



Multi-step joint bilateral depth upsampling

Bram Riemens^a, Om Prakash Gangwal^a, Bart Barenbrug^b, Robert-Paul Berretty^c
^a NXP Research, ^b Philips 3D Solutions, ^c Philips Research
VCIP Conference, January 21, 2009
Contact: bram.riemens@nxp.com



Outline

- ▶ **Introduction**
 - Application context
 - Joint bilateral filtering
- ▶ Image prefiltering and dual resolution depth upsampling
- ▶ Multi-step joint bilateral depth upsampling
- ▶ Computational cost
- ▶ Experimental results
- ▶ Concluding remarks



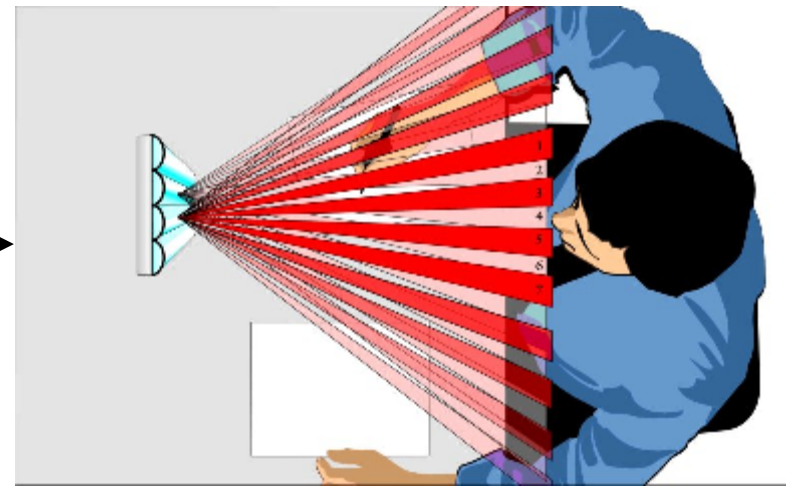
Application context

- ▶ Stereo to multi-view conversion
 - Stereo video input
 - Multi-view display (lenticular technology)

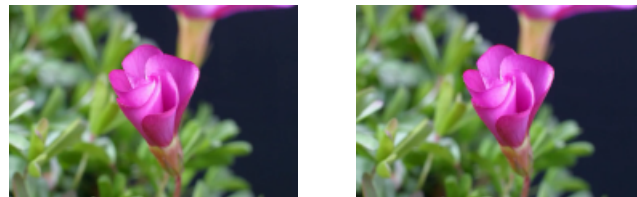
stereo video input



multi-view video output



Application context



Stereo video input:
left and right views

disparity estimation

Based on 8x8 block matching

R-P. M. Berretty, A. K. Riemens, P. F. Machado, "Real-time embedded system for stereo video processing for multiview displays", *Proceedings of SPIE Vol. 6490, 649014 (2007)*



multi-view rendering

Depth dependent pixel shift

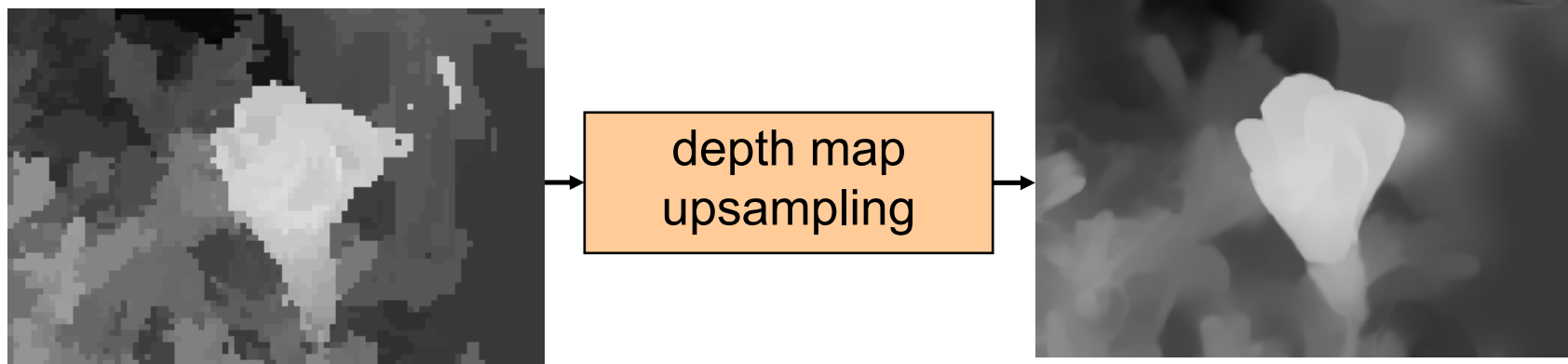
R-P.M. Berretty, F. J. Peters, G. T. J. Volleberg, "Real-time Rendering for Auto-stereoscopic 3D-TV", *Proceedings of ICCE, 2006*



