Computer Vision Final Project

Binary Stereo Matching

B05901062 阮明皓
B05901182 潘彥銘

Cost Computation

- BRIEF descriptor:

$$B(x) = \sum_{1 \leq i \leq n} 2^{i-1} \tau(p_i, q_i) \qquad (1)$$

- Each pair (p_i , q_i) is sampled by Gaussian distribution in an S × S window, which is centered on pixel x.

.

Cost Computation

- And τ (p_i , q_i) is a binary function which is defined as:

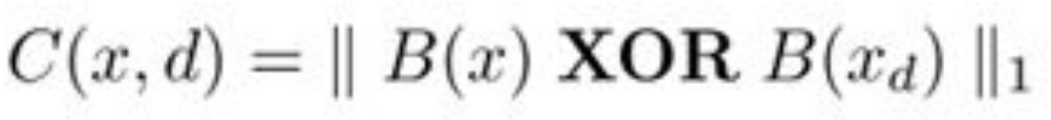
$$\tau(p_i, q_i) = \begin{cases} 1 & : \quad I(p_i) > I(q_i) \\ 0 & : \quad I(p_i) \leq I(q_i) \end{cases} \qquad (2)$$

- I(x) denotes the intensity of pixel x.

.

Cost Computation

- Cost volume:



$$C(x, d) = \| \; B(x) \; \mathbf{XOR} \; B(x_d) \; \|_1 \qquad (3)$$

- x_d is the corresponding pixel of x with disparity d in another view
- C(x, d) measures the hamming distance between two binary strings.

.

Cost Aggregation

- Binary mask:

$$\Phi(x) = \sum_{1 \leq i \leq n} 2^{i-1} \delta(x, p_i, q_i) \qquad (6)$$

.

Cost Aggregation

- Bitwise mask function for a given pair (p_i , q_i) :

$$\delta(x, p_i, q_i) = \begin{cases} 1 & : \quad w(x, p_i, q_i) \leq T \\ 0 & : \quad w(x, p_i, q_i) > T \end{cases} \qquad (5)$$

- where T is set to be the quarter smallest value in the sequence w(x, p_1, q_1), w(x, p_2, q_2), ..., w(x, p_n, q_n).

.

Cost Aggregation

- Weight function for pixel pair (p_i , q_i) as:

$$w(x, p_i, q_i) = \max(SAD(x, p_i), SAD(x, q_i)) \quad (4)$$

- SAD(x, y) = $\sum_{c\in[L,A,B]}|I_{ç}(x)-I_{ç}(y)|$ is the sum of absolute difference between two pixels in the CIELAB color space.

.

Cost Aggregation

- Incorporating the binary mask into (3), the new cost volume:

$$C(x, d) = \|B(x) \ \mathbf{XOR} \ B(x_d) \ \mathbf{AND} \ \Phi(x)\|_1 \qquad (7)$$

.

Disparity Optimization

- We implemented the Winner-Take-All(WTA) method mentioned in class.

.

Disparity Refinement

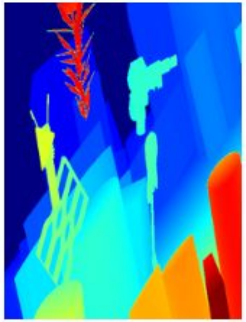
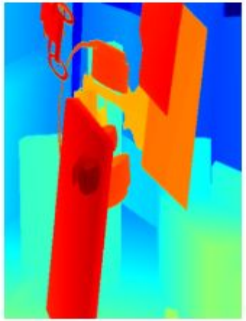
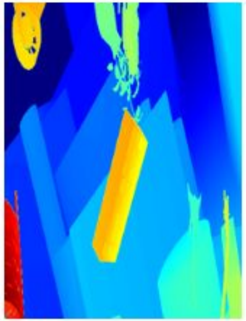
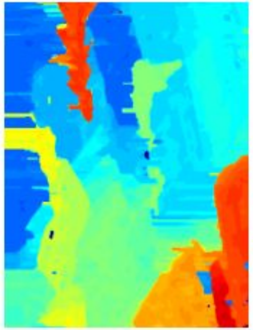
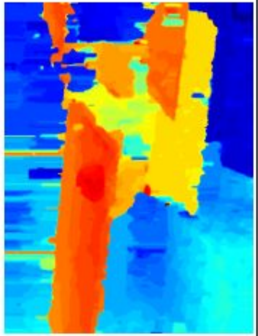
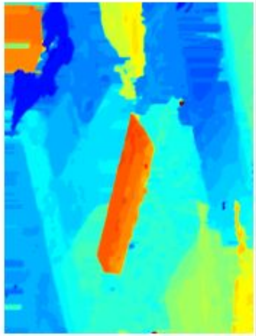
- Apply a left/right consistency check
- Classify depth results into two categories: valid and invalid.
- For an invalidated pixel p, we search its closest valid pixel to the left and to the right. We select the lower of the two as p’s refined disparity.

