Supplementary for "Matching in the Dark: A Dataset for Matching Image Pairs of Low-light Scenes"

Wenzheng Song[1] Masanori Suganuma[1,2] Xing Liu[1]
Noriyuki Shimobayashi[1] Daisuke Maruta[3] Takayuki Okatani[1,2]
[1]GSIS, Tohoku University [2]RIKEN Center for AIP [3]Socionext Inc.
{song,suganuma,ryu,nshimobayashi,okatani}@vision.is.tohoku.ac.jp
maruta.daisuke@socionext.com

A. Distinction from Existing Datasets

Figure 7 shows example images of our dataset and RobotCar [33]. Most of our images are darker than the darkest one of RobotCar. The standard raw image processing yields mostly black images from them. Nevertheless, one can derive sufficient info from their RAW signals, when treating them properly.

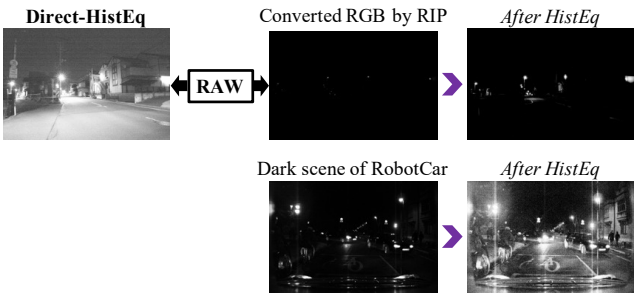

Figure 7. Comparison between RobotCar [33] and our dataset.

B. Performance Comparison to a Manual Adjustment

Figure 8 shows the results of using SuperPoint with the three image-enhancing methods and the result obtained by using SuperPoint on a manually converted 8-bit image from the same RAW image. To be specific, we manually chose the range of 14-bit RAW signal and converted it into an 8-bit image. The values of 'dR' and 'dT' indicate the rotation and translation errors for each method. It is observed that the manual method yields significantly better results than others and indicates that there is stil much room for improvement.

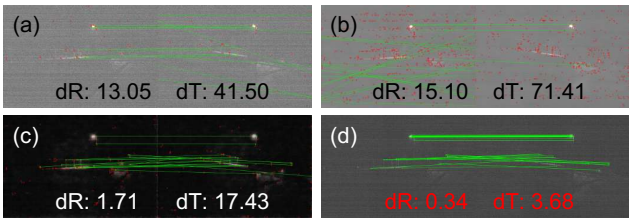

Figure 8. Matching results of **SP** with (a) **Direct-HistEq**, (b) **Direct-BM3D**, (c) **SID**, and (d) Images obtained by manual adjustment of brightness range in 14-bits RAW signals.

C. All Samples of Scene Images in the Dataset

Figures 9 and 10 show all samples of indoor and outdoor scenes in our dataset, respectively. All images are obtained from the long exposure RAW-format images by the standard RIP.

D. More Results of Image Matching

Figure 11 shows the normalized number N_τ of the exposure settings for which the estimation error is lower than threshold τ averaged over 54 *outdoor* scenes.

Figure 12 shows the average angular errors of the camera pose estimated by the compared 88 methods (i.e., eight image enhancers with eleven image matching methods) over all scenes for each of the 6×8 exposure settings.

E. Visualization of Matching Results

Figure 13 and 14 show examples of the visualization of the matching results by the 88 methods for an indoor and an outdoor scene, respectively.

