Reliable Fusion of ToF and Stereo Depth Driven by Confidence Measures

Additional material

This document contains some additional material for the paper *Reliable Fusion of ToF and Stereo Depth Driven by Confidence Measures* (submission #1430). In particular it contains additional data about the experimental results that was not possible to place in the paper due to the length constraints. The images in this document are also provided as separate files in the *images* folder. Notice that the dataset we used is the same of [1] (the input data can be downloaded from the webiste of the authors of that work). However, in [1] the authors propose a method to combine ToF and stereo data from the point of view of the ToF camera rather than from one of the stereo cameras and also their algorithm works in the depth field, rather than disparity as we do. The output of our algorithm is a disparity map from the left camera viewpoint thus the error metrics are evaluated in the disparity field and are not directly comparable with the results presented in [1] (in order to compare with [1] we warped their data to the left camera viewpoint).

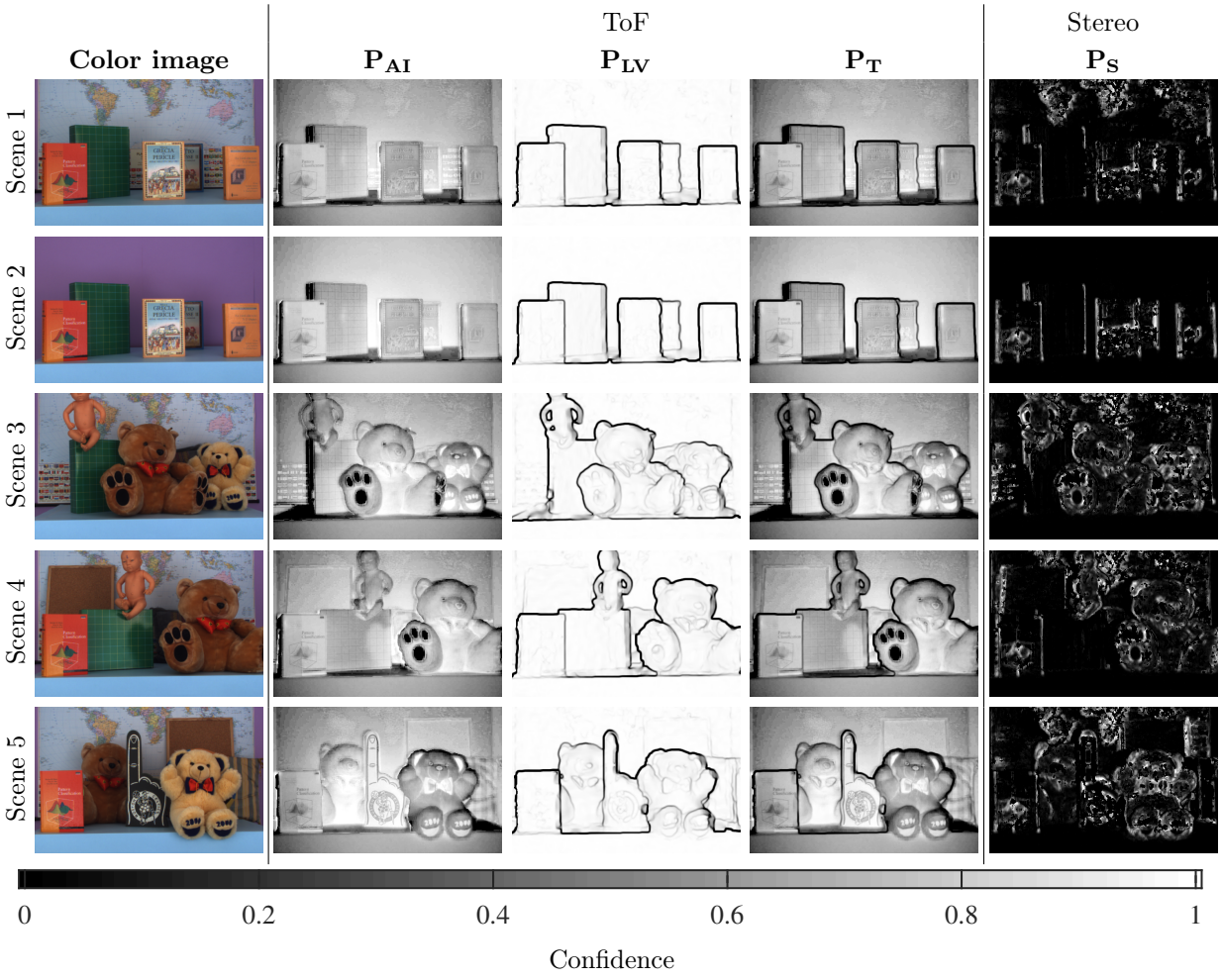
In this document we show:

