MULTIPLE VIEW STEREO BY REFLECTANCE MODELING

Sujung Kim, Henrik Aanæs†, Anders Lindbjerg Dahl†, and Knut Conradsen†*

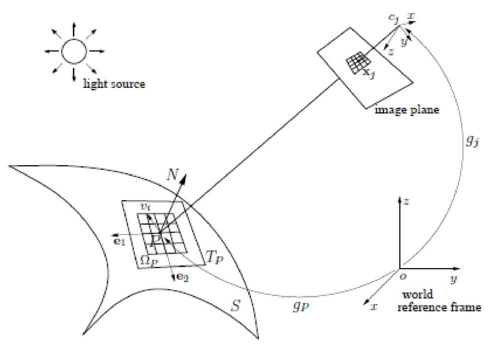
Korea Advanced Institute of Science and Technology (KAIST)*
Denmark Technical University (DTU)†

▪ Motivation

A massive effort has recently been put in multiple view stereo, and this effort have mainly focused on methods for optimization and regularization. The visual metric used have been sums of squared differences(SSD) and normalized cross correlation(NCC) between image pairs. These visual metrics are well suited for diffuse reflecting surfaces, but not for more complex reflecting surfaces with specularities. So we argue that for specular objects a visual metric should be founded on more observations than the degree of freedom(DoF) of the reflectance model. Based on this realization, we investigate how to construct a visual metric dealing with diffuse and specular objects considering SIFT methodology known as being superior to NCC.

▪ Visual metrics

1. Radiance tensor [1]
   - is proposed for a reflectance model
   - uses rank discrepancy as a visual metric

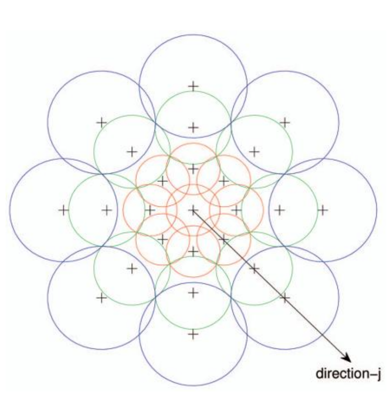


$$\mathbf{R}(\mathbf{x},\mathbf{n}) = [\mathbf{r}_1(\mathbf{x},\mathbf{n}) \quad \mathbf{r}_2(\mathbf{x},\mathbf{n}) \quad \cdots \quad \mathbf{r}_n(\mathbf{x},\mathbf{n})]$$

The singular values of $\mathbf{R}(\mathbf{x},\mathbf{n})$: $\{\sigma_1, \sigma_2, \ldots, \sigma_n\}$

→ **Visual metric :** $J(\mathbf{x},\mathbf{n}) = \sum_{i=3}^{n} \sigma_i^2$

2. DAISY descriptor [2]
   - has similar characteristics to SIFT
   - is computationally efficiency



▪ Visual metrics for specular surfaces

1. Daisy tensor

$$\mathbf{D}(\mathbf{x},\mathbf{n}) = [\mathbf{d}_1(\mathbf{x},\mathbf{n}) \quad \mathbf{d}_2(\mathbf{x},\mathbf{n}) \quad \cdots \quad \mathbf{d}_n(\mathbf{x},\mathbf{n})]$$

The singular values of $\mathbf{D}(\mathbf{x},\mathbf{n})$: $\{\varsigma_1, \varsigma_2, \ldots, \varsigma_n\}$

→ **Visual metrics :** $D_1(\mathbf{x},\mathbf{n}) = \sum_{i=2}^{n} \varsigma_i^2$ **and** $D_2(\mathbf{x},\mathbf{n}) = \sum_{i=3}^{n} \varsigma_i^2$

2. Minimal vs all

☼ **The size of minimal sets is one plus the DoF of the model**

In 2 DoF case, $\{i, j, k\} \in C_3 \rightarrow [\mathbf{d}_i(\mathbf{x},\mathbf{n}) \quad \mathbf{d}_j(\mathbf{x},\mathbf{n}) \quad \mathbf{d}_k(\mathbf{x},\mathbf{n})]$

The rank discrepancy of a single triplet is denoted as

$$\Gamma_{ijk}^3(\mathbf{x},\mathbf{n}) = \varsigma_3^2 = \min_{\mathbf{v}_1,\mathbf{v}_2} \sum_{m\in\{i,j,k\}} \left\| \mathbf{d}_m(\mathbf{x},\mathbf{n}) - [\mathbf{v}_1 \mathbf{v}_2][\mathbf{v}_1 \mathbf{v}_2]^T \mathbf{d}_m(\mathbf{x},\mathbf{n}) \right\|_2^2$$