Multi-view video output

Problem definition

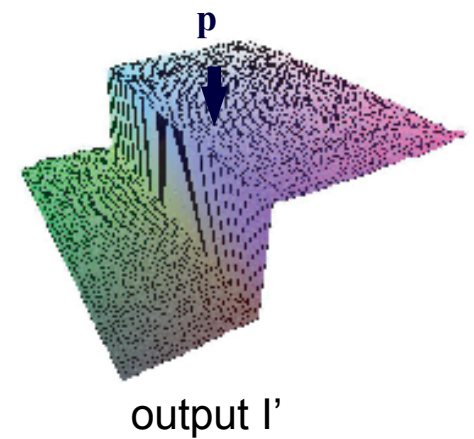
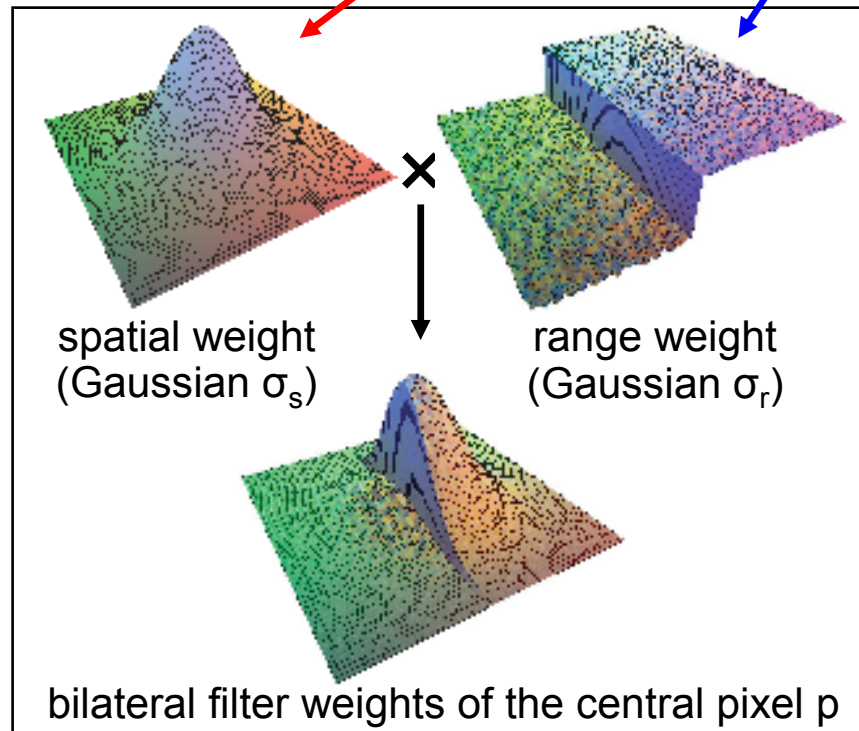
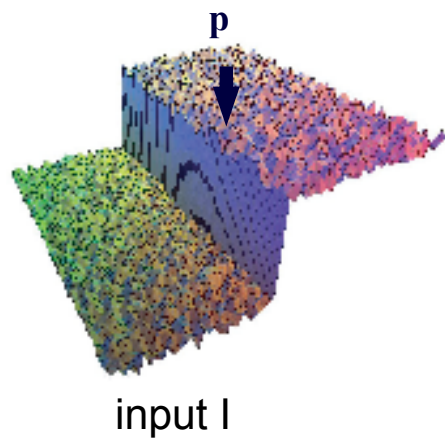
- ▶ Disparity estimation delivers depth map on 8×8 grid
- ▶ Rendering requires full resolution depth map
- ▶ Main objectives / requirements:
 - Cost-effective (potential for real-time at consumer electronics platforms)
 - Align depth map edge with image edge
 - Remove inaccuracies (caused by heuristics in disparity estimation)
 - Temporal stability (esp. at edges and areas with detail)



Bilateral filtering

$$I'_p = \frac{\sum_{q \in S} w_{p,q} I_q}{\sum_{q \in S} w_{p,q}}$$

$$w_{p,q} = s(\| \mathbf{p} - \mathbf{q} \|) r(I_p - I_q)$$



C. Tomasi, R. Manduchi, "Bilateral Filtering for Gray and Color Images", *Proc. ICCV, 1998*
 Paris et al. "A Gentle Introduction to Bilateral Filtering and its Applications", *SigGraph 2007 tutorial*

Joint bilateral filtering

- ▶ Multi-modal filtering
 - Range term defined by one modality
 - Filtering performed on an other modality
- ▶ Propagates properties from one to an other modality
- ▶ Edge preserving properties

depth map filter

use of image properties

$$D'_p = \frac{\sum_{q \in S} w_{p,q} D_q}{\sum_{q \in S} w_{p,q}}$$
$$w_{p,q} = s(\| \mathbf{p} - \mathbf{q} \|) r(I_p - I_q)$$

E. Eisemann, F. Durand, "Flash Photography Enhancement via Intrinsic Relighting", *SigGraph 2004*
G. Petschnigg, et al., "Digital, photography with flash and no-flash image pairs", *SigGraph 2004*

Joint bilateral upsampling

- ▶ First publication on bilateral filters for upsampling at SigGraph 2007
 - Johannes Kopf, Univ. of Konstanz (Germany) provided reference sw.
- ▶ Our solution developed in parallel with Kopf et al.
- ▶ Our proposed solution benchmarked against [Kopf2007]
- ▶ [Kopf2007] solution:
 - High resolution image in range term
 - Low resolution input \rightarrow high resolution output

high resolution result low resolution input high resolution image

$$D_p^h = \frac{\sum_{q \in S} w_{p,q} D_q^l}{\sum_{q \in S} w_{p,q}}$$
$$w_{p,q} = s(\| \mathbf{p} - \mathbf{q} \|) r(I_p^h - I_q^h)$$

Kopf et al., "Joint Bilateral Upsampling", *SigGraph 2007*

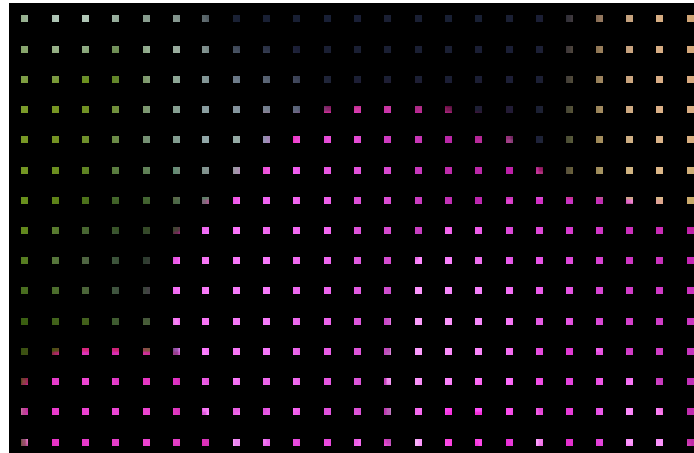


Details on [Kopf2007] – pixel dropping

- ▶ High resolution image data is used at the low resolution depth grid
 - Pixel dropping in the image
 - Lack of pre-filter → aliasing artifacts
- ▶ Straightforward pre-filtering of image data in the range term blurs edges
 - Edge alignment gone



$$r(I_p^h - I_q^h)$$



Outline

- ▶ Introduction
 - Application context
 - Joint bilateral filtering
- ▶ Image prefiltering and dual resolution depth upsampling
- ▶ Multi-step joint bilateral depth upsampling
- ▶ Computational cost
- ▶ Experimental results
- ▶ Concluding remarks



