PCM -Nanjing 2013

A Complementary Aggregation Approach for Local Stereo Matching

Ran Ju

Yang Yang

Xiangyang Xu

Chunrong Xia

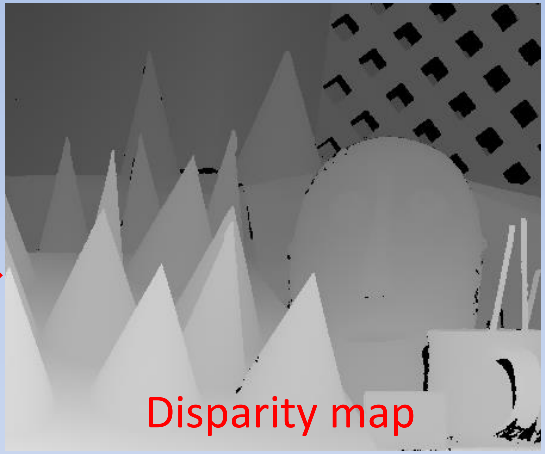
Gangshan Wu

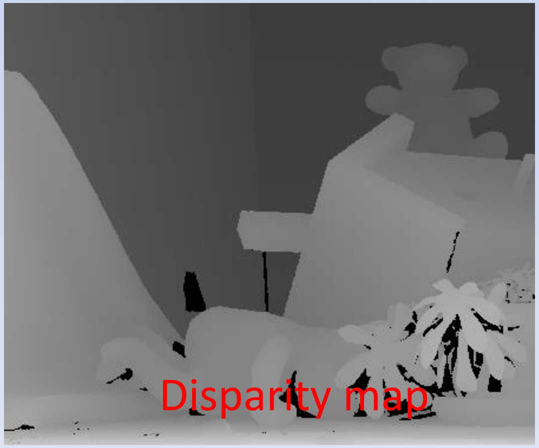
State Key Laboratory for Novel Software Technology, Nanjing University

juran@smail.nju.edu.com

Stereo matching – disparity map

- Stereo matching——find corresponding pixels in two images.



