Accelerating Cost Volume Filtering Using Salient Subvolumes and Robust Occlusion Handling

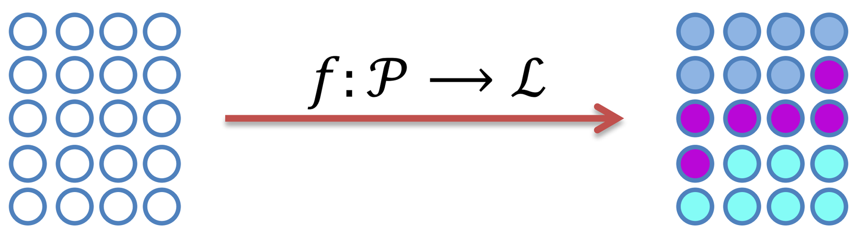
Mohamed A. Helala and Faisal Z. Qureshi {mohamed.helala,faisal.qureshi}@uoit.ca

http://www.vclab.ca

1.Introduction

- Given a set of images I with a set of pixels $\mathcal{P}$.

A pixel labeling problem $f:\mathcal{P} \longrightarrow \mathcal{L}$ assigns every pixel $(x,y) \in \mathcal{P}$ to a label $l \in \mathcal{L}$.



$\mathcal{L} = \{l_1, \ldots, l_m\}$

- Solutions should be spatially smooth, obey label costs and preserve edge discontinuities.
- Traditionally solved using Markov Random Fields (MRF's), however, Cost Volume Filtering (CF) [2] is a fast alternative.
- CF is slow for a large label space. So, our objective is to accelerate CF.

2.Cost Volume Filtering (CF)

- We are interested in pixel labeling for stereo where CF solves using the following steps:

1. Construct a cost volume $C(x,y,l)$.
2. Filter the cost volume.
3. Apply Winner-Takes-All.
4. Identify pixels with incorrect labels (gaps).
5. Post-process gaps using Row Filling (RF).
6. Refine using Weighted Median Filtering.

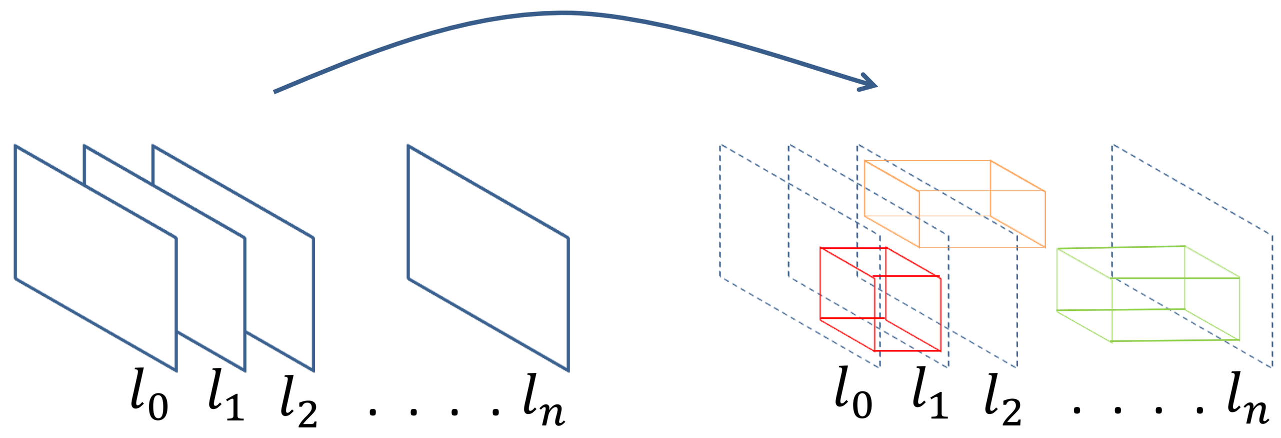
3.Accelerated Cost Volume Filtering (ACF)

- **Main Idea**: Identify salient subvolumes in the cost volume and restrict filtering to the selected subvolumes.