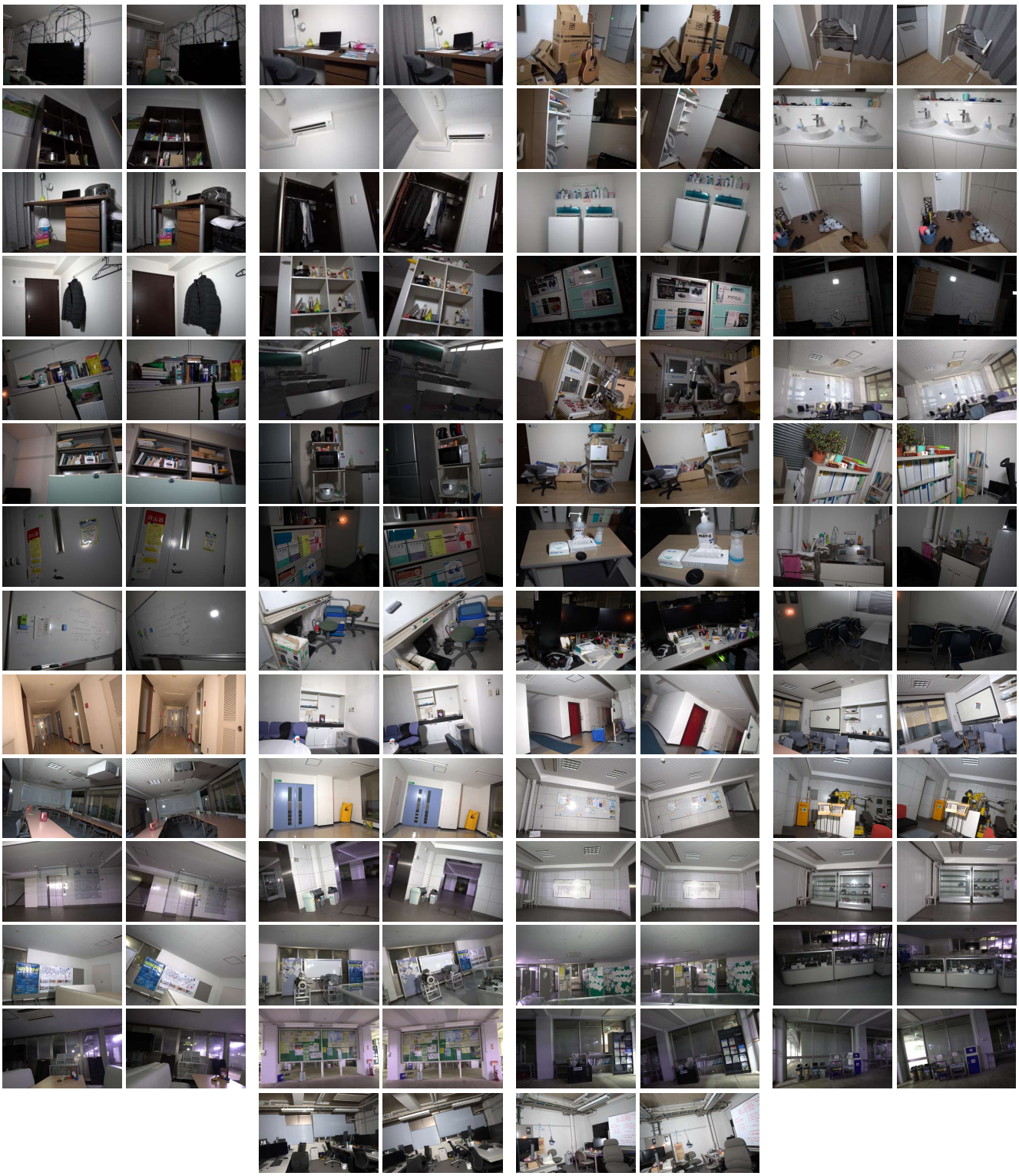

Figure 9. Samples of all image pairs (long exposure versions) of the indoor scenes.