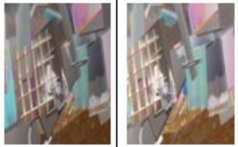
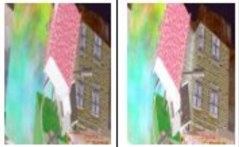
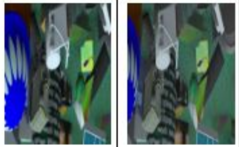
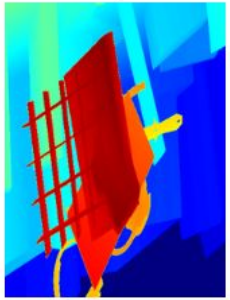
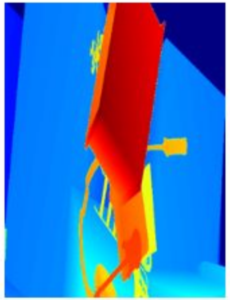
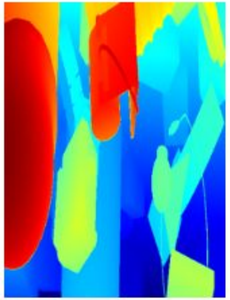
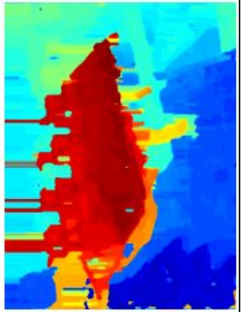
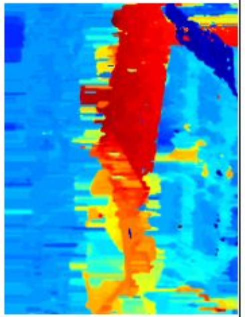
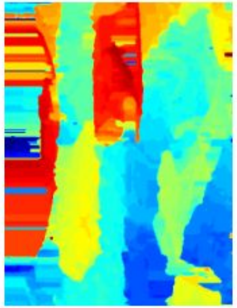
Disparity Refinement

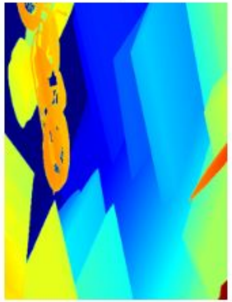
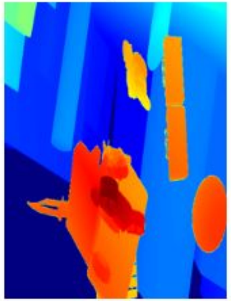
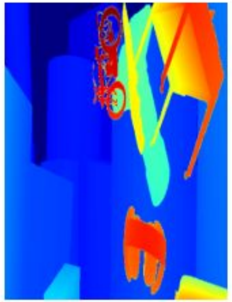
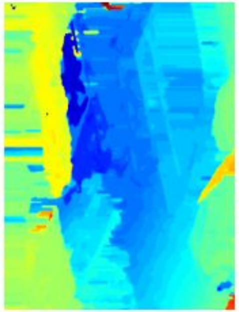
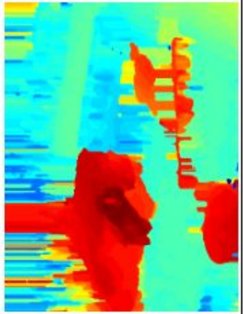
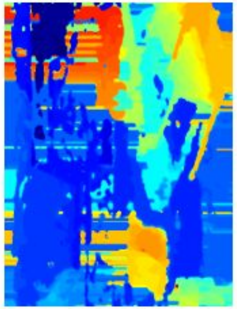
- Last, we apply 5x5 median and bilateral filter to get our final disparity map.