Then, the 2 DoF minimal visual metric is $M_2(\mathbf{x},\mathbf{n}) = \sum_{\{i,j,k\}\in C_3} \Gamma_{ijk}^3(\mathbf{x},\mathbf{n})$

In a similar way, the 1 DoF minimal visual metric is $M_1(\mathbf{x},\mathbf{n}) = \sum_{\{i,j\}\in C_2} \Gamma_{ij}^2(\mathbf{x},\mathbf{n})$

3. Regularization

$$D_{1.5}(\mathbf{x},\mathbf{n}) = (1-\beta) D_2(\mathbf{x},\mathbf{n}) + \beta D_1(\mathbf{x},\mathbf{n}) = \beta\varsigma_2^2 + \sum_{i=3}^{n} \varsigma_i^2$$

$$M_{1.5}(\mathbf{x},\mathbf{n}) = \sum_{\{i,j,k\}\in C_3} \Gamma_{ijk}^{2.5}(\mathbf{x},\mathbf{n}) \quad \text{where} \quad \Gamma_{ijk}^{2.5}(\mathbf{x},\mathbf{n}) = \beta\varsigma_2^2 + \varsigma_3^2$$

4. Investigated visual metrics

: J^{11}, J^{31} **and** $D_1, D_{1.5}, D_2, M_1, M_{1.5}, M_2$

✓ **If 1 DoF, 2 DoF or a regularized alternative should be used ?**

✓ **If the metric should be based directly on all relevant images or a combination of minimal subsets ?**

5. Local minima of visual metrics

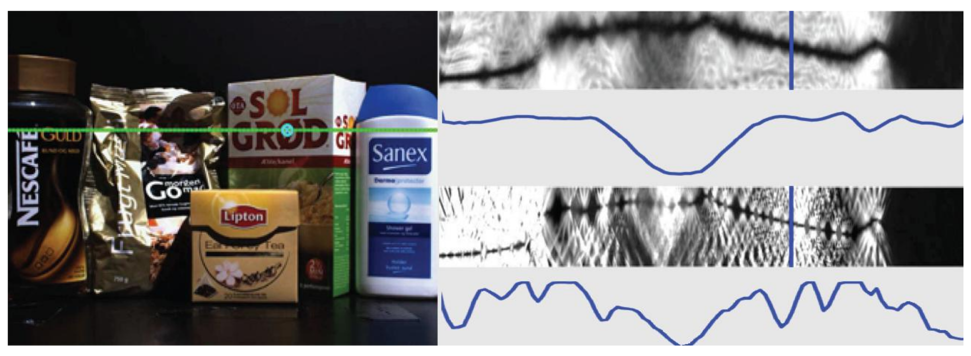


Fig. 1. The cross-sections of the visual metrics D_2 and J^{11}: Slices are extracted from the horizontal green line on the left image[3] and depth is equal to the vertical blue line on the slice images. Note that D_2 has much fewer local minima than J^{11}

▪ Experimental results

Fig. 2. Some scenes of our investigation: various daily objects (strongly reflected region, widely reflected region, almost diffused region) are included.

VM	Scene 1		Scene 2		Scene 3		Scene 4		Scene 5		Scene 6	
	mean	std.	mean	std.	mean	std.	mean	std.	mean	std.	mean	std.
D_1	9.47	14.98	17.84	32.82	7.39	15.77	6.39	15.34	4.86	8.02	15.49	34.89
D_2	7.29	13.52	16.61	31.26	6.27	12.14	**5.55**	**13.73**	4.67	7.92	13.24	33.32
$D_{1.5}$	8.36	14.30	**12.58**	**26.36**	7.16	15.34	6.24	15.45	4.62	6.99	15.30	35.83
M_1	6.23	8.04	14.86	32.67	5.47	**7.20**	5.99	16.45	4.29	5.72	10.27	23.68
M_2	**5.64**	**7.38**	13.07	29.48	5.52	7.36	5.79	15.42	**4.26**	5.70	**9.06**	24.45
$M_{1.5}$	6.39	8.44	14.08	31.03	**5.43**	7.23	5.78	14.90	4.28	**5.59**	9.40	**22.33**
J^{11}	25.47	40.66	78.99	84.75	30.06	48.61	25.95	60.15	12.89	28.48	67.96	79.95
J^{31}	20.24	37.46	39.30	56.43	19.21	42.36	12.10	34.45	7.66	18.26	46.13	73.02

Table 1. Average reconstruction errors and standard deviation(in mm). If we assume a few hundred *independent* observations the main differences between the means are significant.

