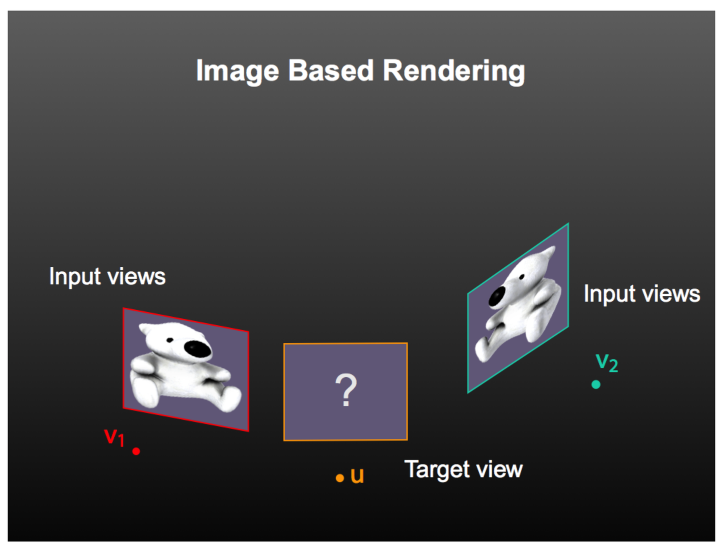
Bayesian View Synthesis and Image-Based Rendering Principles

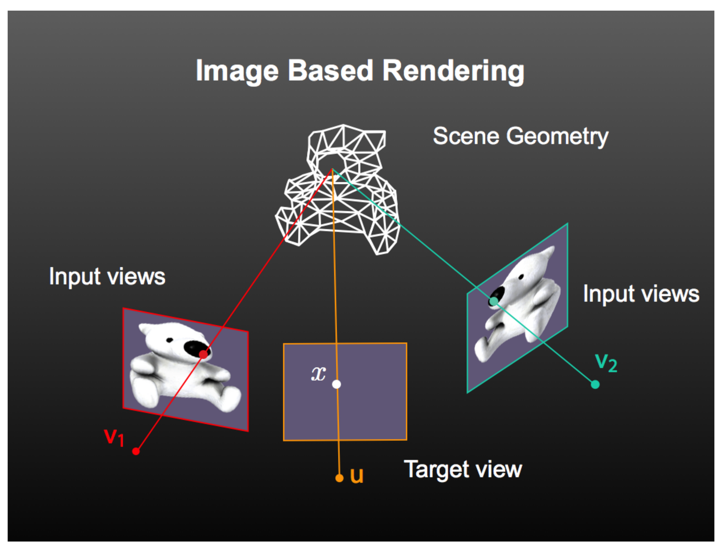
Sergi Pujades[1] - Frédéric Devernay[1] - Bastian Goldluecke[2]

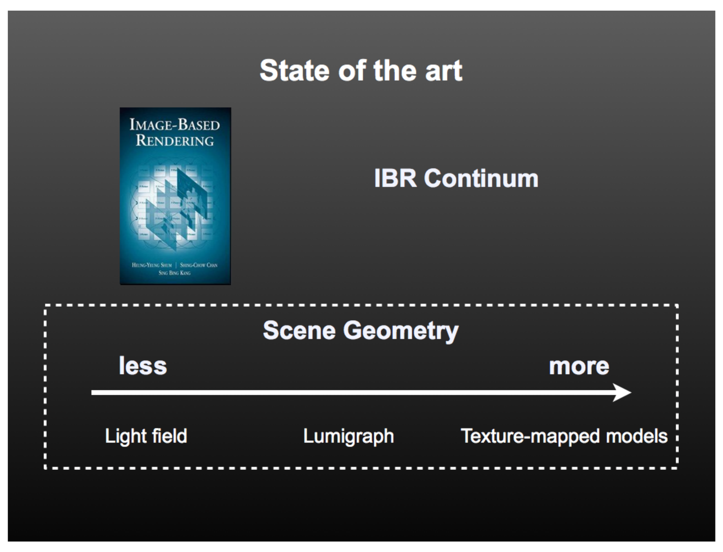
[1] Inria - PRIMA Team, Univ. Grenoble Alpes, LIG, F-38000 Grenoble, France, CNRS, LIG, F-38000 Grenoble, France.