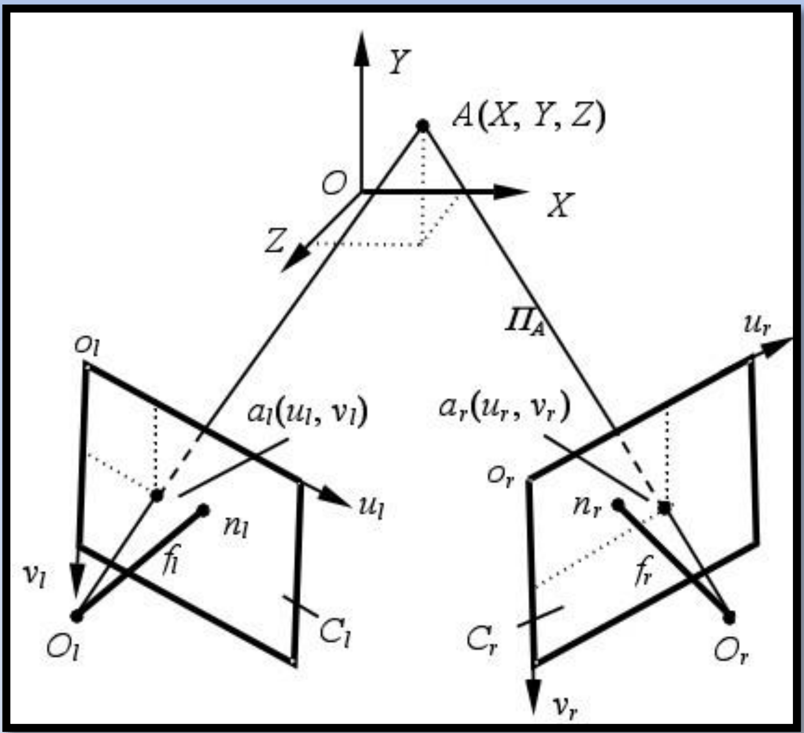


Left image

Right image

Stereo rectification

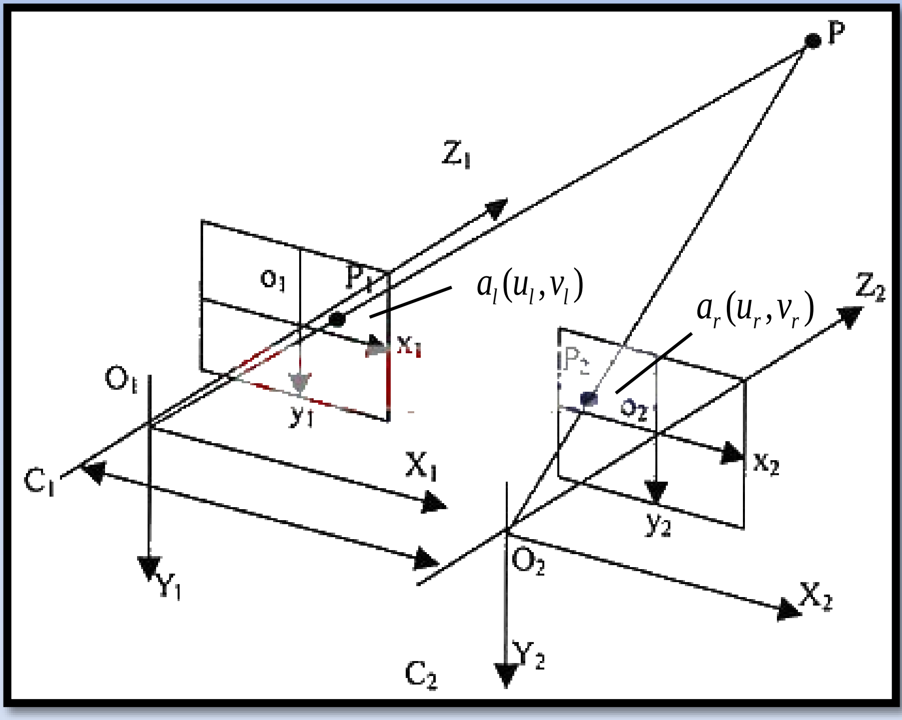
- Uncalibrated



$$d_u = u_l - u_r$$

$$d_v = v_l - v_r$$

- Calibrated



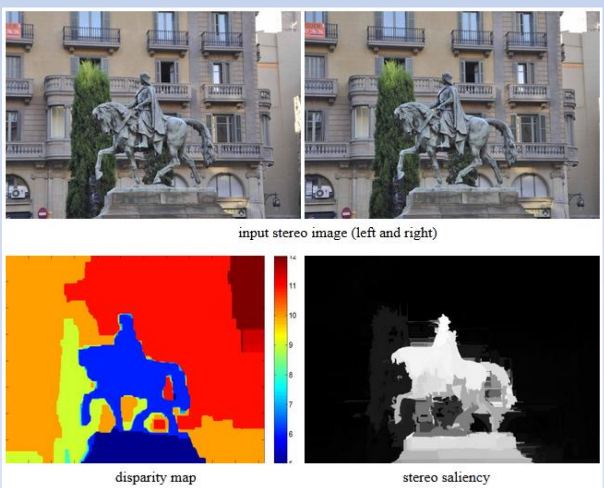
$$d_u = 0$$

$$d_v = v_l - v_r \longrightarrow disparity$$

Applications

View synthesis. C. L. Zitnick. ToG 2004.

Stereo image editing. S. Gupta. Siggraph 2010.

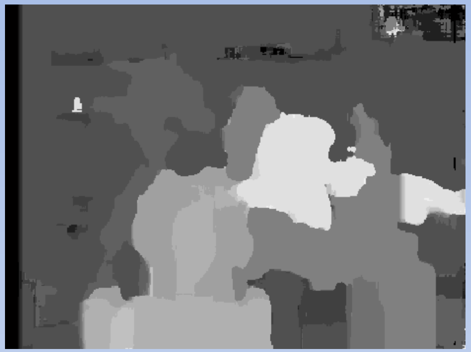


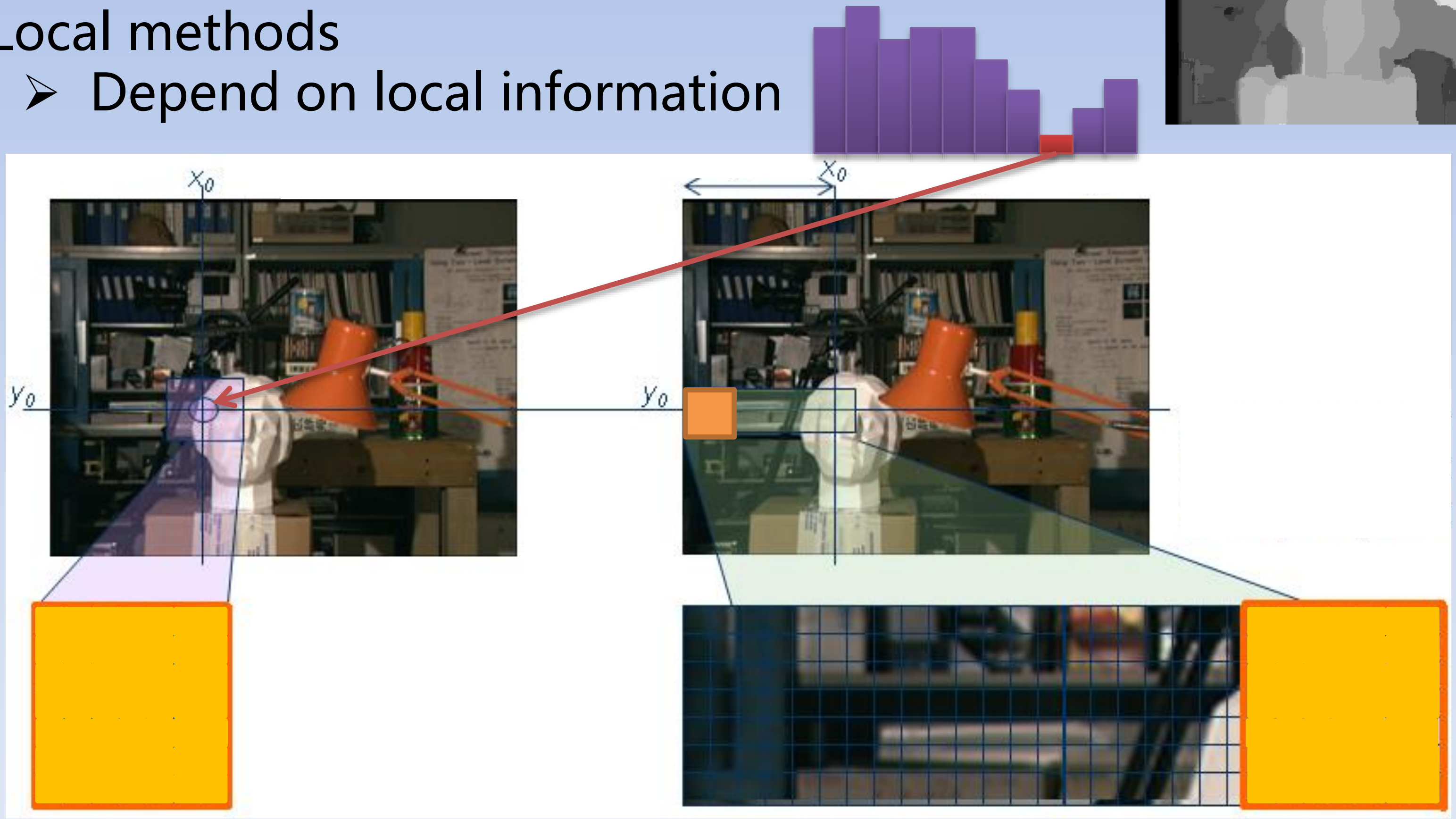
Saliency analysis. Y. Niu. CVPR 2012.

Scene analysis. S. Gupta. CVPR 2013.