Dual resolution bilateral filter

- ▶ Our solution: **dual resolution** approach
 - Intermediate step: create downsampled version of the image
 - Apply anti-alias pre-filtering

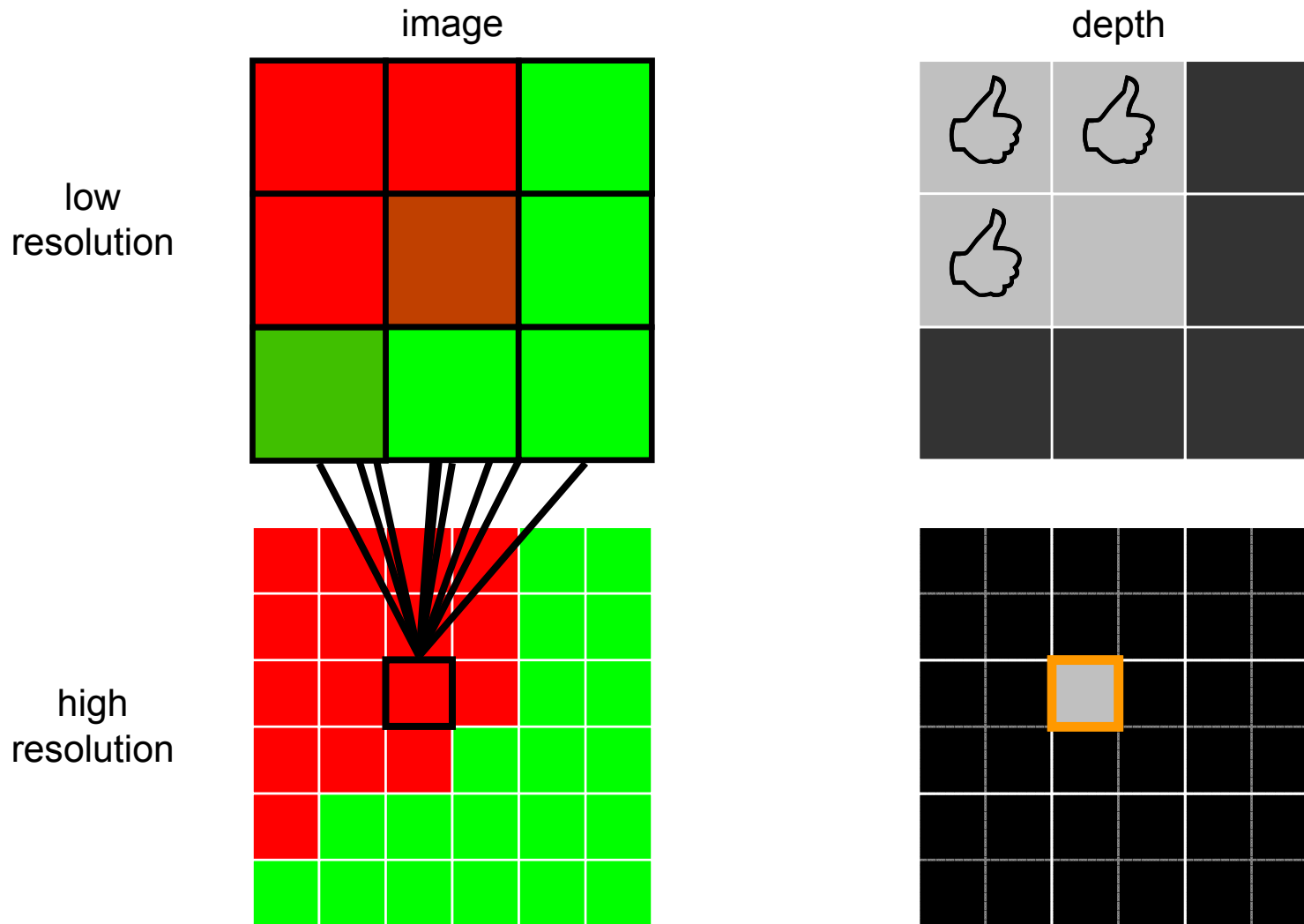
$$D_p^h = \frac{\sum_{q \in S} w_{p,q} D_q^l}{\sum_{q \in S} w_{p,q}}$$

