Supplementary Material

In this supplementary document we provide more experimental results and discussion on both benchmark and the proposed ICD datasets.

1. Discussion on benchmark datasets

In this section we discuss more about benchmark datasets correspond to Section 4.2 and Table 2 in the paper.

1.1. Evaluation of separate performance metrics

We investigate the reliability of ILISD on "Cubicle" sequence from CDnet dataset. We choose "Cubicle" as a sample because size of the object is changing through the sequence and at the same time, illumination and shadows are also changing. In this sequence the size of the object is decreased as the frame number increases. Fig. 1 shows *precision*, *recall*, *false-positive*, and *F-measure* for 50 consecutive frames of the sequence "Cubicle" which contain a moving object (i.e. person). Fig. 1(a) shows the precision of the proposed method against all other methods, and as shown ILISD has highest precision and the difference with other methods is more significant when the size of the object becomes smaller because the effect of illumination changes and shadow increases. Fig. 1(b) illustrates that the recall of SSGoDec, MBRMF, DECOLOR, and LSD are higher than our method when the size of the object is large. The reason is that DECOLOR, MBRMF, and LSD use MRF and structural information. This helps the large object to be connected component and therefore recall increases. Although these methods provide higher recall than our method, they also produce lots of false positives in comparison with our method which is shown in Fig. 1(c). We also show the $F - measure$ of the results of methods on all 50 frames, that shows the reliability of the proposed method.

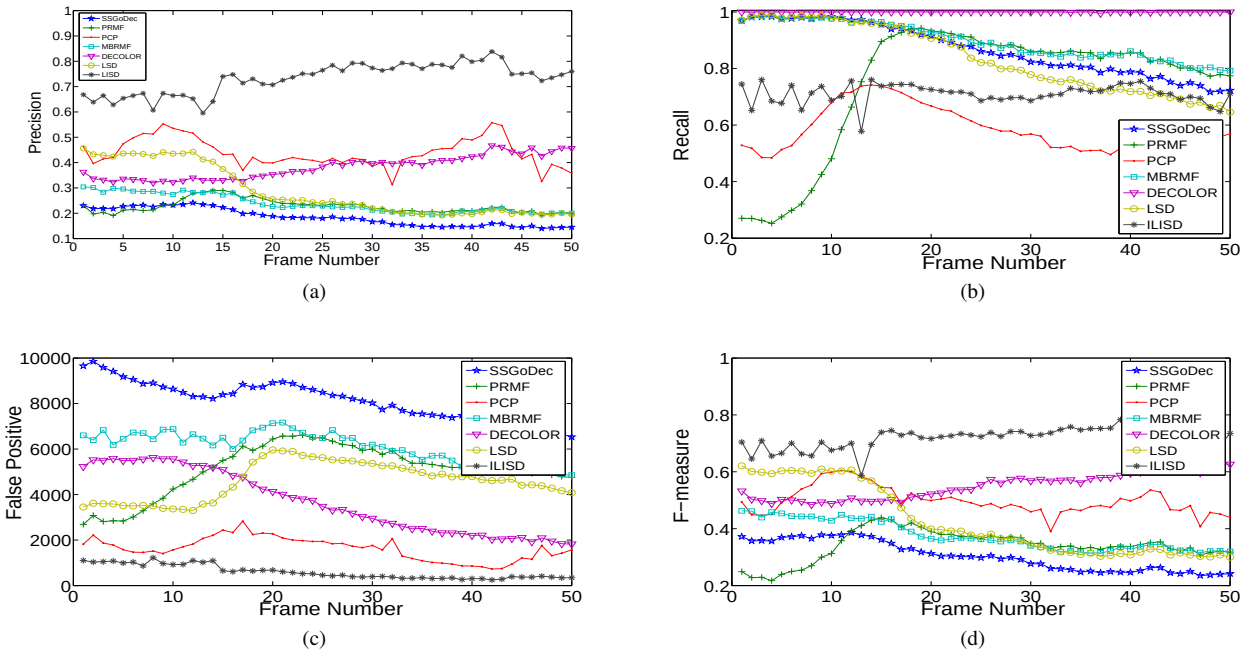


Figure 1: Precision, Recall, False positive and F-measure comparisons of sequence "Cubicle" from CDnet dataset.

1.2. Effect of sudden illumination changes

As we mentioned in the paper, for the "lobby" sequence ground-truth is available for some selected frames, but the ground-truth is not provided while illumination is changing. As illustrated in Table 2 of the paper, the accuracy of our method in this sequence is in the second place after DECOLOR. Here, we consider one more evaluation to show the reliability of our method against DECOLOR and all other competing methods in the "Lobby" sequence as a sample sequence with sudden

illumination change. we select 50 consecutive frames while illumination is changing. These frames have no objects and any foreground pixel is considered as false positive. Fig. 2 compares the false positives of ILISD with all other methods. Fig. 2 clearly shows that ILISD and PCP produce much smaller number of false positives while illumination is changing. Therefore, although DECOLOR shows a little better accuracy on some specific ground-truth frames, unlike our method, it is not reliable while illumination is changing.