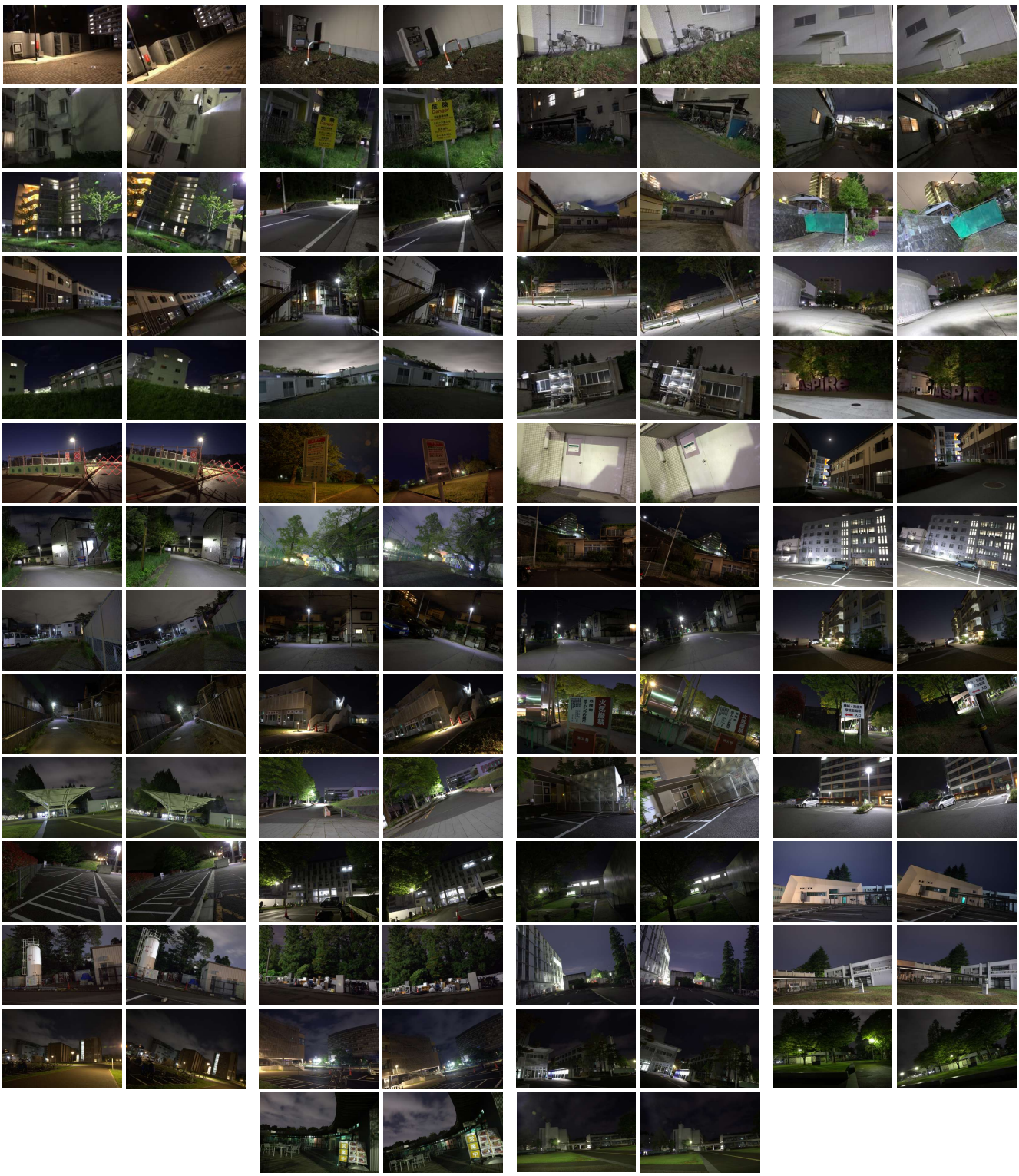

Figure 10. Samples of all image pairs (long exposure versions) of the outdoor scenes.