$$w_{p,q} = s(\| \mathbf{p} - \mathbf{q} \|) r(I_p^h - I_q^l)$$

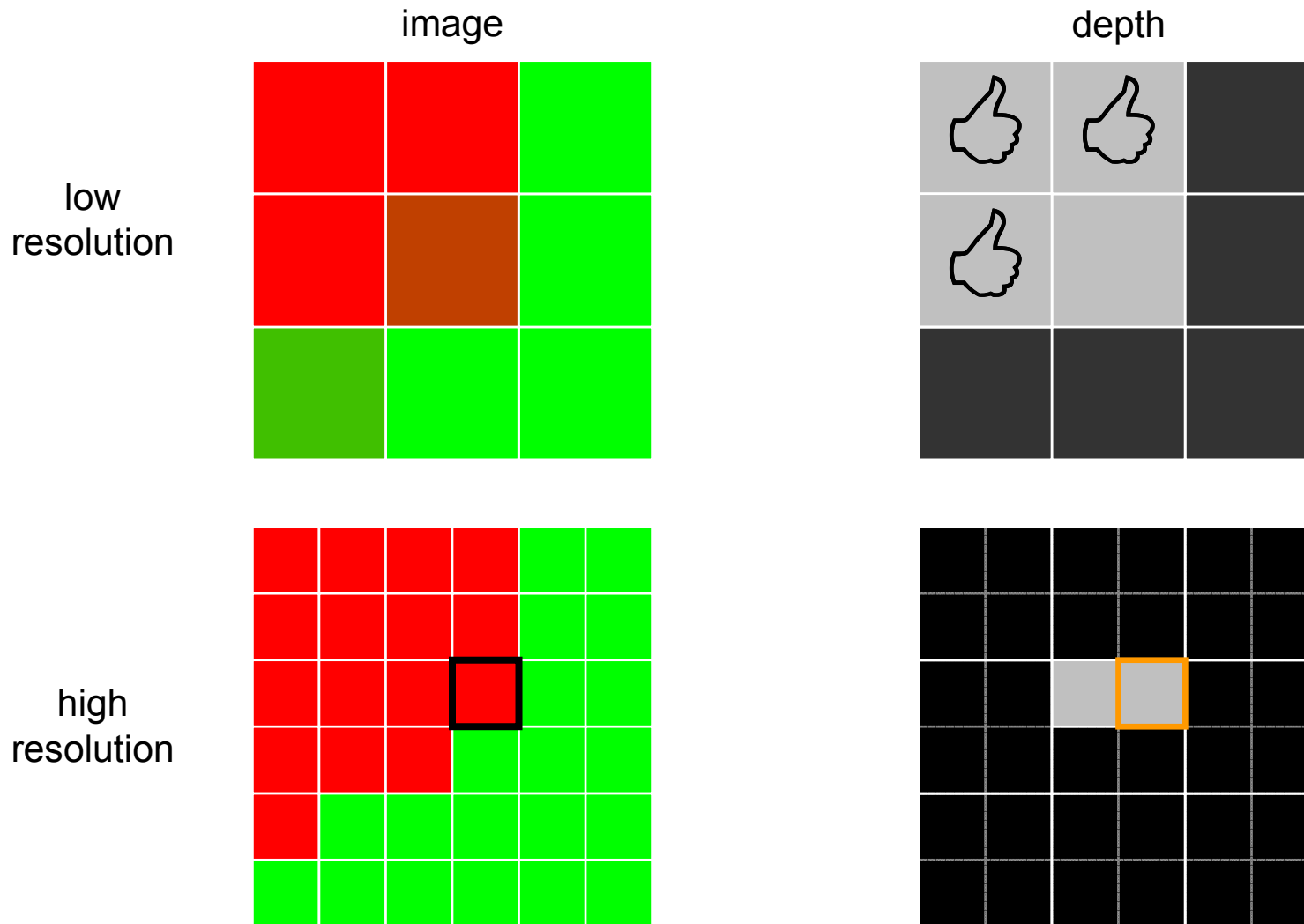
low resolution image

high resolution image

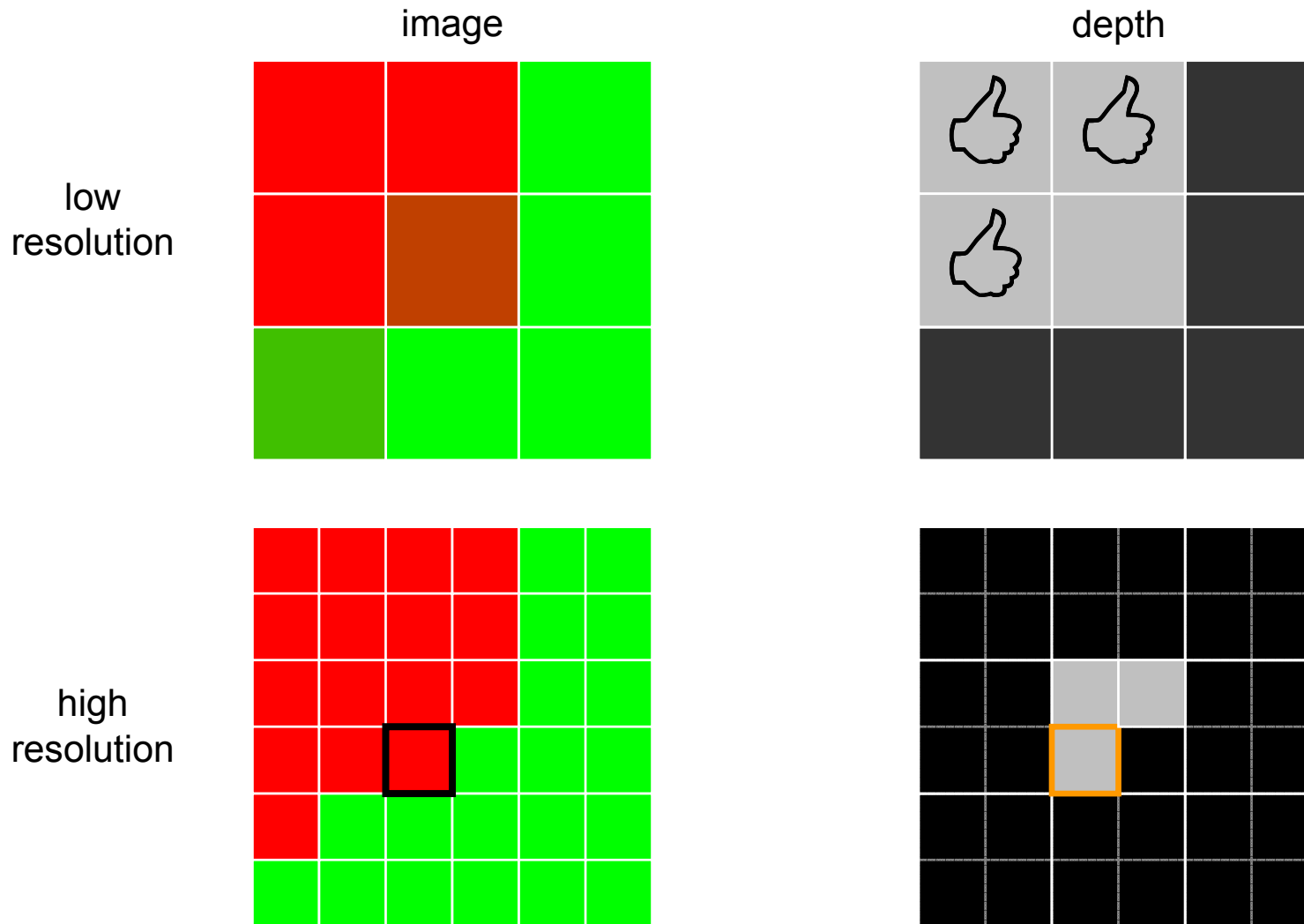
Animation of 2x2 upsampling



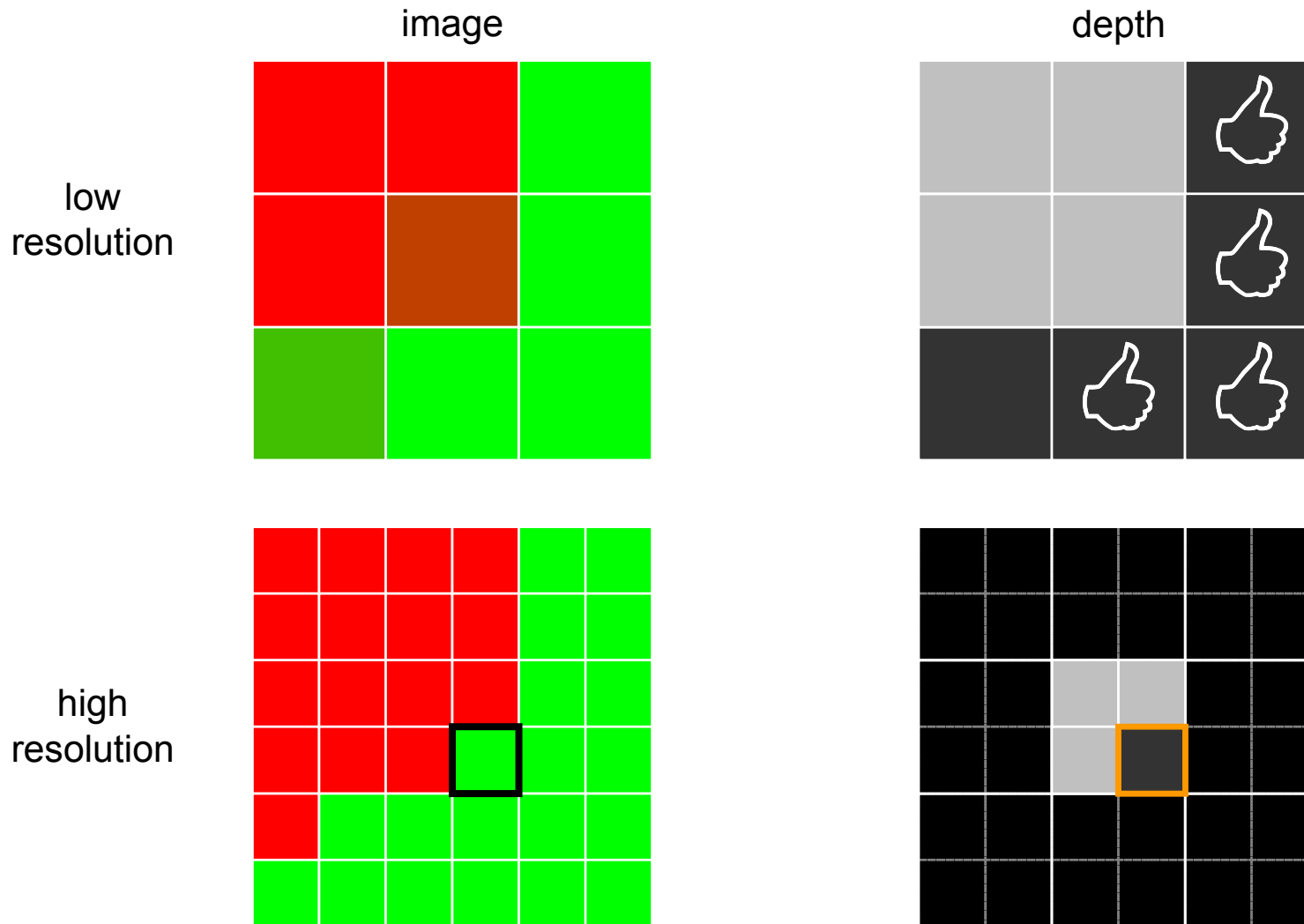
Animation of 2x2 upsampling



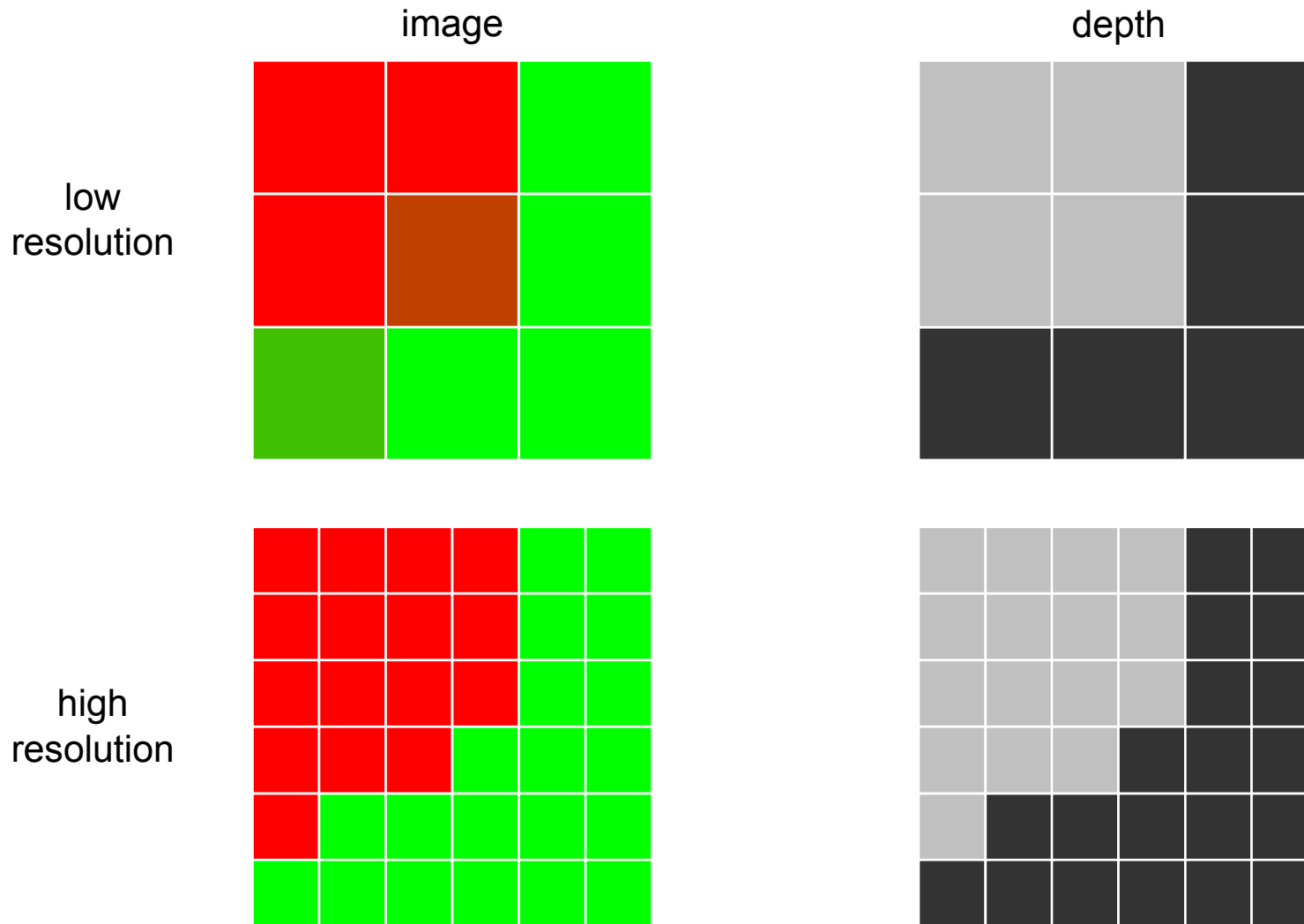
Animation of 2x2 upsampling



Animation of 2x2 upsampling



Animation of 2x2 upsampling



Anti-alias experiment

- ▶ Input “Moebius” from Middlebury test set
- ▶ Both image and ground-truth depth map shifted one pixel per frame
- ▶ Depth map downscaled to 8x8 grid after shifting



image and depth input



Kopf et al. original and proposed solution

Outline

- ▶ Introduction
 - Application context
 - Joint bilateral filtering
- ▶ Image prefiltering and dual resolution depth upsampling
- ▶ Multi-step joint bilateral depth upsampling
