (right) from Light-field images (left)