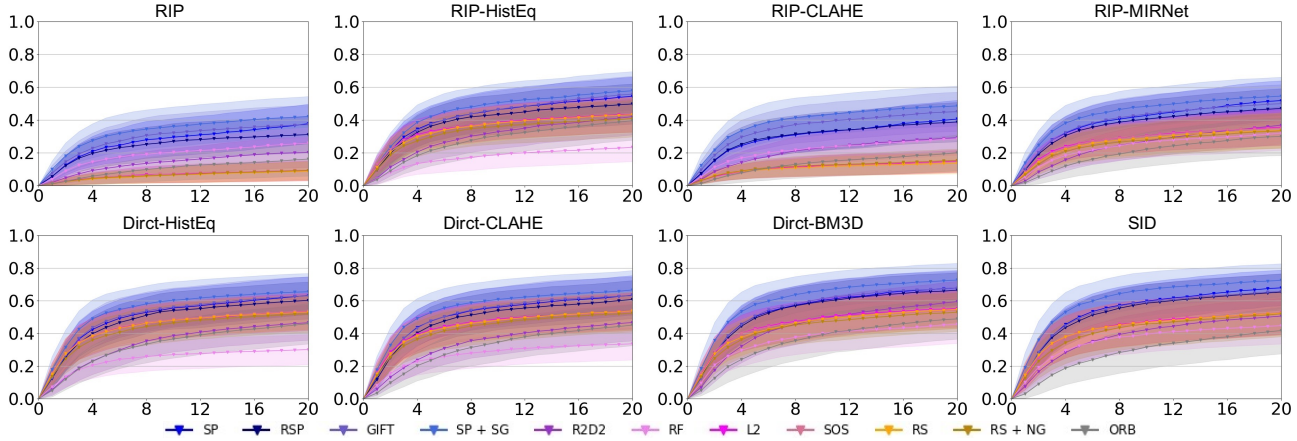

Figure 11. The normalized number N_τ of the exposure settings (the vertical axis) for which the estimation error of each method is lower than threshold τ (the horizontal axis). Each panel shows the means and standard deviations over 54 *outdoor* scenes for the eleven image matching methods for an image-enhancing method.

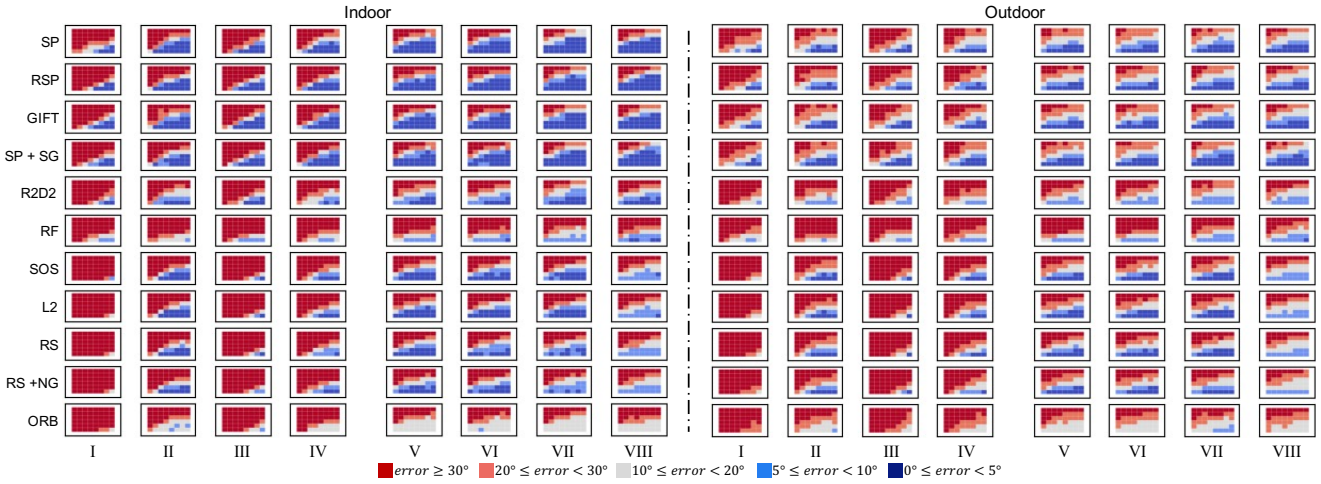

Figure 12. Average angular errors of the camera pose estimated by the 88 methods (i.e., eight image enhancers with eleven image matching methods) over all the 54 scenes for each of the 6×8 exposure settings. (I) **RIP**. (II) **RIP-HistEq**. (III) **RIP-CLAHE**. (IV) **RIP-MIRNet**. (V) **Direct-HistEq**. (VI) **Direct-CLAHE**. (VII) **Direct-BM3D**. (VIII) **SID**.