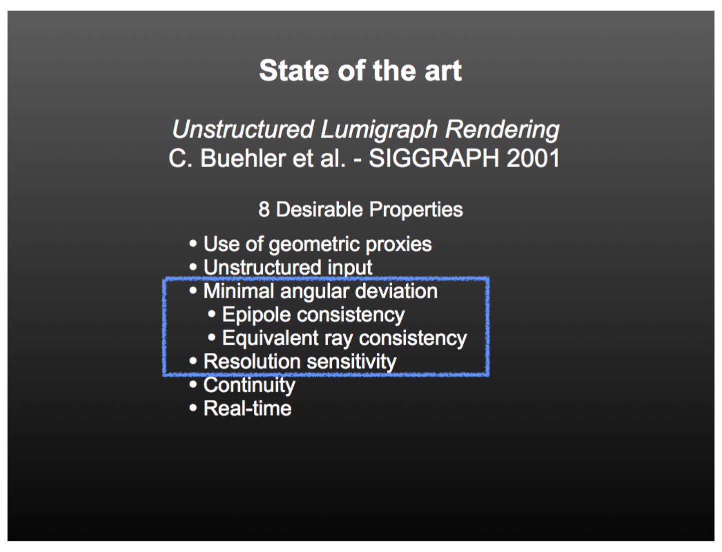
[2] University of Konstanz

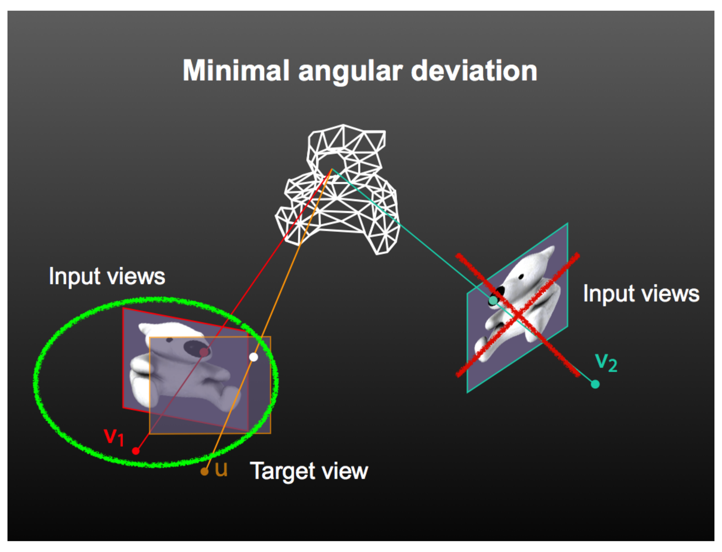
Image-Based Rendering

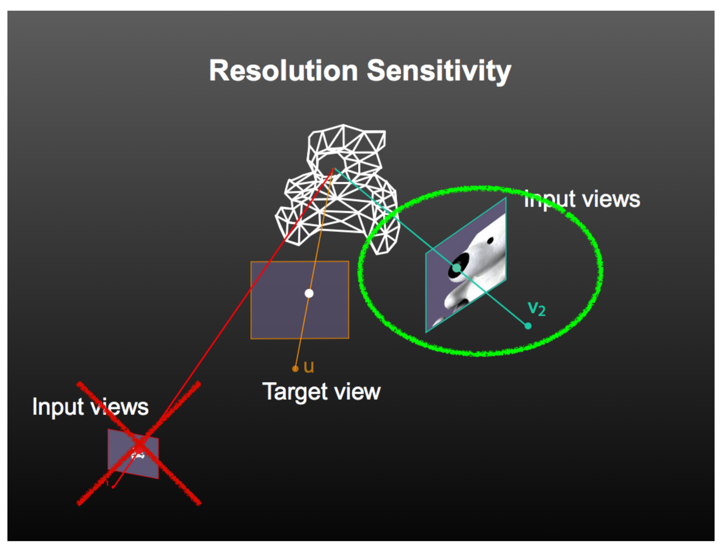


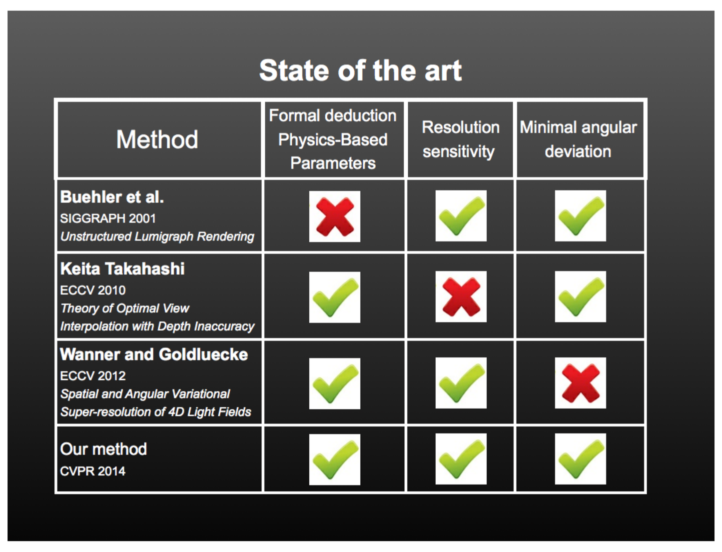




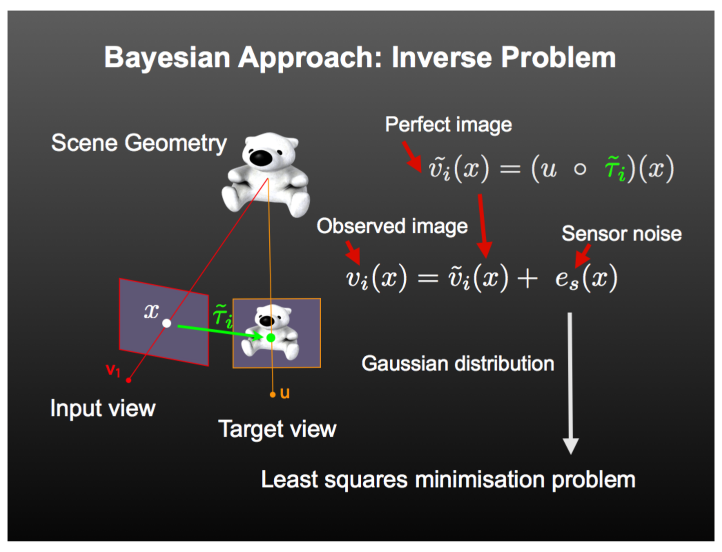


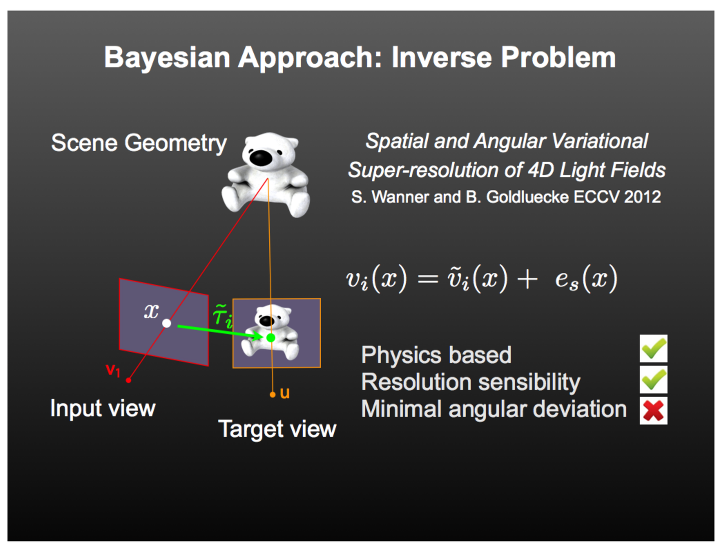


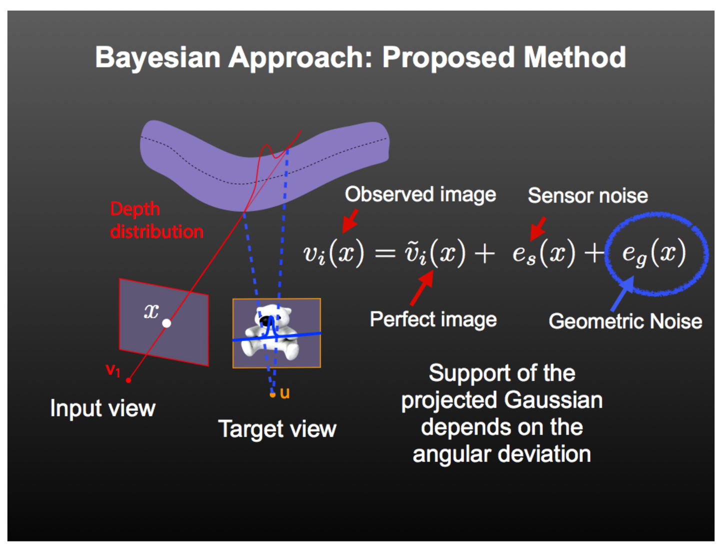


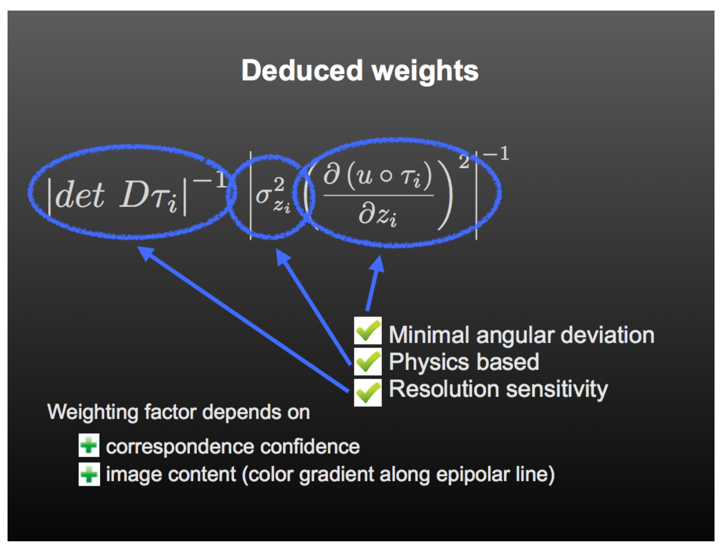


Previous and Proposed Generative Model









Equations

Notation

