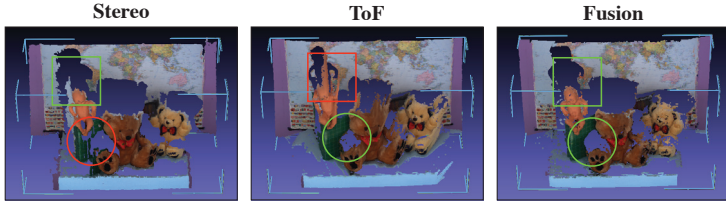


RELIABLE FUSION OF TOF AND STEREO DEPTH DRIVEN BY CONFIDENCE MEASURES

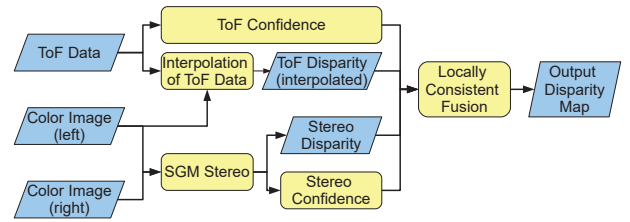


Giulio Marin, Pietro Zanuttigh - University of Padova, Italy
Stefano Mattoccia - University of Bologna, Italy

Motivation



Proposed Method



ToF Disparity Map and Confidence

Disparity map

ToF data are projected on the left color camera, converted to disparity and interpolated combining cross bilateral filtering and color segmentation.

Confidence

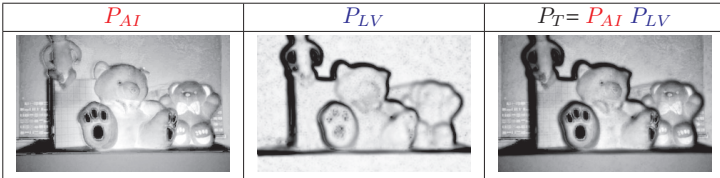
Both **photometric** and **geometric** properties of the scene are used for the confidence estimation.

- P_{AI} : amplitude and intensity of ToF signal.

$$\sigma_z = \frac{c}{4\pi f_{mod}} \frac{\sqrt{I/2}}{A} \Rightarrow \sigma_d = bf \frac{\sigma_z}{z^2 - \sigma_z^2} \quad (1)$$

- P_{LV} : accounts for local depth variance.

$$D_i^{TOF} = \frac{1}{|\mathcal{N}(p_i^{TOF})|} \sum_{j \in \mathcal{N}(p_i^{TOF})} |z_i - z_j| \quad (2)$$



Stereo Disparity Map and Confidence

Disparity map

A high resolution disparity map is inferred by global or semi-global stereo vision algorithms.

Confidence

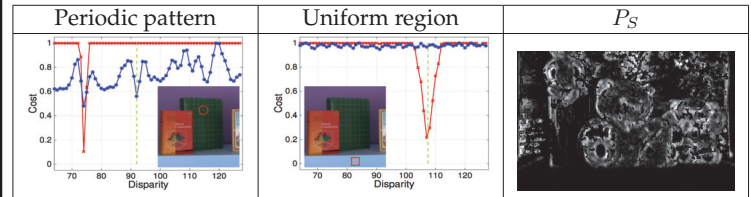
The stereo confidence map is a function of both the **point-wise local cost function**, and of the **comparison of the local with the globally optimized cost**.

$$P_S = \frac{\Delta C^l}{C_1^l} \left(1 - \frac{\min\{\Delta d^l, \gamma\}}{\gamma} \right) \left(1 - \frac{\min\{\Delta d^g, \gamma\}}{\gamma} \right) \quad (3)$$

- $C_1 = C(d_1) = \min_d C(d)$, C_2 s.t. $|d_2 - d_1| \leq 1$

- l : Local cost, g : Global cost

- $\Delta C^l = C_2^l - C_1^l$, $\Delta d^l = |d_2^l - d_1^l|$, $\Delta d^g = |d_1^g - d_1^l|$