Related Work

- Global methods
  - ➢ Smoothness assumption
  - ➢ Optimization problem
- Local methods
  - ➢ Depend on local information

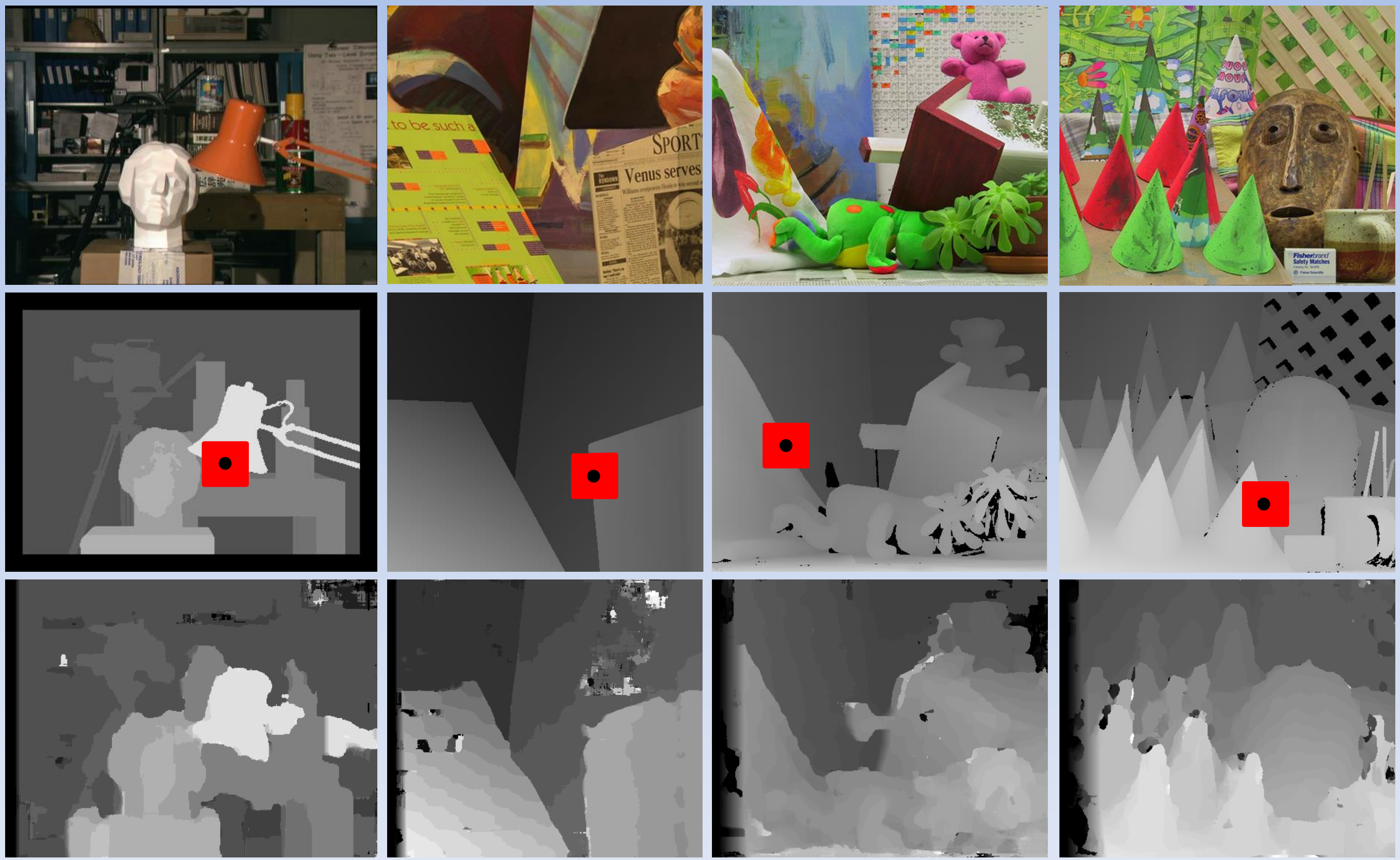




Scharstein, D., Szeliski, R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. IJCV 2002.