Target image: $u : \Gamma \to \mathbb{R}$ (or $\mathbb{R}^3$)

Input images: $v_i : \Omega_i \to \mathbb{R}$ (or $\mathbb{R}^3$)

Backwards warp: $\tau_i : \Omega_i \to \Gamma$

Binary occlusion mask: $m_i : \Omega_i \to \{0,1\}$

Visibility set: $V_i \subset \Omega_i \mid m_i = 1$

Forward warp: $\beta_i : \tau_i(V_i) \to \Omega_i \mid \beta_i \circ \tau_i = Id$

Maximum a Posteriori energy

$$E(u) = E_{\text{data}}(u) + \lambda E_{\text{prior}}(u)$$

$$E_{\text{data}}(u) = \sum_{i=1}^{n} \frac{1}{2} \int_{\Omega_i} \omega_i(u)\, m_i (b * (u \circ \tau_i) - v_i)^2 \, dx,$$

$$\text{with } \omega_i(u) = \left(\sigma_s^2 + \sigma_{g_i}^2\right)^{-1}$$

$$E_{\text{prior}}(u) = \int_{\Gamma} |Du|$$

Relation between image error on geometric error

$$\sigma_{g_i} = \sigma_{z_i} \left| b * \frac{\partial (u \circ \tau_i)}{\partial z_i} \right| = \sigma_{z_i} \left| b * \left((\nabla u \circ \tau_i) \cdot \frac{\partial \tau_i}{\partial z_i} \right) \right|$$