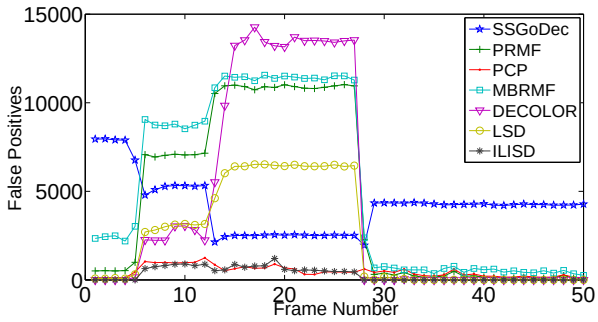
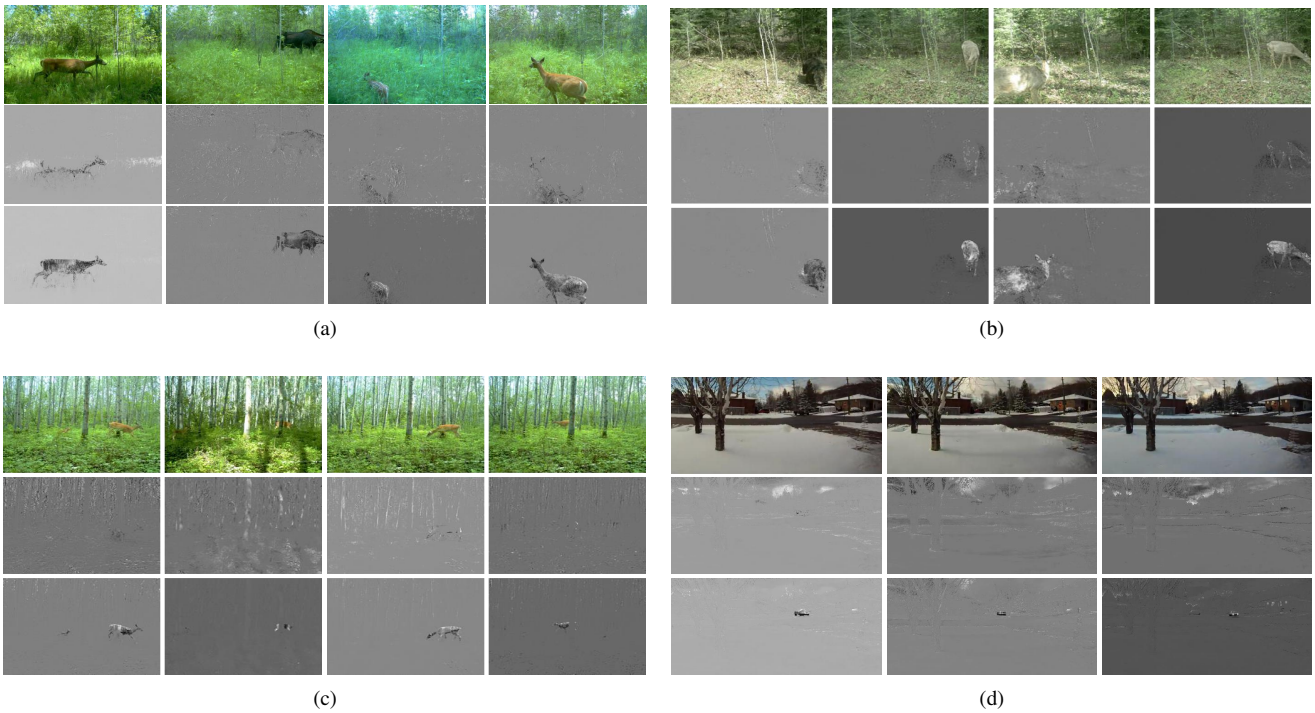


Figure 2: Comparison of number of false positives between the proposed method and the competing methods for sequence "Lobby" from I2R dataset.

2. Discussion on illumination change dataset (ICD)

In this section, we show extensive qualitative results of our proposed method and then compare our method with the competing methods. Fig. 3 shows more results of detecting objects and separating them from illumination changes corresponding to Fig. 4 in the paper. The second and the third rows in each subfigure show the sparse outliers C and S corresponding to illumination change and moving objects, respectively.



(a) (b)

(c) (d)

Figure 3: First row of each sub-figure: sample images with different illumination for sequences a)"Wildlife1", b)"Wildlife2", c)"Wildlife3", and d)"WinterStreet". Second row of each sub-figure: sparse outliers C. Third row of each sub-figure: detected objects S.