Local methods

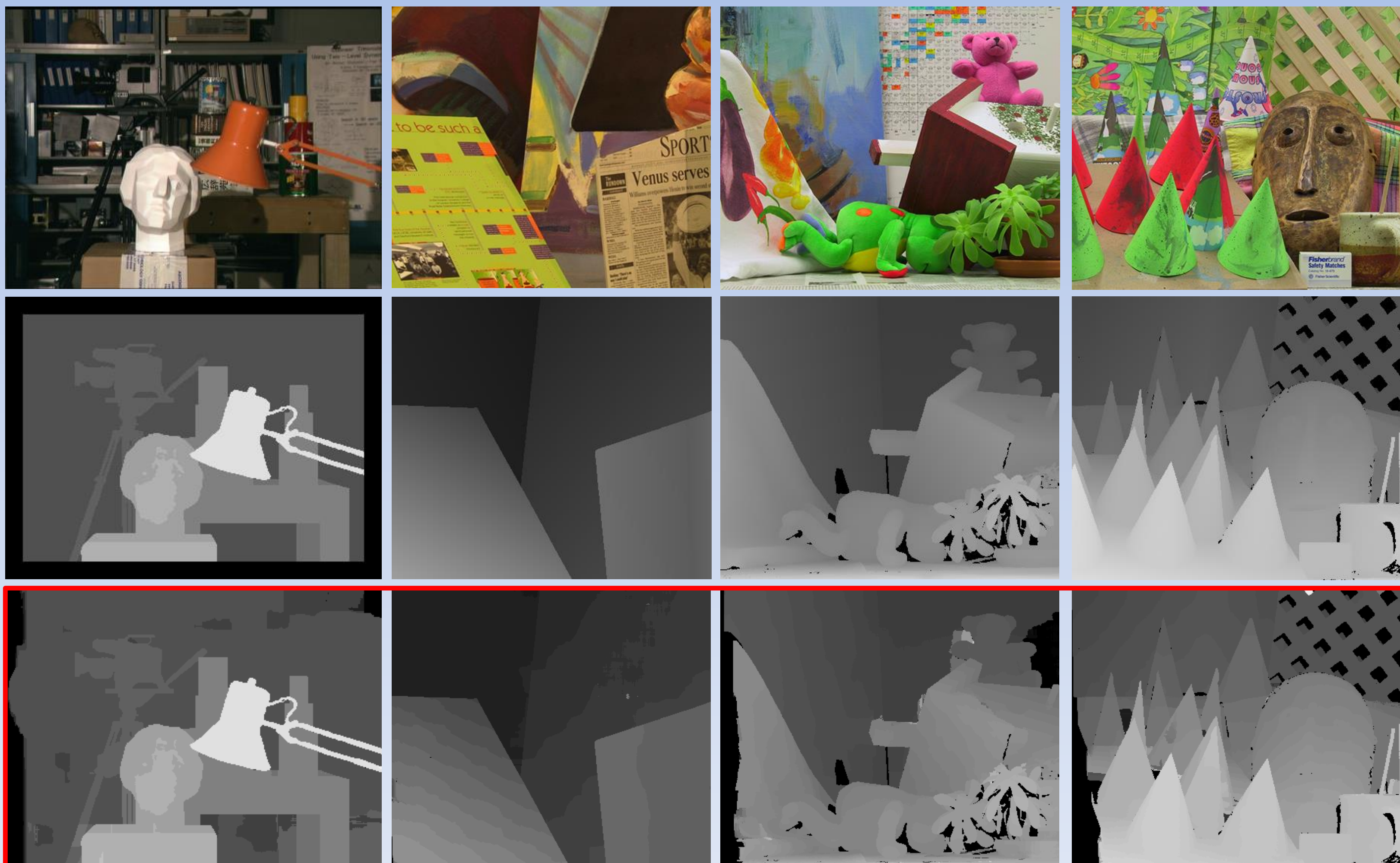
- Challenges



Middlebury dataset [1]

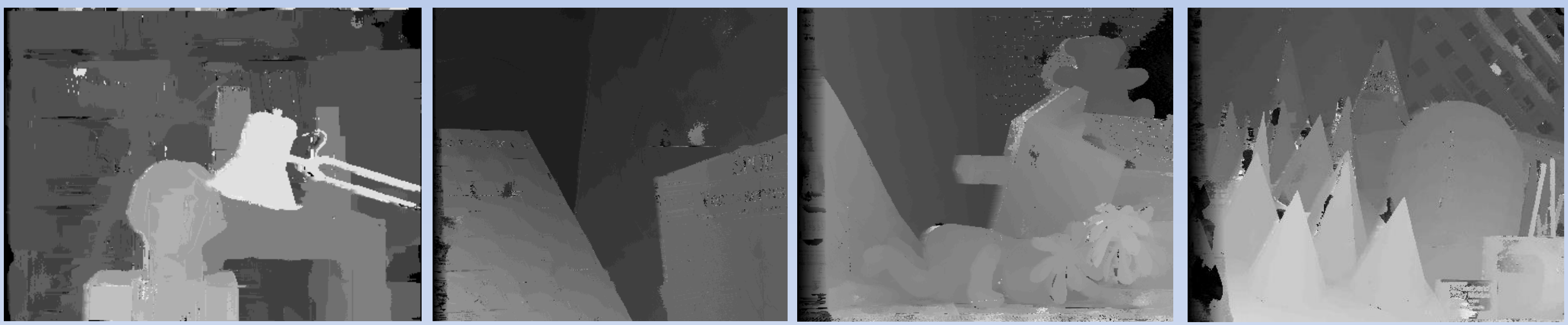
Ideal aggregation

YOUR METHOD	32.2	0.56 1	0.55 1	0.07 1	0.64 86	0.80 68	4.75 86	6.59 60	12.1 65	11.4 14	1.33 1	5.77 2	3.05 1	3.97

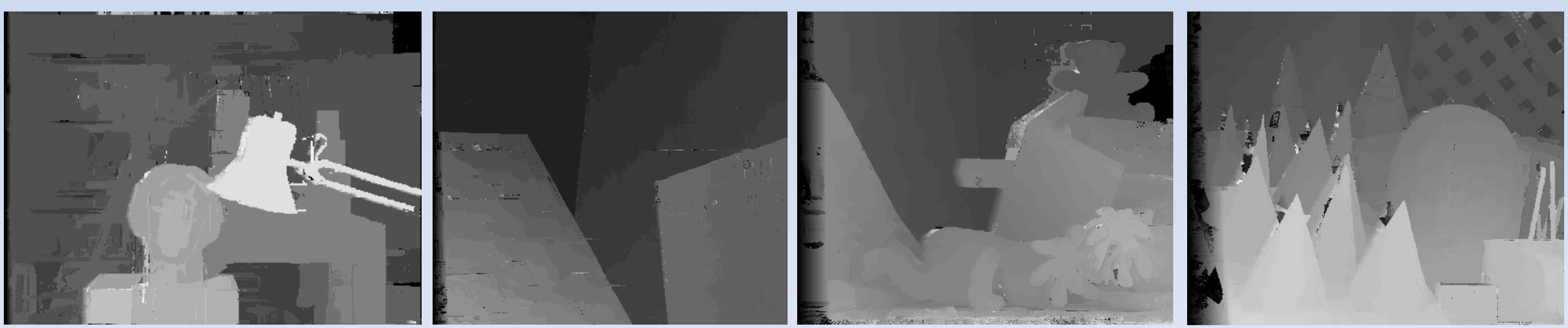


Approximations

- Pixels inside a local region are likely to have similar depth.
- Adaptive weight (Yoon, 2006)
  - Pixels with similar colors are likely to have similar depth.
- Segment support (Tombari, 2007)
  - Pixels inside a segment are likely to have similar depth.



Adaptive weight [2]



Segment support[2]