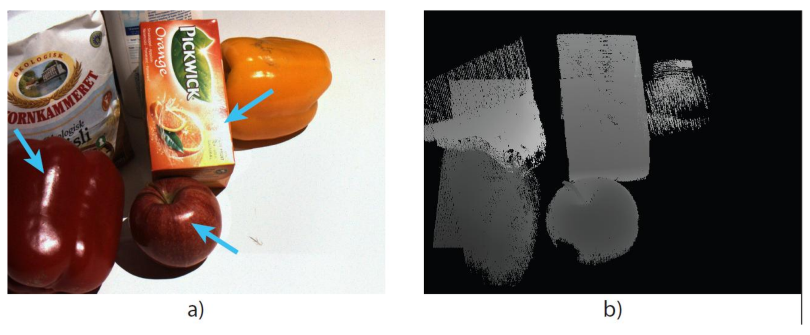
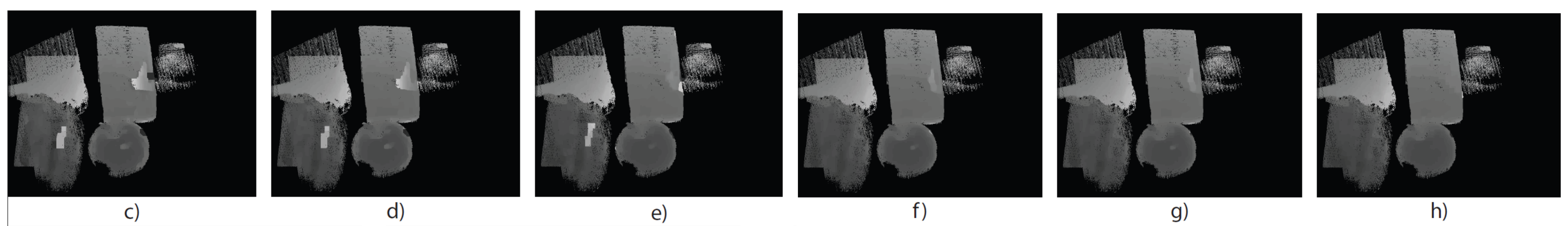
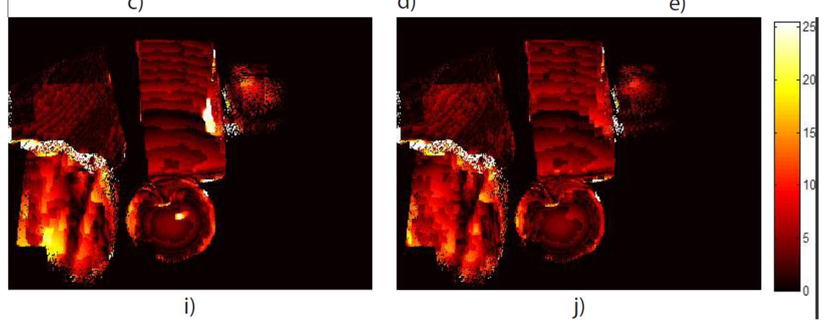


Fig. 3. The reconstruction results of scene 1: a) sample image, b) ground truth, c)-h) reconstructed depth maps for the following visual metrics D_1 , $D_{1.5}$, D_2, M_1, $M_{1.5}$, M_2 respectively, i) and j) the reconstruction error of M1 and M2 respectively, i.e. (b-f) and (b-h)





***In our experiments, for obtaining the performance of visual metrics, we computed depth-maps via the alpha-expansion algorithm of [4].**

✓ Discussion

- The daisy based visual metrics outperform the raw based visual metrics
- No single visual metric performs best for all scenes
- A minimal case(2DoF, regularized version) is preferable.

▪ Future work

- Compare it with the state of the art lambertian multiview stereo methods
- Investigate the effects when there are plenty of available images

※ References

[1] Jin, H., Soatto, S., Yezzi, A.: *Multi-view stereo reconstruction of dense shape and complex appearance*. IJCV (2005) 175-189

[2] Tola, E., Lepetit, V., Fua, P.: *Daisy: an efficient dense descriptor applied to wide-baseline stereo*. IEEE PAMI (2010) 815-830

[3] Aanæs, H., Dahl, A., Steenstrup Pedersen, K.: *Interesting interest points*. IJCV (2011) 1-18

[4] Boykov, Y., Veksler, O., Zabih, R.: *Fast approximate energy minimization via graph cuts*. IEEE PAMI (2001) 1222-1239