- Building salient subvolumes:

1. Find matched SIFT keypoints (x,y).
2. Identify locations (x,y,l') within the cost volume.
3. For each salient location (x,y,l'), define a window $b_l(x,y)$ centered on each disparity l for $\|l-l'\| \leq u$.

- ✓ **Pros**: 2.2 times speedup on Middlebury benchmark dataset, and up to 4 times on high resolution datasets.
- ✓ **Cons**: For some scenes, there exist a marginal increase in occluded pixels.



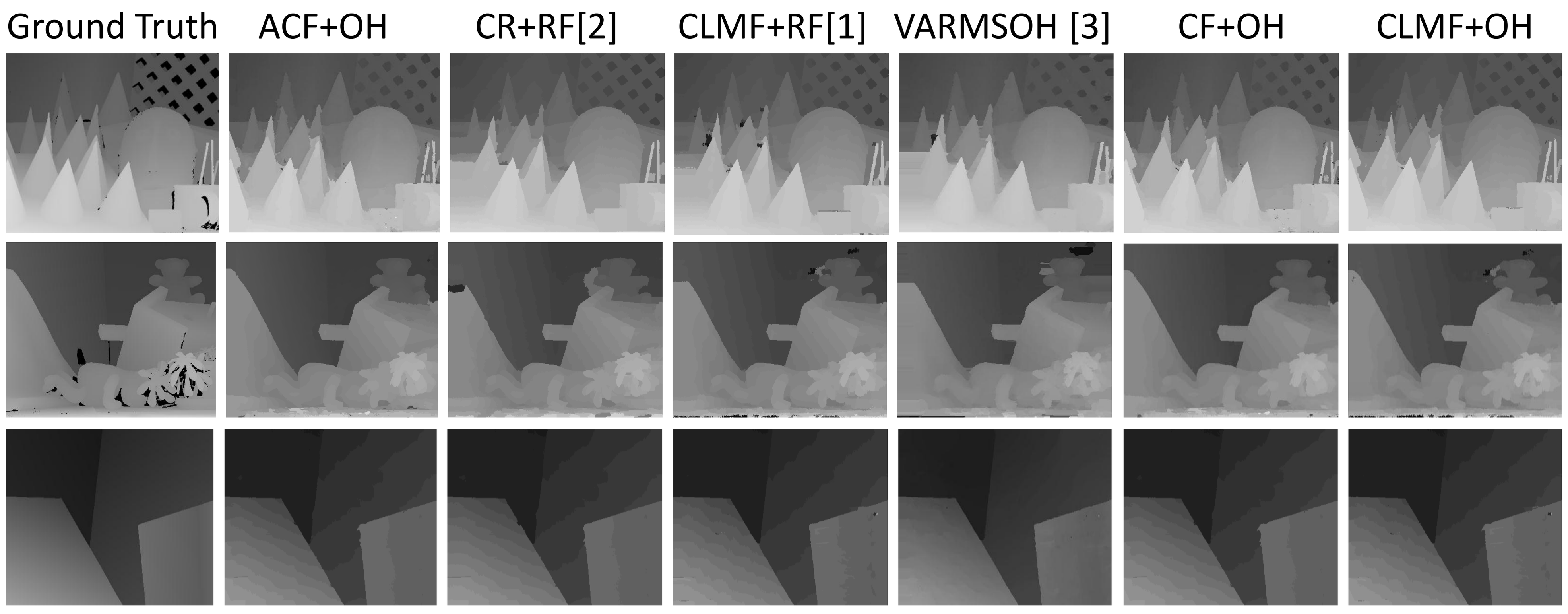
Disparity = 24.9 Disparity = 24.9

4.Occlusion Handling (OH)