- the confidence maps produced by the methods proposed in Sections 4 and 5 of the paper. This document contains both the ToF and stereo system confidence maps on the considered dataset;
- the disparity maps produced by the ToF upsampling, by the stereo vision system, by the proposed fusion algorithm and the ones produced by the other methods compared in Section 7 in the paper ([1, 2, 4, 5])
- the error maps of each method, both in terms of absolute differences and of mean squared error.

Contents

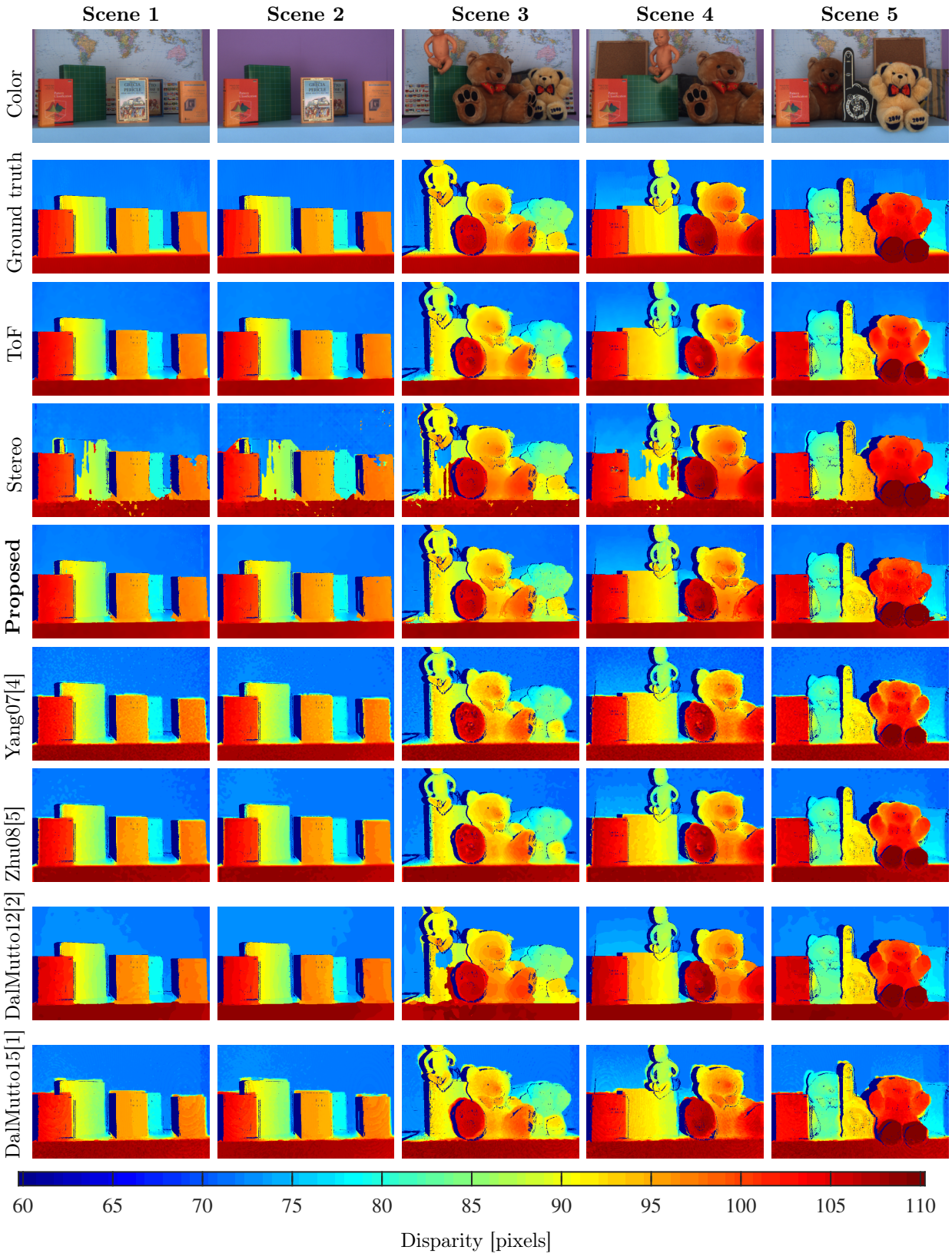
1 Confidence maps

For each scene of the dataset that we used for the comparison, we show the confidence maps associated to ToF and stereo data. The first column shows the reference color images, the second, third and fourth columns show the confidence maps associated to the ToF and the last column shows the confidence maps of stereo data. For ToF data, we show the confidence from amplitude and intensity values P_{AI}, the confidence from local variance P_{LV} and their product $P_T = P_{AI}P_{LV}$. The last column shows the confidence of the stereo system, i.e., P_S. As shown in the color map below, dark values correspond to low confidence and bright values correspond to higher confidence values.