- ▶ Computational cost
- ▶ Experimental results
- ▶ Concluding remarks

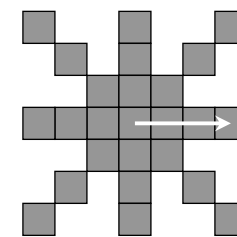


Proposed solution

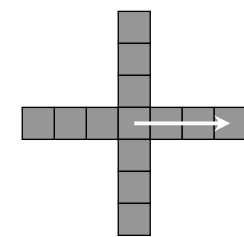
- ▶ Spatial filter s is box filter \rightarrow removed from the equation
- ▶ Range filter r is Gaussian, $\sigma_r = 0.1$
- ▶ Stepwise upsampling: 3 steps of 2×2
- ▶ Image downsampling in 3 steps of 2×2 ; each with 1,3,3,1 pre-filtering
- ▶ Filter aperture S : only cross or star pattern
- ▶ Optional extra step of joint bilateral pre-filtering at 8×8 grid

$$D_p^h = \frac{\sum_{q \in S} w_{p,q} D_q^l}{\sum_{q \in S} w_{p,q}}$$

$$w_{p,q} = r(I_p^h - I_q^l)$$



star pattern
radius 3



cross pattern
radius 3

Filter configurations

- ▶ Joint bilateral filter characterized by: size of filter aperture and σ_s
- ▶ Our stepwise solution: more freedom to adapt to signal conditions
- ▶ σ_r is 0.1 in all cases

Application: high quality depth maps
Main aim: refining object shape

Application: depth maps of lower quality
Additional aim: smooth out depth outliers

"Proposed"	Standard		Large	
	Pattern	Radius	Pattern	Radius
JBF (8×8)	n.a.	n.a.	star	5
Step 1 (8×8)	cross	1	star	2
Step 2 (4×4)	cross	1	cross	1
Step 3 (2×2)	cross	1	cross	1
Total aperture	36 × 36		132 × 132	
"Kopf et al."	Standard		Large	
Aperture σ_s	5×5	0.5	17×17	2.0
Total aperture	40 × 40		136 × 136	

Outline

- ▶ Introduction
 - Application context
 - Joint bilateral filtering
- ▶ Image prefiltering and dual resolution depth upsampling
- ▶ Multi-step joint bilateral depth upsampling
- ▶ Computational cost
- ▶ Experimental results
- ▶ Concluding remarks



Algorithmic operations per output pixel

- ▶ Operation types (here added with equal weights):
 - Basic: add, sub, abs, lut, clip, round, etc.
 - Multiplication
 - Division