Problem

- Insufficient support
- Fail to work in highly textured area

7.03%

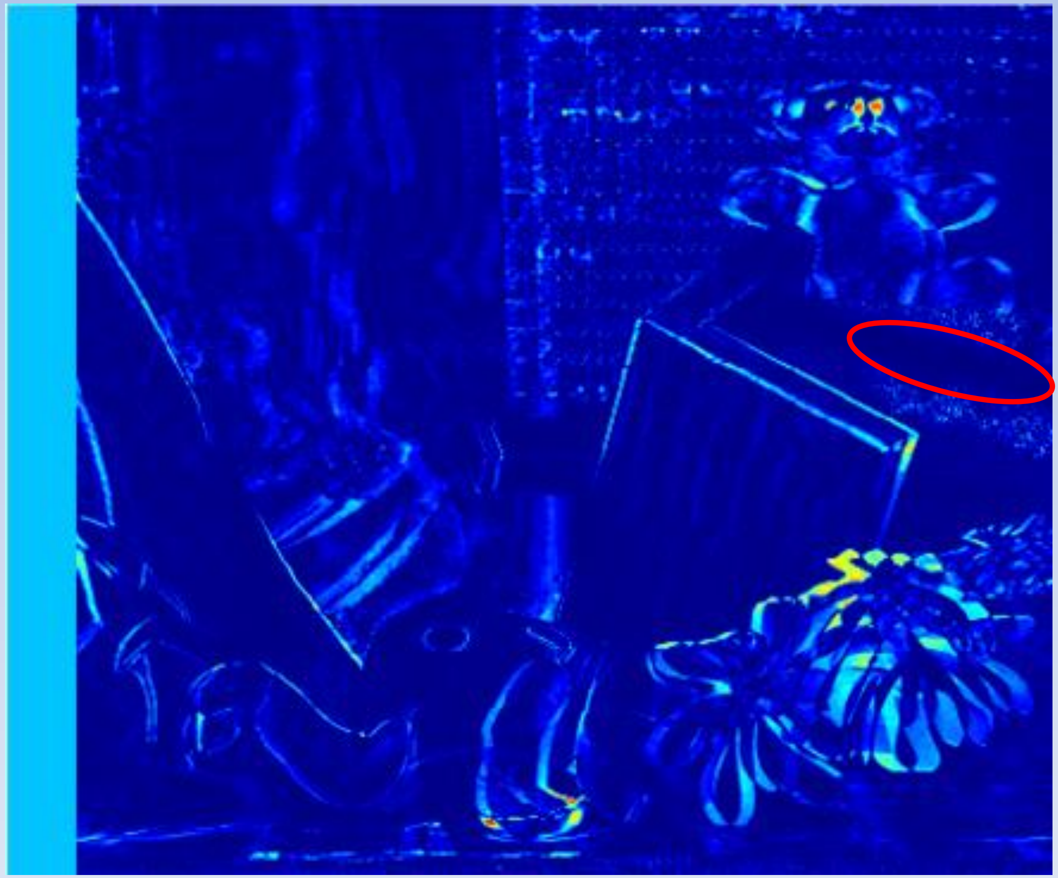
Same depth Similar color	Same depth Dissimilar color
Different depth Similar color	Different depth Dissimilar color

82.52%

Find some other cues…

Correlation cue

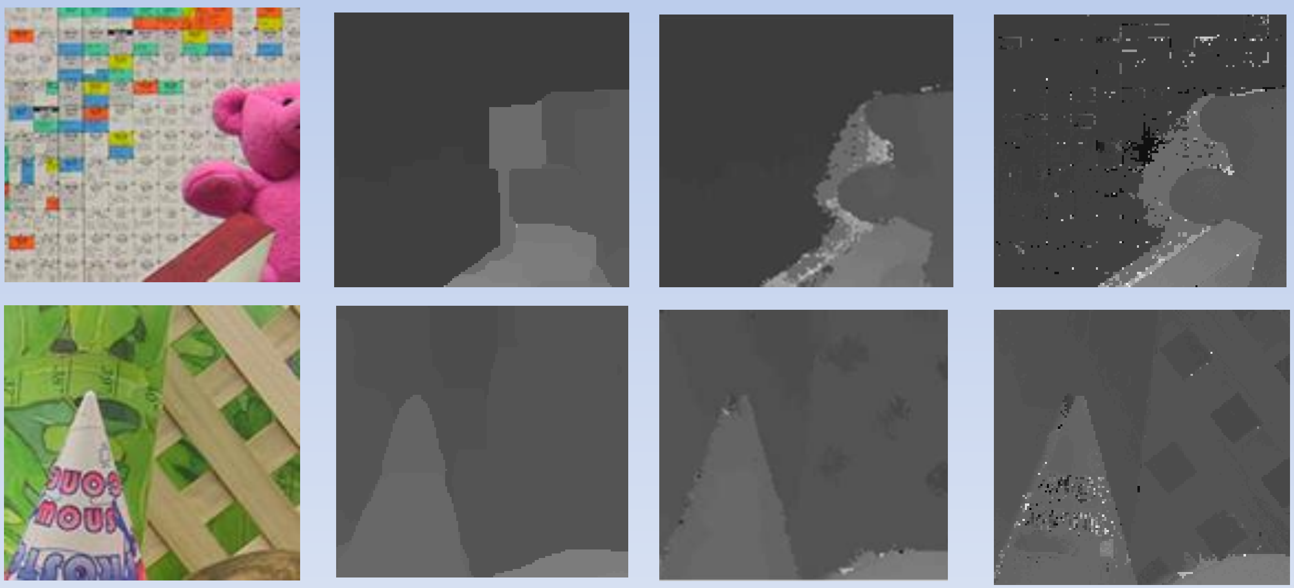
- Pixels with the same depth aggregate to a “dark region” at the correct disparity slice.



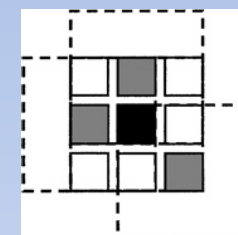
Adaptive window methods

- Search for the surrounding area (sub windows) with highest correlation on all disparity hypothesizes as the support region.

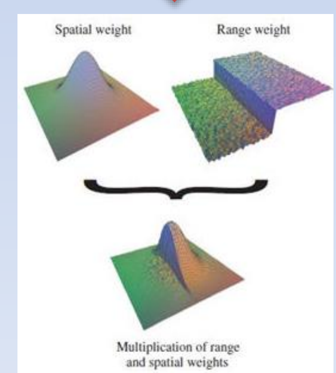
Color image | Shiftable window | Adaptive weight on correlation | Adaptive weight on color



High texture ✓ Discontinuity ✕ | High texture ✓ Discontinuity Δ | High texture ✕ Discontinuity ✓