2 Disparity maps

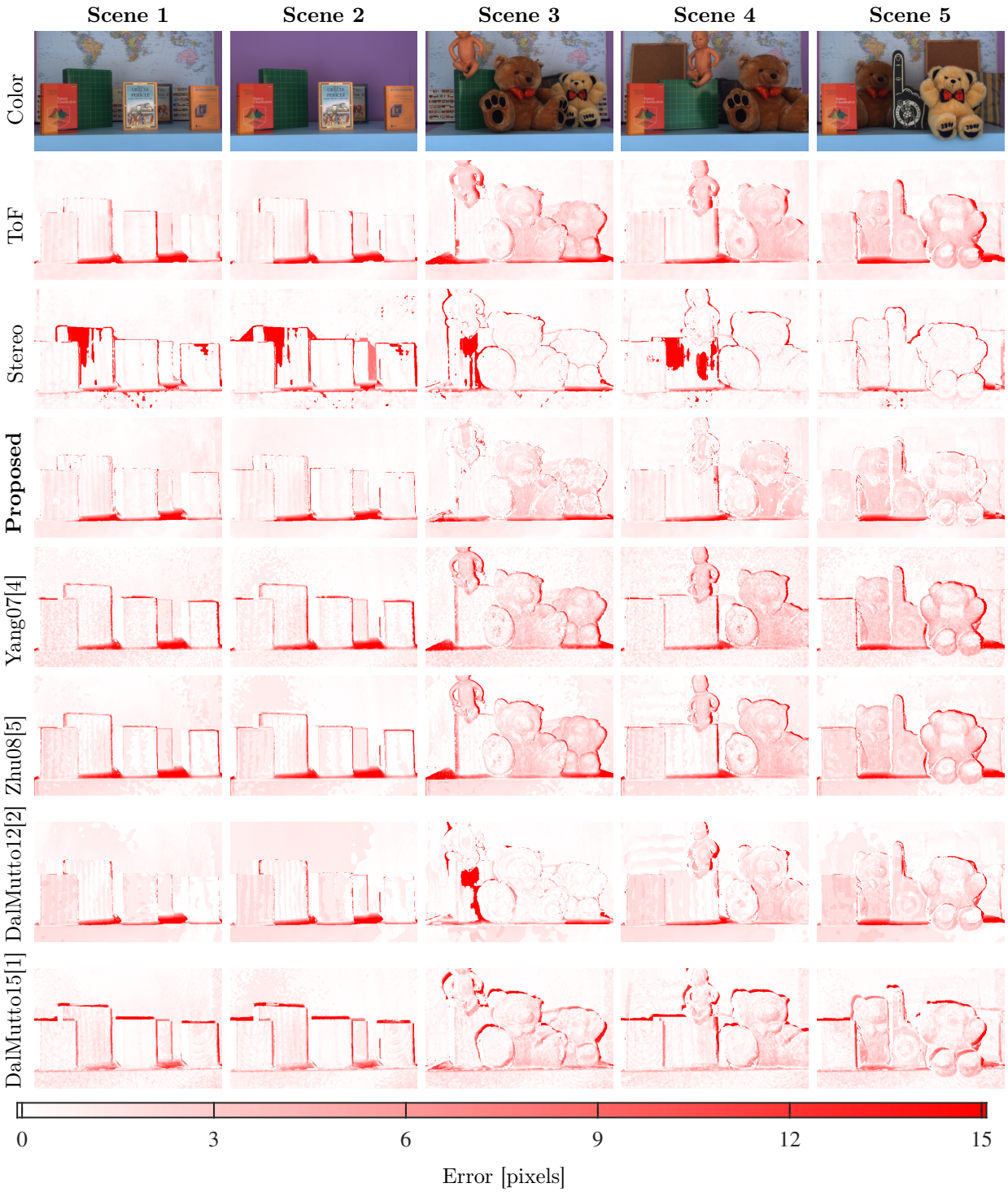
The next table shows a comparison of the disparity maps produced by all the competing algorithms, as well as the intermediate steps of our proposed method, i.e., the upsampled and interpolated map from ToF data and the disparity computed with the SGM [3] algorithm for the stereo system. The color bar at the bottom shows the meaning of the colors, i.e., disparity values from 60 to 110 pixels. Dark blue points are the ones for which disparity values are not available.


Scene 1
Scene 2
Scene 3
Scene 4
Scene 5
Color
Ground truth
ToF
Stereo
Proposed
Yang07[4]
Zhu08[5]
DalMutto12[2]
DalMutto15[1]
60
65
70
75
80
85
90
95
100
105
110
Disparity [pixels]