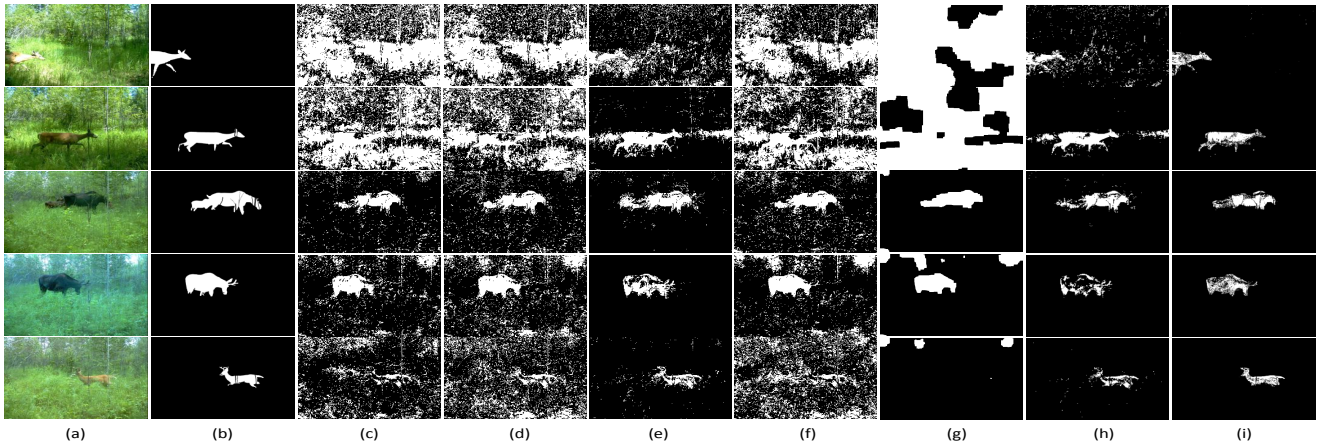


Figure 4: Comparison of qualitative results between our method and six competing methods. a) input, b) ground truth, c) SSGoDec, d) PRMF, e) PCP, f) MBRMF, g) DECOLOR, h) LSD, i) ILISD

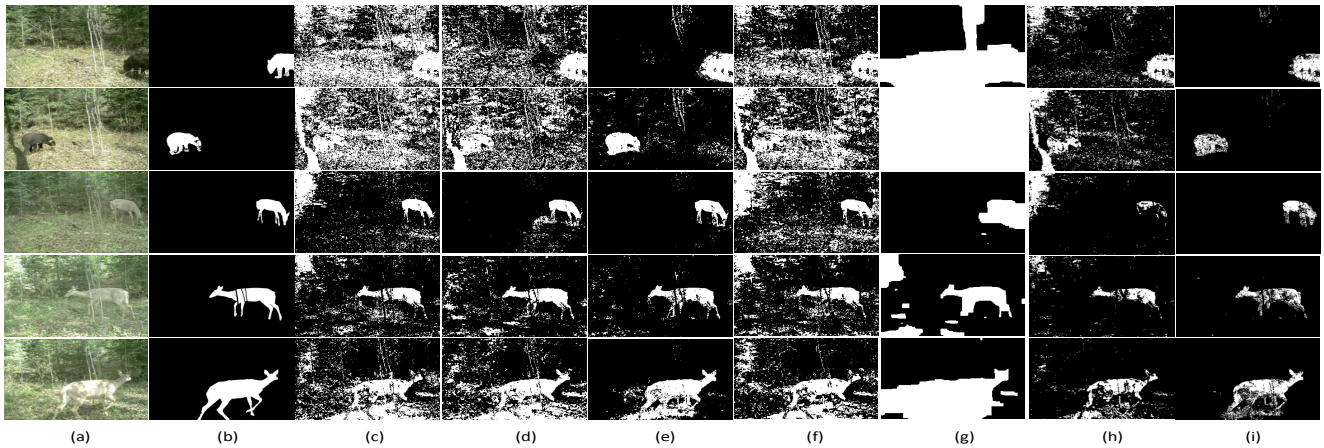


Figure 5: Comparison of qualitative results between our method and six competing methods. a) input, b) ground truth, c) SSGoDec, d) PRMF, e) PCP, f) MBRMF, g) DECOLOR, h) LSD, i) ILISD

For the sake of brevity, in the paper we only showed the qualitative results of "Wildlife3" from wildlife sequences; however, we have shown the numerical results of our method for all sequences and compared them with other competing methods in Table 3 of the paper. Now here, Figs. 4 and 5 illustrate the comparison of qualitative results on selected images of sequences "Wildlife1" and "Wildlife2", respectively.

Fig. 4 shows the comparison of qualitative results on five selected images of sequence "Wildlife1". Since the illumination variations in the time-lapse image sequences are significant, we show five images of the sequence to provide a better comparison. The first two rows depict heavy illumination changes, and only the results of LSD and PCP are comparable with ILISD; however, those methods still have many false positives. The third row shows the results of all methods when the illumination is relatively unchanged and the results of all methods are comparable with ILISD. In the last two rows of Fig. 4 although LSD has not many false positive, its recall is too low, and only PCP is comparable with our method. Fig. 5 demonstrates that in sequence "Wildlife2" only the results of PCP is comparable with our method.