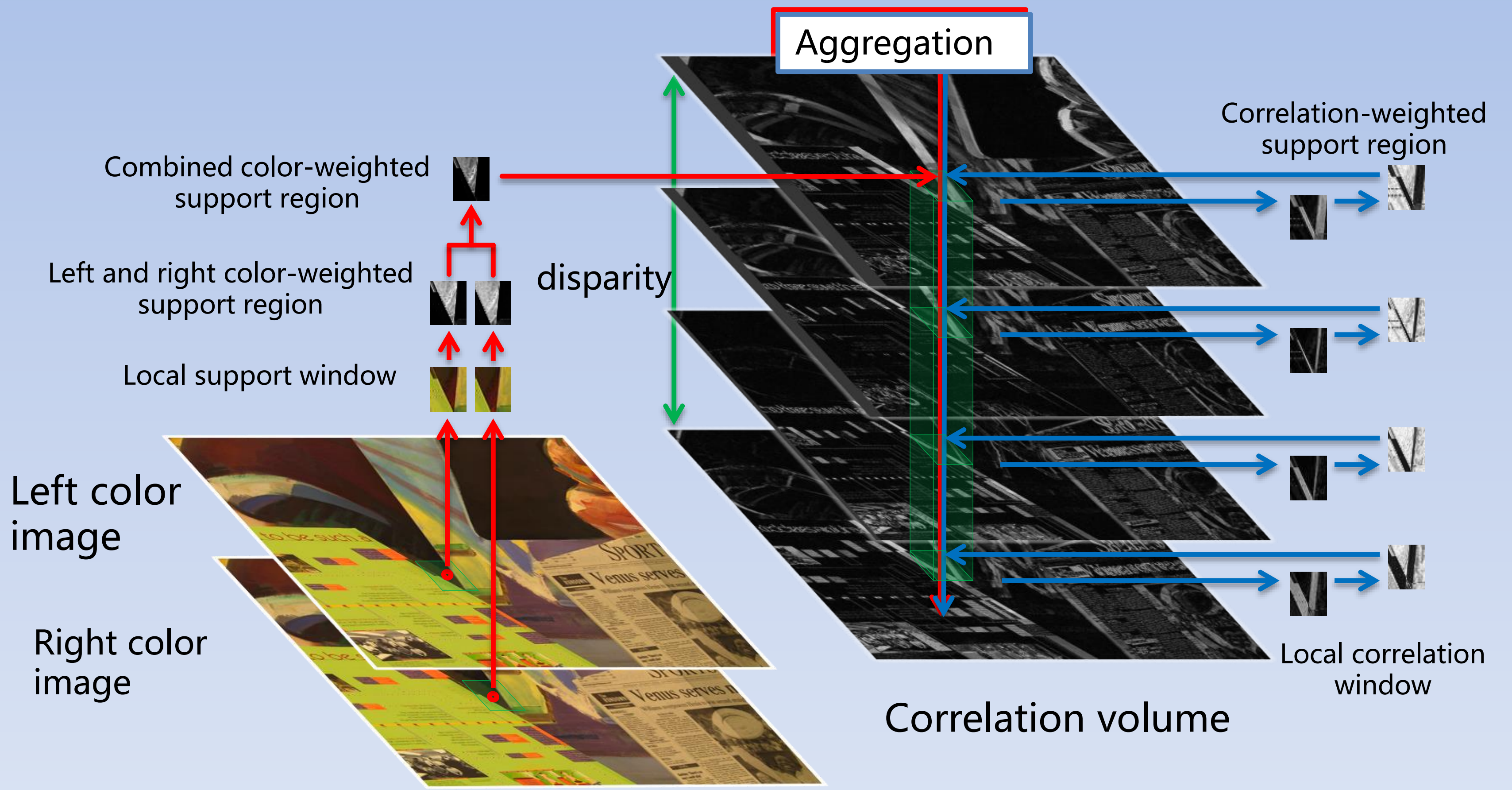


Try all 3x3 shifted windows around the black pixel and take the sub window with minimum matching score.



Edge aware aggregation

The proposed method



Approach

- Matching cost of pixel (p, d)

$$C(p,d) = \frac{1}{3}\sum_{c\in\{R,G,B\}} | I_l^c(p) - I_r^c(p-d) |$$

- Correlation-based Aggregation (Bilateral filter):

$$C_A(p, d) = \frac{\sum_{q\in N(p)} \omega_o(p, q) C(p, d)}{\sum_{q\in N(p)} \omega_o(p, q)}$$

$$\omega_o(p,q) = \exp(-(\frac{|C(q,d) - C(p,d)|}{\gamma_o} + \frac{\| p-q \|_2}{\eta_o}))$$

Approach

- Color-based Aggregation(Guided bilateral filter):

$$C'_A(p,d) = \frac{\sum_{q \in N(p), q_d \in N(p_d)} \omega_c(p,q)\omega_c(p_d,q_d)C_A(q,d)}{\sum_{q \in N(p), q_d \in N(p_d)} \omega_c(p,q)\omega_c(p_d,q_d)}$$

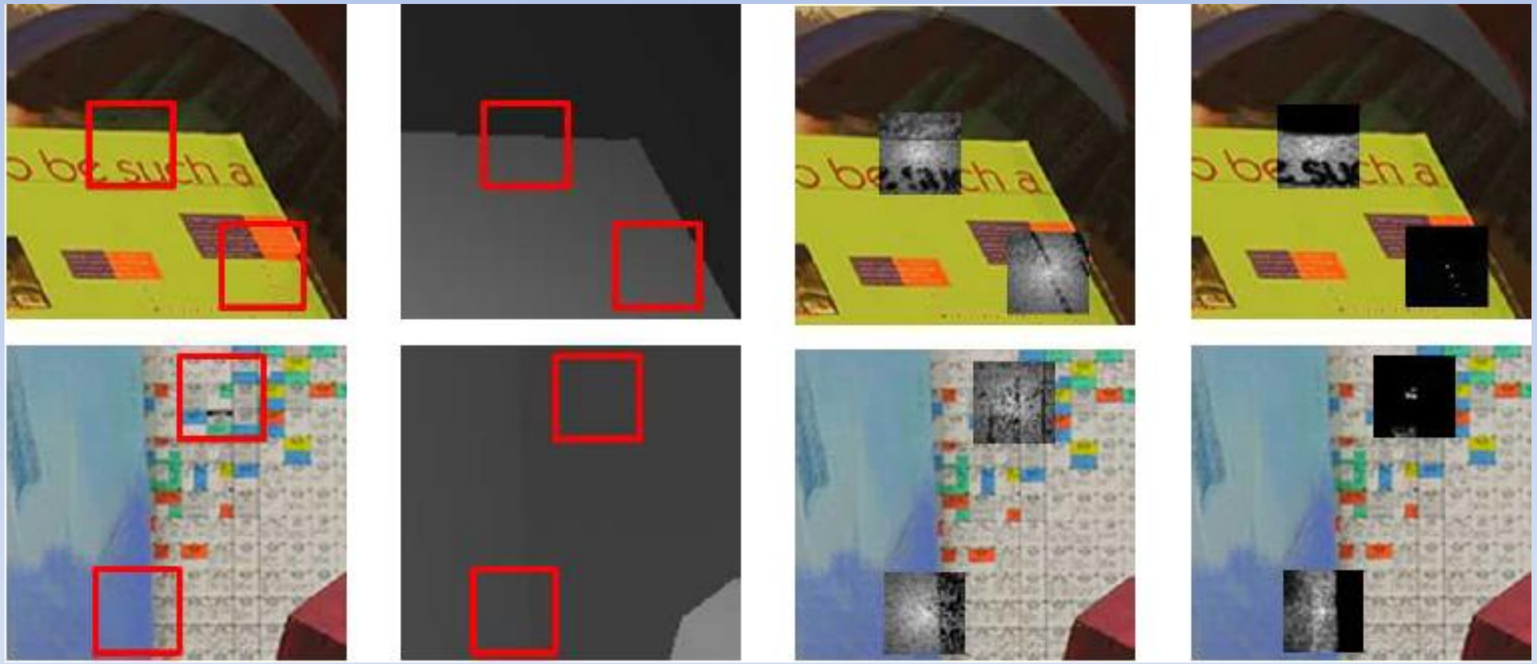
$$\omega_c(p,q) = \exp(-(\frac{\Delta_{pq}}{\gamma_c} + \frac{\| p-q \|_2}{\eta_c}))$$