Data Fusion

Enforce local consistency and weight the two contributions exploiting the confidence information.

$$\Omega'_f(d) = \sum_{g \in A} \left(P_T(g) \mathcal{P}_{f,g,T}(d) + P_S(g) \mathcal{P}_{f,g,S}(d) \right) \quad (4)$$

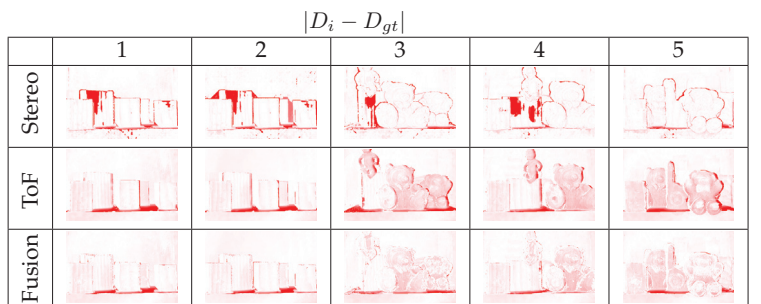
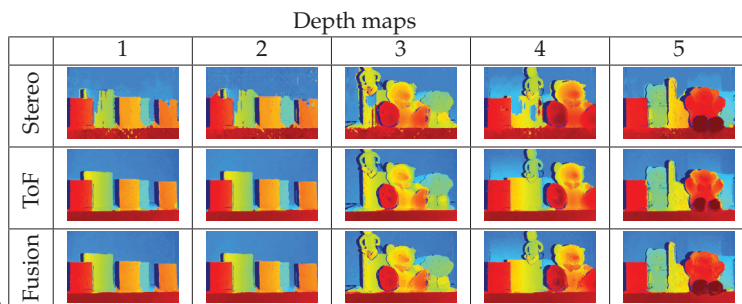
where $\mathcal{P}_{f,g}(d) = e^{-\frac{\Delta_{f,g}}{\gamma_s}} \cdot e^{-\frac{\Delta_{f,g}^\psi}{\gamma_c}} \cdot e^{-\frac{\Delta_{f',g'}}{\gamma_c}} \cdot e^{-\frac{\Delta_{g,g'}}{\gamma_t}}$

- f, g and f', g' : points in the left and right image
- Δ : spatial proximity; $\Delta^\psi, \Delta^\omega$: color similarity

Experimental Results - Comparison

Scene	1	2	3	4	5	Avg.
ToF Int.	9.83	10.33	14.43	8.68	15.12	11.67
Stereo	19.17	27.83	18.06	25.52	11.49	20.42
Fusion	7.40	9.33	6.92	6.30	8.39	7.67
[1]	7.43	9.27	12.60	7.99	13.01	10.06
[2]	8.49	9.92	11.44	9.88	15.19	10.98
[3]	9.04	10.04	13.04	9.52	14.03	11.13
[4]	10.98	13.19	9.83	13.93	13.10	12.21
Ideal	2.50	2.60	3.22	2.42	3.16	2.78

Experimental Results - Images



[1] Dal Mutto, C., Zanuttigh, P., Mattoccia, S., Cortelazzo, G.M.: Locally consistent tof and stereo data fusion. In: Workshop on Consumer Depth Cameras for Computer Vision (ECCV Workshop), 2012.
[2] Yang, Q., Yang, R., Davis, J., Nister, D.: Spatial-depth super resolution for range images. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2007.
[3] Zhu, J., Wang, L., Yang, R., Davis, J.: Fusion of time-of-flight depth and stereo for high accuracy depth maps. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2008.
[4] Dal Mutto, C., Zanuttigh, P., Cortelazzo, G.M.: Probabilistic tof and stereo data fusion based on mixed pixels measurement models. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2015.