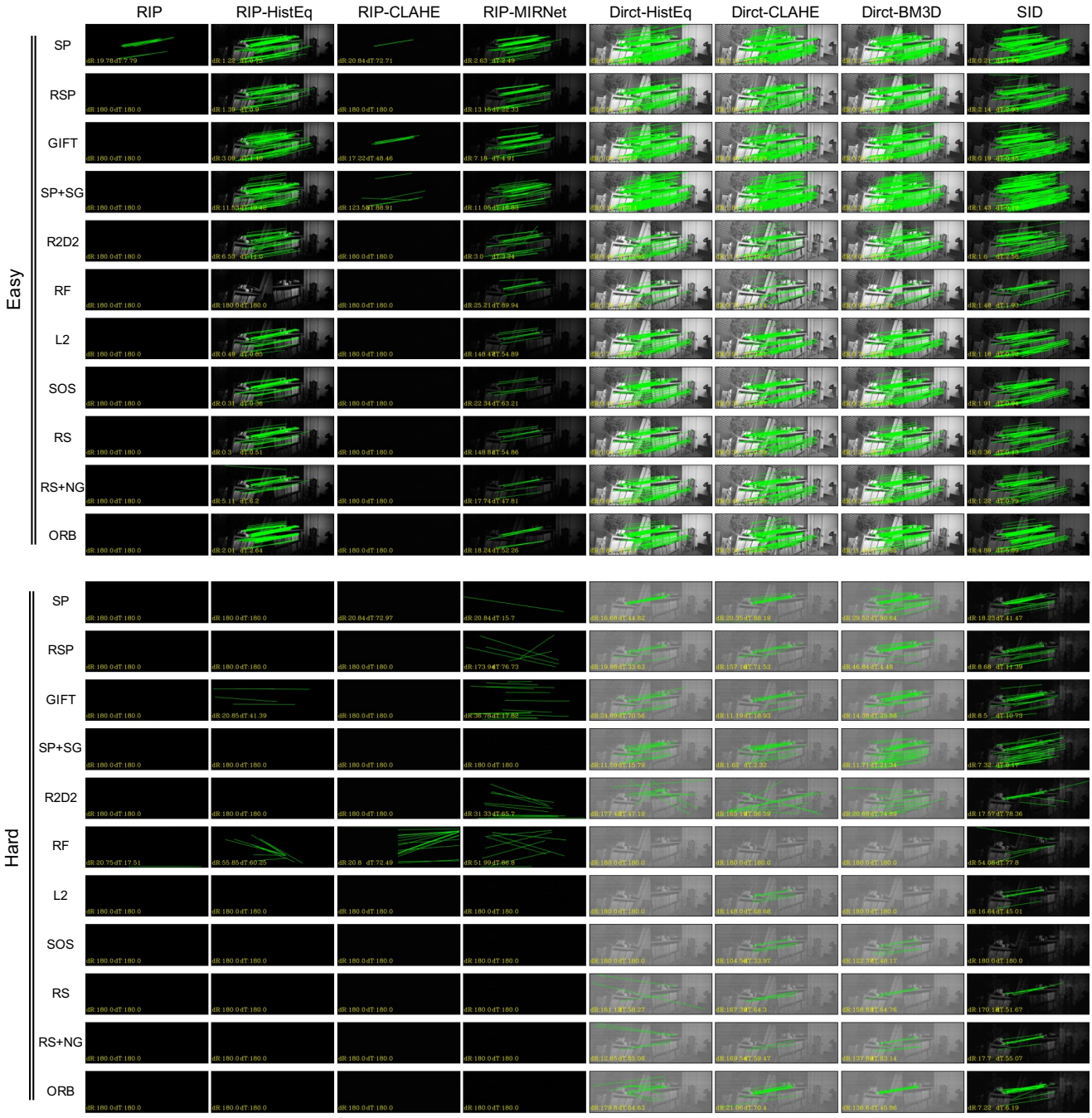

Figure 13. Visualization of the matching results for one of the 54 indoor scenes. Point correspondences judged as inliers are shown in green lines. The combination of eleven matching methods and the eight image enhancing methods are applied to two image pairs with different levels of exposure (i.e., 'Easy' and 'Hard').