- The WTA (winner-take-all) strategy to compute the final disparity map:

$$d(p) = \arg \min_{0 \le d \le d_{MAX}} C'_A(p,d)$$

Approach

- Local supports of correlation-based and color-based aggregation



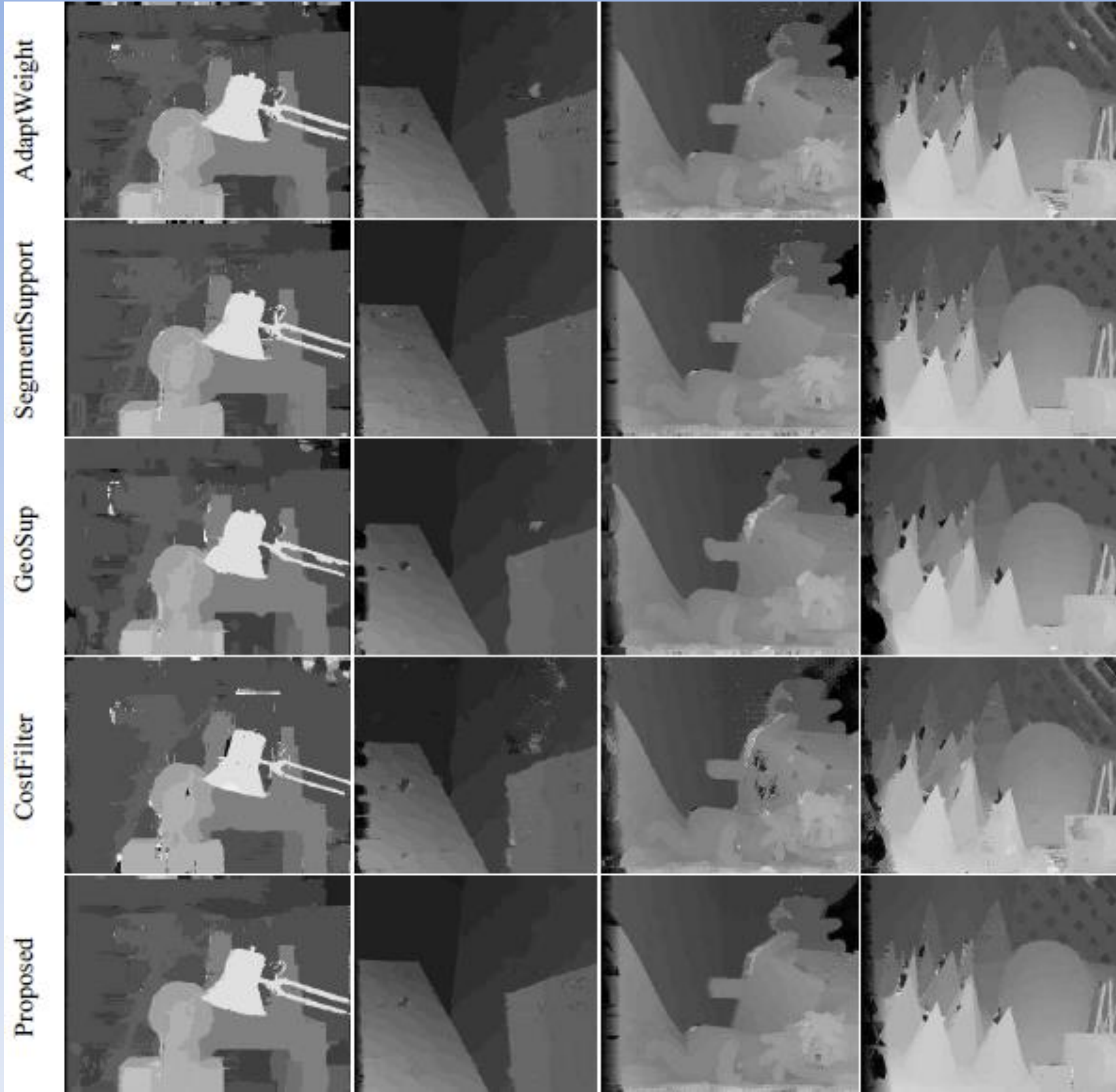
Color image | Ground truth | local supports of correlation-based aggregation | local supports of color-based aggregation

Evaluation

- Dataset
  - http://vision.middlebury.edu/stereo/data/
  - Stereo photos taken in real world

- The benchmark
  - The percent of bad pixels

- The state-of-the-art methods to compare
  - Adaptive weight (Yoon, PAMI, 2006)
  - Segment support (Tombari, PSIVT, 2007)
  - Geosup (Hosni, ICIP, 2009)
  - Cost filter (Rhemann, PAMI, 2011)