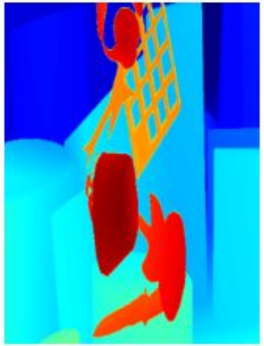
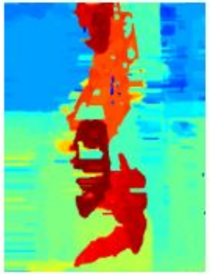
Result

- Synthetic

	0	1	2
L/R			
GT			
Result			
Error	1.70	1.99	2.61

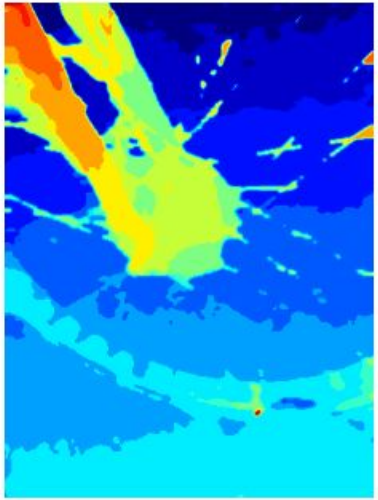
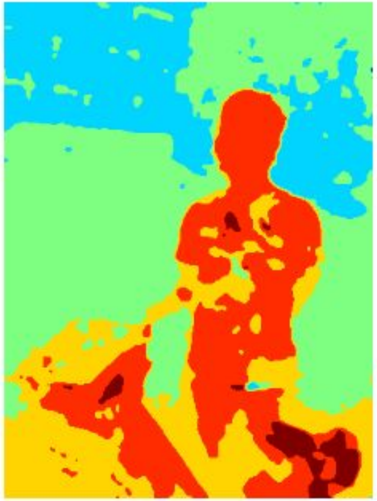
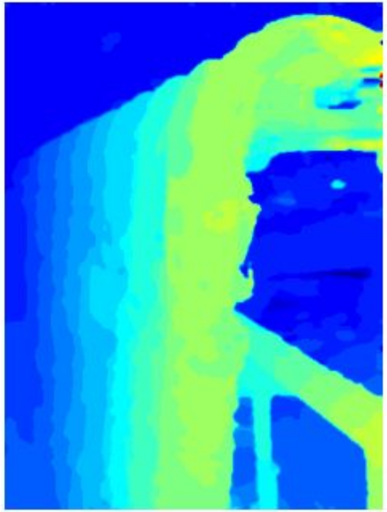
	3	4	5
L/R			
GT			
Result			
Error	3.94	3.34	4.10

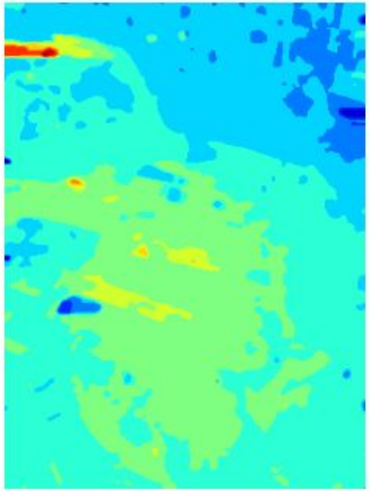
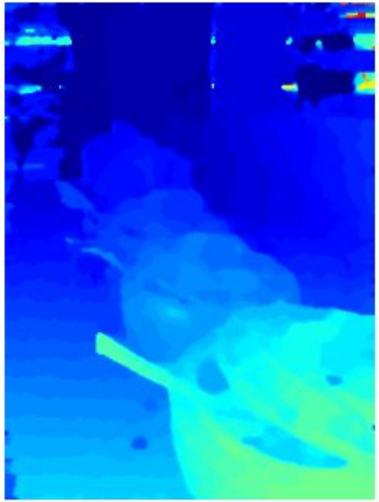
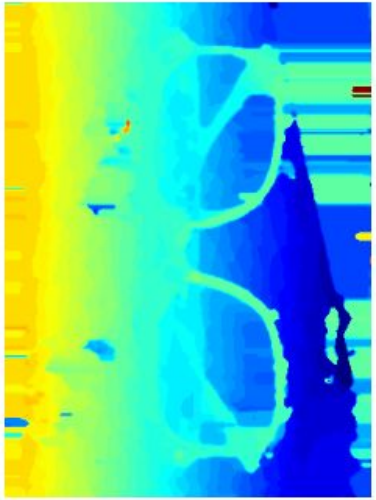
	6	7	8
L/R	 	 	
GT			
Result			
Error	1.65	3.47	3.56