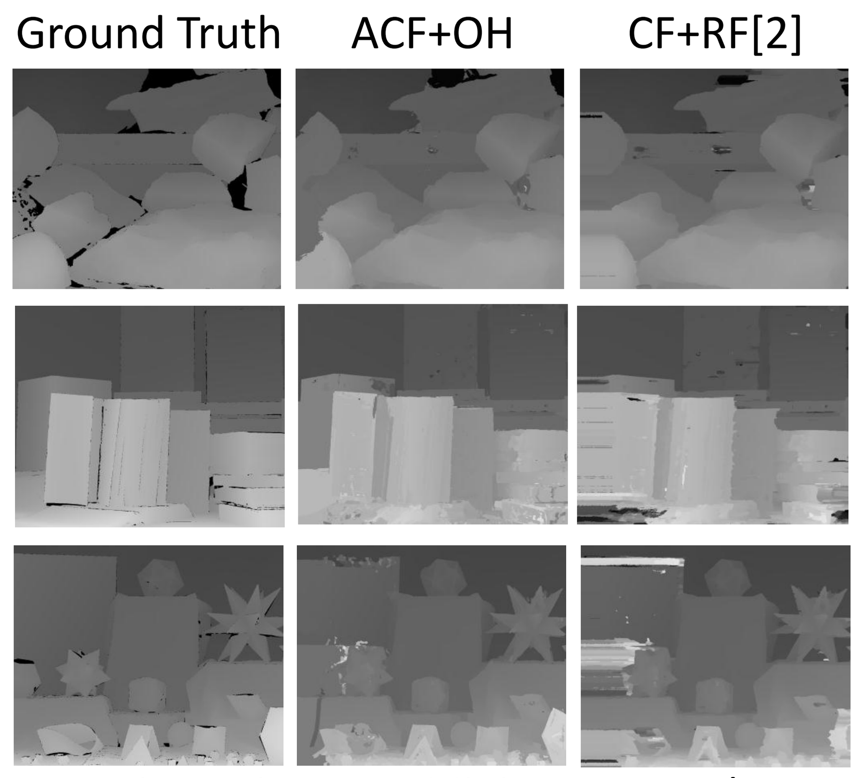
- **Main Idea**: Gaps in filtering methods have a large impact on the output accuracy. A better OH method than the ubiquitous RF strategy will improve the accuracy.

- We fill gaps while preserving edge discontinuities by building a set of compact superpixels and propagating labels using an inspired simulated annealing method.

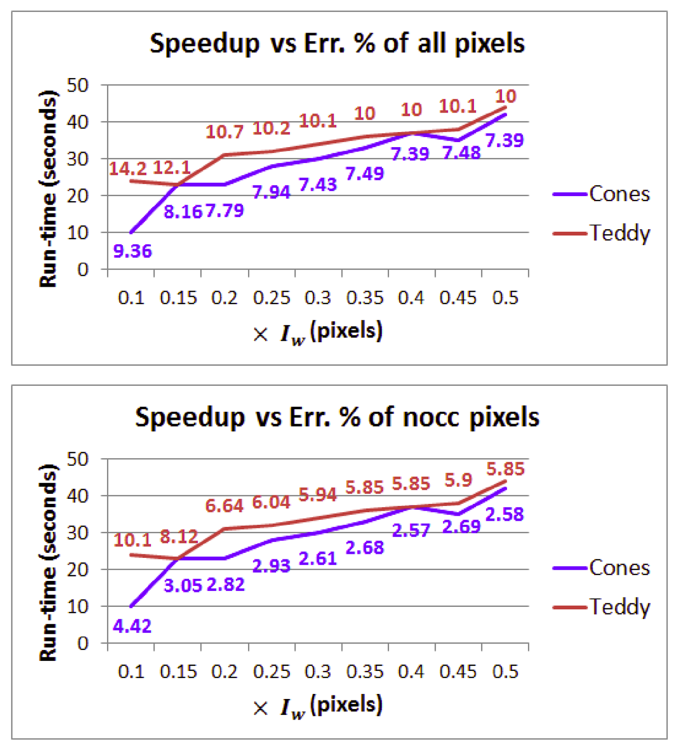
5. Experimental Results



Results on three Middlebury benchmark datasets.