	Standard				Large			
	basic	mul	div	total	basic	mul	div	total
Proposed	75.5	9.6	1.3	86	90.7	10.9	1.3	103
Kopf et al.	323.0	125.0	1.0	449	3755.0	1445.0	1.0	5201
Gain factor				5.2				50.5

Larger footprint mostly on 8×8 grid, therefore only slight increase on ops/pixel for large increase in footprint

Linear scaling with number of filter taps (approx. factor of 12)

Huge benefit in computational cost

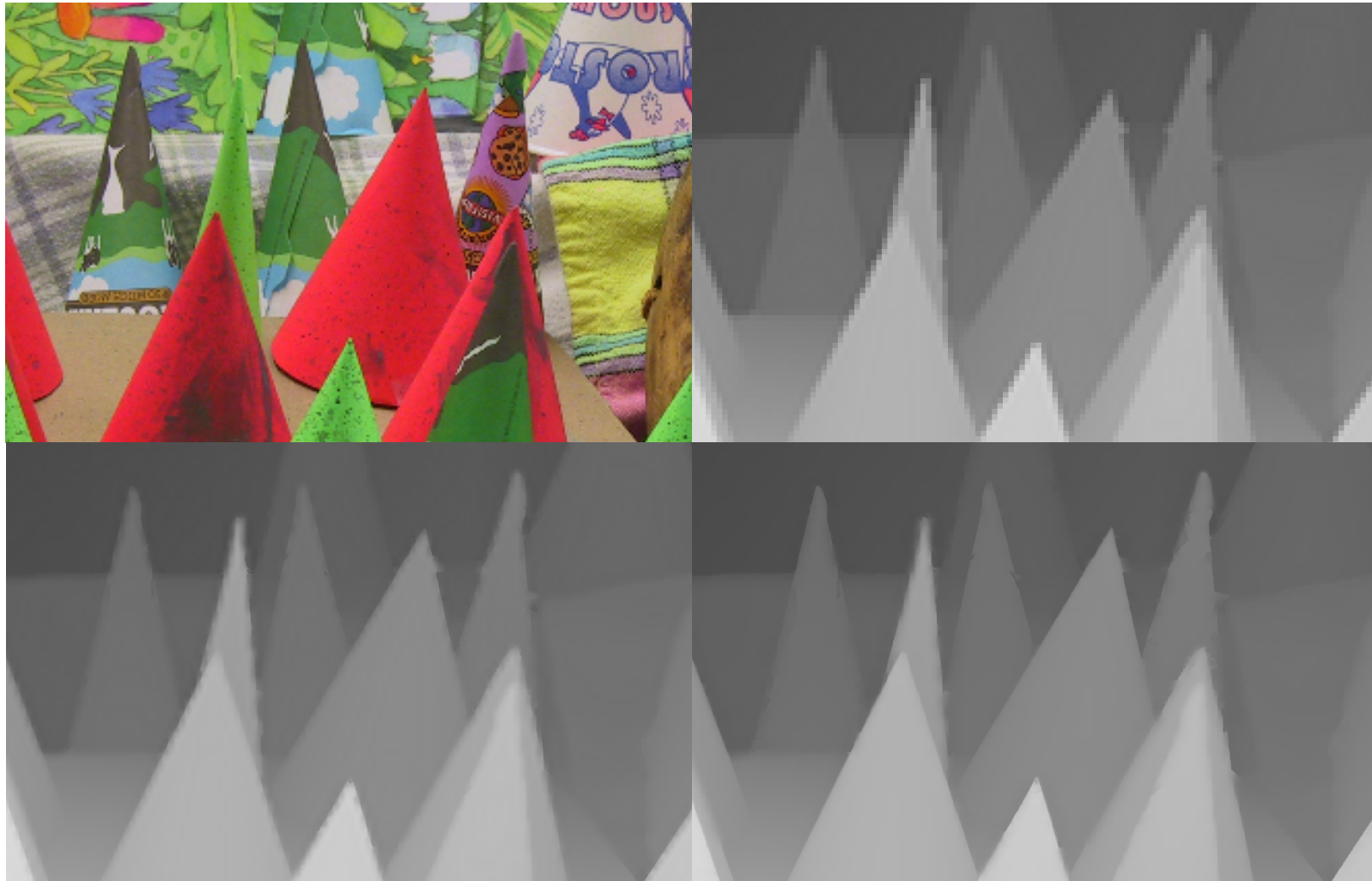


Outline

- ▶ Introduction
 - Application context
 - Joint bilateral filtering
- ▶ Image prefiltering and dual resolution depth upsampling
- ▶ Multi-step joint bilateral depth upsampling
- ▶ Computational cost
- ▶ **Experimental results**
- ▶ Concluding remarks



Processing high quality depth maps (standard)



Kopf et al.: standard aperture

Proposed: standard configuration

Processing lower quality depth maps (large)



Kopf et al.: large aperture

Proposed: large configuration

Outline

- ▶ Introduction
 - Application context
 - Joint bilateral filtering
- ▶ Image prefiltering and dual resolution depth upsampling
- ▶ Multi-step joint bilateral depth upsampling
- ▶ Computational cost
- ▶ Experimental results
- ▶ Concluding remarks



Other applications

- ▶ Depth map from range camera
- ▶ Depth map enhancement
- ▶ Process in downsampled domain for efficiency
- ▶ Other multi-modal image enhancement
 - Image colorization
 - Adaptive tone mapping



Summary and conclusions

- ▶ Showed multi-step joint bilateral filter with anti-alias image prefiltering
- ▶ Application on depth maps for 8×8 upsampling
- ▶ Dual-resolution range term improves temporal consistency over state-of-the-art
- ▶ Our approach shows similar or better image quality compared to state-of-the-art on two use-cases: high quality and lower quality inputs
- ▶ Multi-step approach shows favorable compute load scalability
- ▶ For lower quality inputs, our method favors a reduction of computational cost by a factor of 50