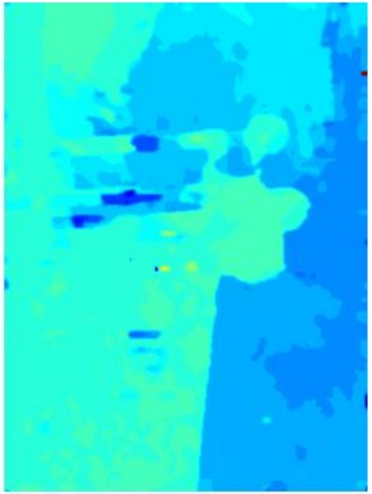
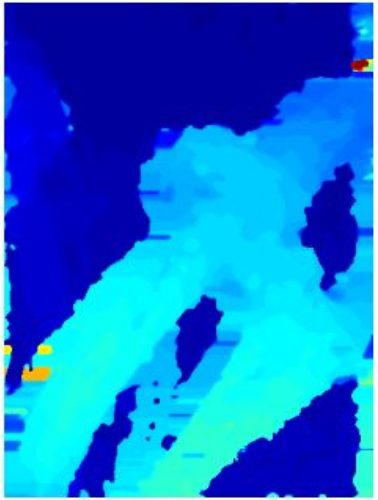
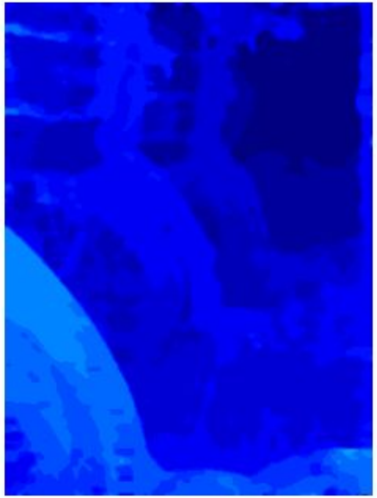
	9	Average Error: 2.93 Runtime of Image 0: 243 sec (Machine Spec: Intel Core i7 4-core CPU)
L/R		
GT		
Result		
Error	2.96	

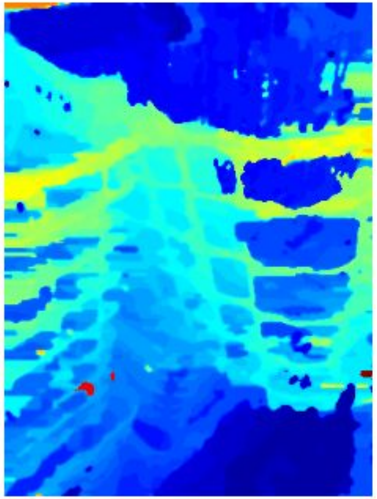
Result

- Real

	0		1		2	
L/R						
Result						

	3		4		5	
L/R						
Result						

	6		7		8	
L/R						
Result						

<table>
<tr><td></td><td colspan="2">9</td><td rowspan="3">Runtime of Image 0: 165 sec
(Machine Spec: Intel Core i7 4-core CPU)</td></tr>
<tr><td>L/R</td><td> </td><td> </td></tr>
<tr><td>Result</td><td colspan="2"></td></tr>
</table>

Reference

[1] Kang Zhang, Jiyang Li, Yijing Li, WeiDong Hu, Lifeng Sun, and Shiqiang Yang. Binary stereo matching. CoRR, abs/1402.2020, 2014.

[2] Michael Bleyer, Christoph Rhemann, and Carsten Rother. Patchmatch stereo - stereo matching with slanted support windows. In BMVC, January 2011.

[3] J. Zbontarˇ and Y. LeCun, “Stereo matching by training a convolutional neural

Thank you for your listening!