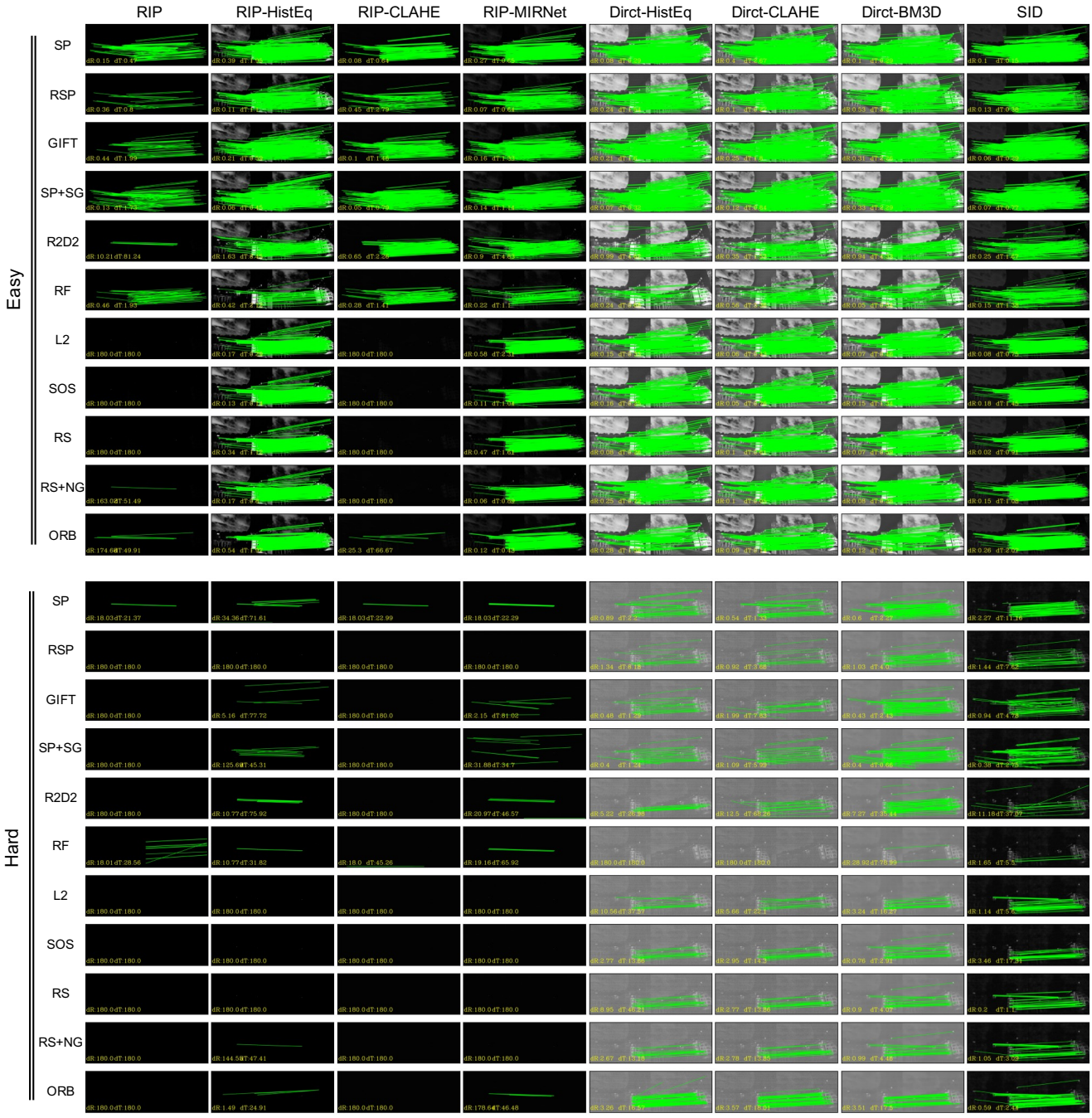

Figure 14. Visualization of the matching results for one of the 54 outdoor scenes. Point correspondences judged as inliers are shown in green lines. The combination of eleven matching methods and the eight image enhancing methods are applied to two image pairs with different levels of exposure (i.e., 'Easy' and 'Hard').