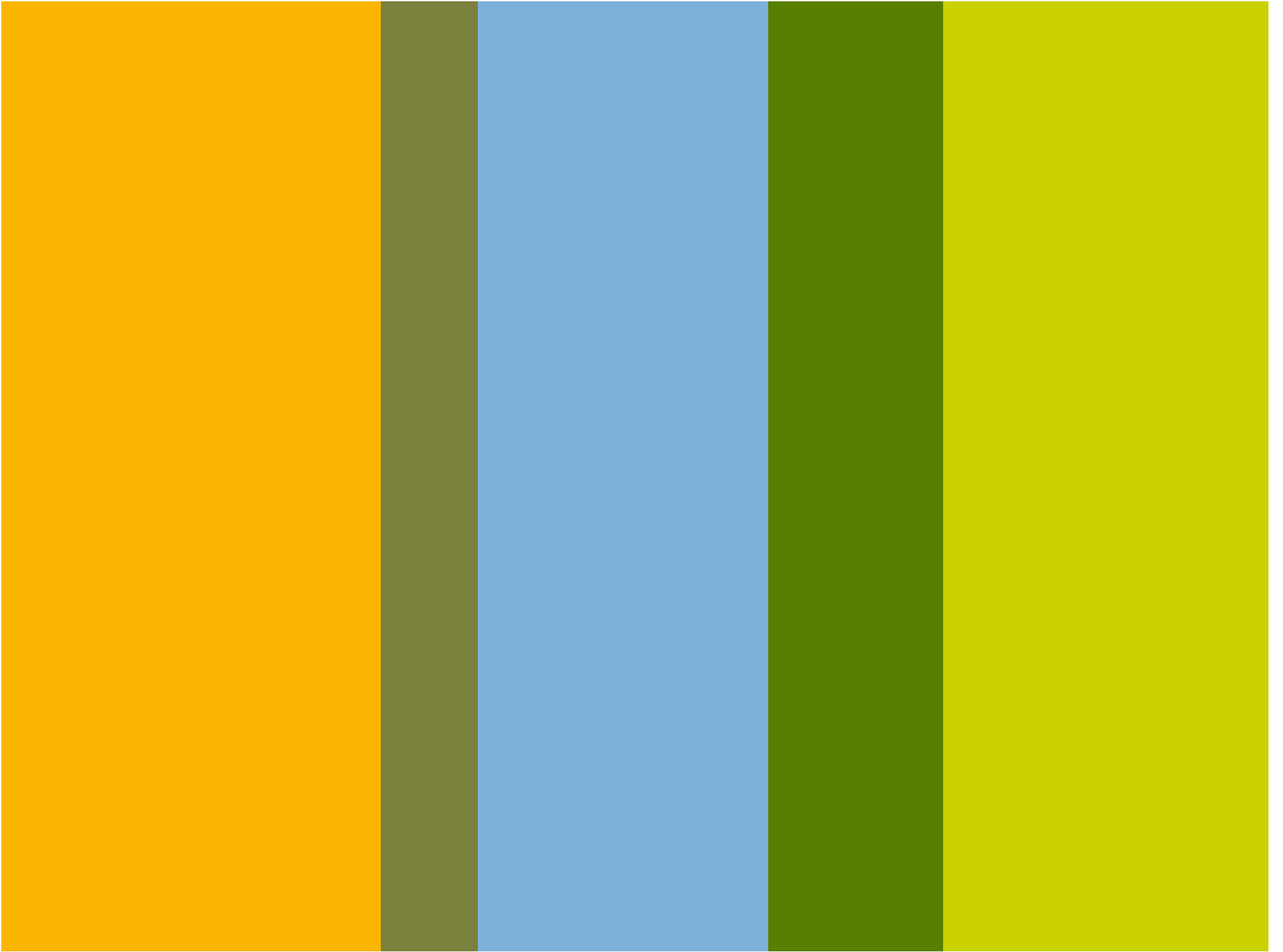


Acknowledgements

- ▶ Johannes Kopf (Univ. of Konstanz, Germany) for sharing code of SigGraph 2007 paper
- ▶ Chris Varekamp (Philips Research, Netherlands) for inspiring initial discussions

... and finally ... questions???





Bonus: flower with generated view

- ▶ Processed depth map used to generate a new view



stereo input

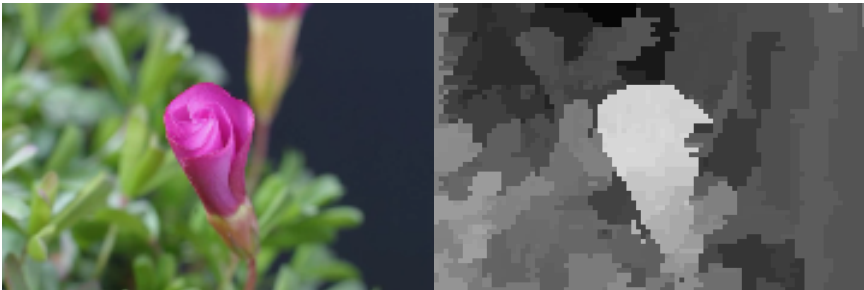
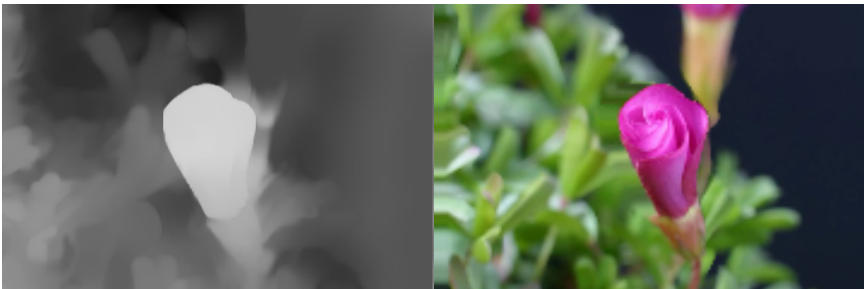


image and estimated depth map



filtered depth map and
generated view