3 Error maps

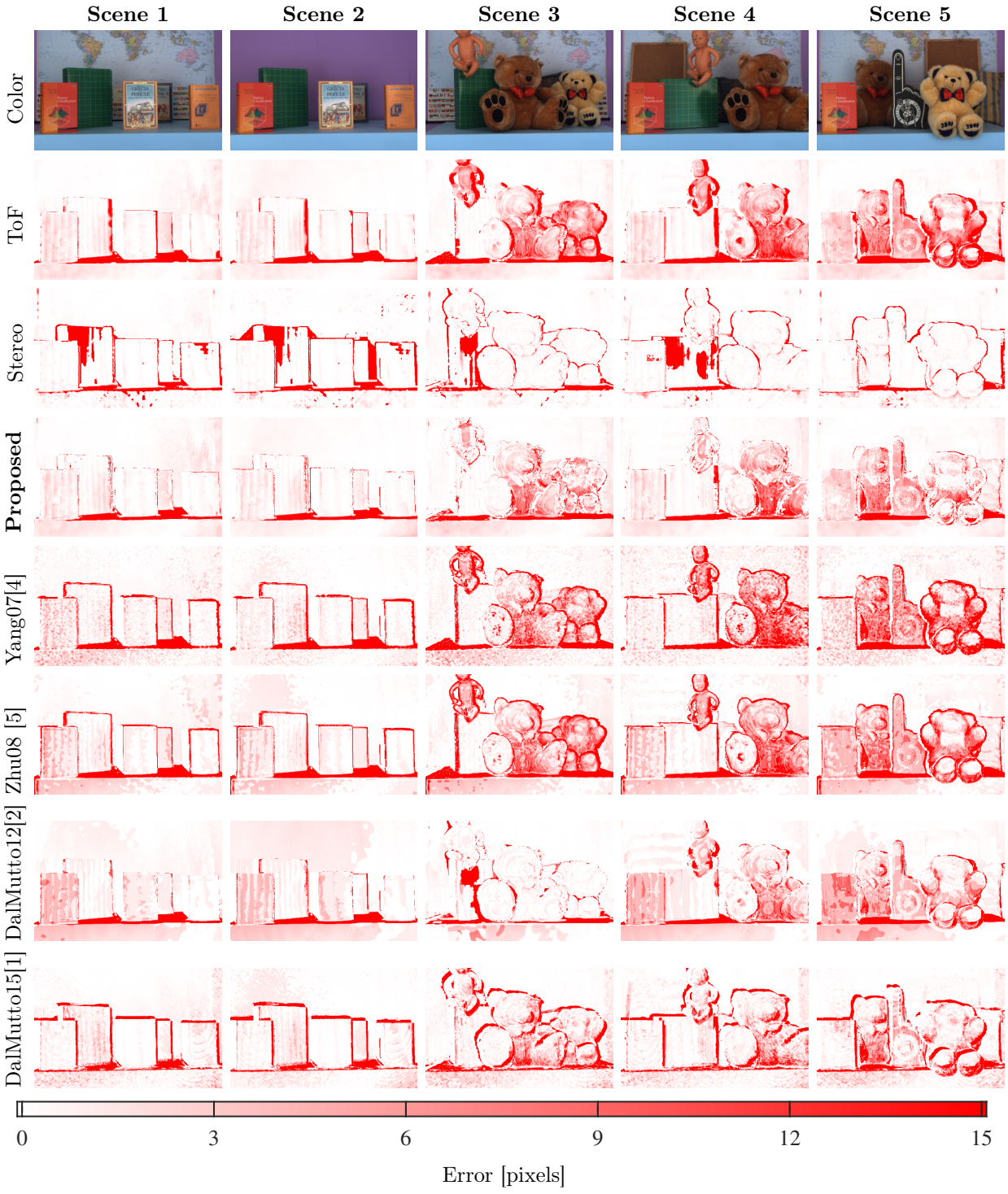
3.1 Absolute difference

The images show the absolute difference between the output disparity maps and the ground truth, i.e., $|D_i - D_{GT}|$, where D_i is the considered disparity map i for the evaluation, and D_{GT} is the ground truth.



3.2 Mean squared error

The images show the mean squared error between the output disparity maps and the ground truth, i.e., $|D_i - D_{GT}|^2$, where D_i is the considered disparity map i for the evaluation, and D_{GT} is the ground truth. With respect to the absolute difference, the MSE penalizes more larger errors and less errors smaller than 1 pixel.



References

[1] C. Dal Mutto, P. Zanuttigh, and G.M. Cortelazzo. Probabilistic tof and stereo data fusion based on mixed pixels measurement models. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 37(11):2260–2272, 2015.

[2] C. Dal Mutto, P. Zanuttigh, S. Mattoccia, and G.M. Cortelazzo. Locally consistent tof and stereo data fusion. In *ECCV Workshop on Consumer Depth Cameras for Computer Vision*, pages 598–607. Springer, 2012.

[3] Heiko Hirschmuller. Stereo processing by semiglobal matching and mutual information. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 2008.

[4] Q. Yang, R. Yang, J. Davis, and D. Nister. Spatial-depth super resolution for range images. In *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*, pages 1–8, 2007.

[5] J. Zhu, L. Wang, R. Yang, and J. Davis. Fusion of time-of-flight depth and stereo for high accuracy depth maps. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, 2008.