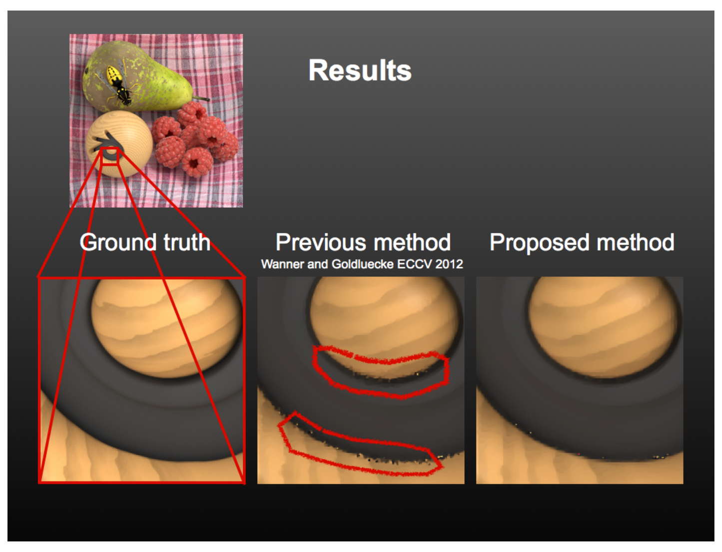
Minimization with a re-weighted iterative method

- Use estimate $\tilde{u}$ of u and consider $\omega_i(\tilde{u})$ constant during iteration: the simplified energy is convex.
- Minimize simplified energy using FISTA.
- Update weights with current solution and iterate.

First iteration: $\tilde{u} = \frac{1}{n} \sum v_i \circ \beta_i$

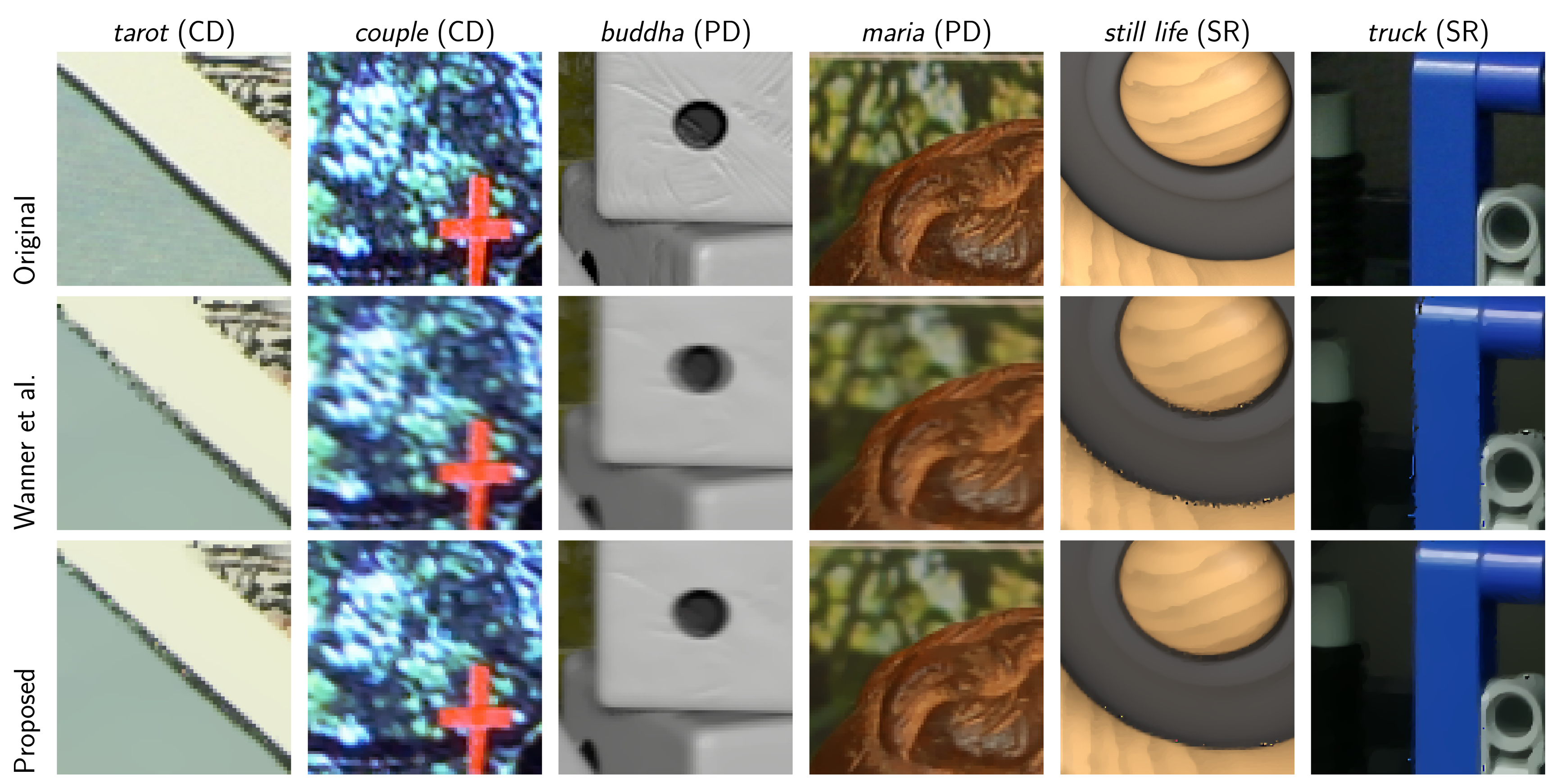
Code available at
http://sourceforge.net/projects/cocolib/