AdaptWeight
SegmentSupport
GeoSup
CostFilter
Proposed

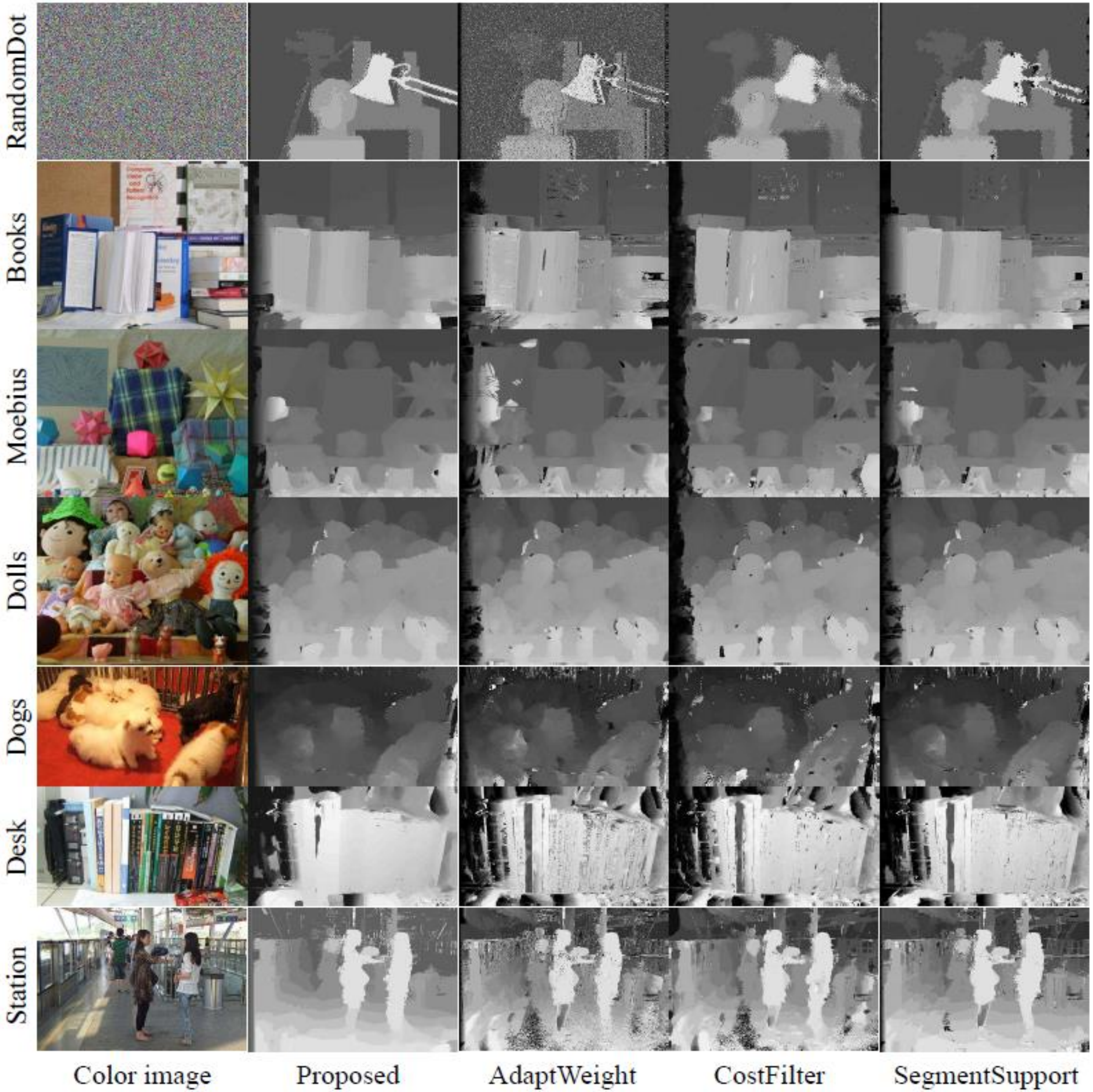
accuracy

Table 2. Bad pixels in non-occluded regions

Algorithm	Tsukuba	Venus	Teddy	Cones	Average	
AdaptWeight [2]	2.82	2.76	12.1	9.66	6.84	2.02 ↑
SegmentSupport [3]	2.05	1.47	10.8	5.08	4.85	0.03 ↑
GeoSup [4]	3.16	2.74	11.6	5.11	5.65	0.83 ↑
CostFilter [5]	3.23	4.53	13.7	15.3	9.19	4.37 ↑
Proposed	2.01	1.25	11.1	4.93	4.82	

Table 3. Bad pixels in discontinuous regions

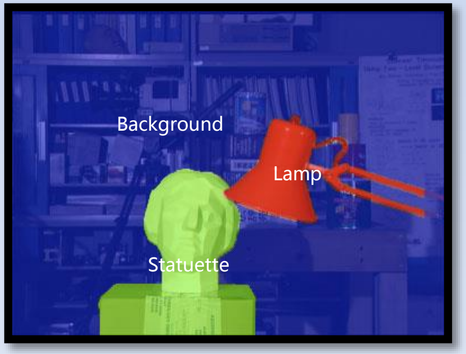
Algorithm	Tsukuba	Venus	Teddy	Cones	Average	
AdaptWeight [2]	7.38	10.4	21.4	15.9	13.77	2.39 ↑
SegmentSupport [3]	7.14	10.5	21.7	12.5	12.96	1.58 ↑
GeoSup [4]	10.1	17.3	22.9	12.3	15.65	4.17 ↑
CostFilter [5]	9.16	20.8	24.1	22.1	19.04	7.66 ↑
Proposed	7.07	5.86	21.2	11.4	11.38	


RandomDot
Books
Moebius
Dolls
Dogs
Desk
Station
Color image
Proposed
AdaptWeight
CostFilter
SegmentSupport

Conclusion and Future work

- A complementary aggregation approach using correlation and color cues
- An adaptive weight approach for correlation-based aggregation

- Incorporate our approach with more robust matching cost measurement for real applications
- Supply a higher level description using the disparity map for semantic analysis.



Object level description

References

[1] Scharstein, D., Szeliski, R.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. International Journal of Computer Vision 47(1) (2002) 7–42

[2] Yoon, K., Kweon, I.: Adaptive support-weight approach for correspondence search. Pattern Analysis and Machine Intelligence, IEEE Transactions on 28(4) (2006) 650–656

[3] Tombari, F., Mattoccia, S., Di Stefano, L.: Segmentation-based adaptive support for accurate stereo correspondence. Advances in Image and Video Technology (2007) 427–438

[4] Hosni, A., Bleyer, M., Gelautz, M., Rhemann, C.: Local stereo matching using geodesic support weights. In: Image Processing (ICIP), 2009 16th IEEE International Conference on, 2093–2096

[5] Rhemann, C., Hosni, A., Bleyer, M., Rother, C., Gelautz, M.: Fast cost-volume filtering for visual correspondence and beyond. In: Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, 3017–3024

Thank you!

Q&A