Result on three Middlebury 2005/06 high resolution datasets.



Run-time vs. accuracy comparison.

Algorithm	Error thrreshold = 1		Error thrreshold = 0.5	
	Rank	% error	Rank	% error
CF+OH	**25**	5.22	30	12.9
CLMF+OH	38	5.14	66	16.9
ACF+OH ($r=.3$)	30	5.26	33	13
ACF+OH ($r=.2$)	39	5.45	37	13.3
CF+RF [2]	42	5.55	27	12.8
CLMF+RF [1]	37	**5.13**	64	16.7
ACF+RF($r=.3$)	64	5.99	45	13.4
ACF+RF($r=.2$)	60	5.92	42	13.6
VARMSOH [3]	116	8.17	**21**	**11.8**

Quantitative evaluation on Middlebury benchmark datasets.

Algorithm	Tsukuba			Venus			Teddy			Cones		
	nocc	all	disk	nocc	all	disk	nocc	all	disc	nocc	all	disk
CF+OH	**1.45**	**1.75**	7.37	**0.19**	**0.37**	2.24	5.85	**10**	16.1	2.6	7.41	7.31
CLMF+OH	2.39	2.69	6.53	0.26	**0.37**	2.23	**5.49**	10.7	**14.2**	2.46	**7.22**	7.10
ACF+OH ($r=.3$)	1.45	1.75	7.37	0.19	0.37	2.24	5.94	10.1	16.4	2.61	7.43	7.23
ACF+OH ($r=.2$)	1.45	1.75	7.37	0.19	0.37	2.24	6.64	10.7	16.3	2.82	7.79	7.74
CF+RF [2]	1.51	1.85	7.61	0.2	0.39	2.42	6.16	11.8	16	2.71	8.24	7.66
CLMF+RF [1]	2.46	2.78	**6.26**	0.27	0.38	**2.15**	5.50	10.6	**14.2**	**2.34**	7.82	**6.80**
ACF+RF ($r=.3$)	1.51	1.85	7.61	0.2	0.39	2.42	6.94	11.3	18.5	3.38	8.49	9.3
ACF+RF ($r=.2$)	1.51	1.85	7.61	0.2	0.39	2.42	6.96	11.1	17.1	3.66	9.06	9.8
VarMSOH [3]	3.97	5.23	14.9	0.28	0.76	3.78	9.34	14.3	20	4.14	9.91	11.4

Stereo Evaluation results on Middlebury benchmark using error threshold equals 1.

Algorithm	Average % occluded pixels		Run-time (seconds)	
	Standard	High Resolution	Standard	High Resolution
ACF(r=0.2)	14.2	-	16.117	-
ACF(r=0.3)	14.39	26.1	18.717	159.82
CF	13.6	26.9	28.2	505
RF	-	-	0.11	1.4
OH	-	-	0.131	0.2

A comparison of ACF and CF without any post-processing steps on Middlebury standard and high-resolution datasets. Runtimes for RF and OH post-processing steps are also provided.

6.Conclusions and Future Work

- This work develops accelerated cost volume filtering with occlusion handling for stereo disparity estimation.
- Our method outperforms state-of-the-art techniques: CF, CLMF and VARMSOH, on Middlebury datasets.
- We hope to apply our method to other discrete labeling problems, such as optical flow computation, etc.

7.References

1. Lu, J., Shi, K., Min, D., Lin, L., Do, M.: Cross-based local multipoint filtering. In CVPR (2012).
2. Hosni, A., Rhemann, C., Bleyer, M., Rother, C., Gelautz, M.: Fast cost-volume filtering for visual correspondence and beyond. PAMI 25 (2013).
3. Ben-Ari, R., Sochen, N.: Stereo matching with mumford-shah regularization and occlusion handling. PAMI 32 (2010).