Example for Intuition



Numerical and Visual Results

	HCI light fields, raytraced				HCI light fields, gantry				Stanford light fields, gantry					
	still life		*buddha*		*maria*		*couple*		*truck*		*gum nuts*		*tarot*	
Estimated disparity														
Wanner et al. ECCV 2012	30.13	58	**42.84**	**17**	40.06	53	26.55	226	33.75	408	31.82	1439	28.71	60
Proposed	**30.45**	**55**	42.37	18	**40.10**	53	**28.50**	**178**	**33.78**	**407**	**31.93**	**1437**	**28.88**	**58**
Planar disparity														
Wanner et al. ECCV 2012	21.28	430	34.28	74	31.65	144	20.07	725	32.48	419	30.55	1403	22.64	278
Proposed	**22.24**	**380**	**37.51**	**44**	**34.38**	**99**	**22.88**	**457**	**33.79**	**386**	**31.30**	**1378**	**23.78**	**218**
Super-resolution														
Wanner et al. ECCV 2012	24.93	230	**34.50**	**122**	35.18	129	**25.54**	**287**	**33.11**	**378**	31.80	1475	**26.66**	**113**
Proposed	**25.12**	**228**	34.44	123	**35.20**	129	25.34	289	33.08	379	**31.89**	**1471**	26.54	117



Top: Numerical results for synthetic and real-world light fields. We compare our method to Wanner et al. ECCV 2012 work with respect to same-resolution view synthesis for estimated disparity and a flat plane proxy, as well as super-resolved view synthesis. For each light field, the first value is the PSNR (bigger is better), the second value is DSSIM in units of 10^{-4} (smaller is better). The better value is highlighted in bold.

Right: Visual comparison of novel views obtained for different light fields. From top to bottom, the rows present closeups of the ground truth images, the results obtained by Wanner et al. ECCV 2012 work, and our results. CD stands for computed disparity, PD for planar disparity and SR for super-resolution. The results obtained by the proposed method are visibly sharper, in particular along color edges.

University of Konstanz

CVPR 2014