The 36th British Machine Vision Conference (BMVC) 2025

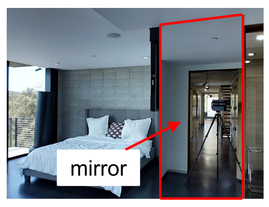
Depth Inconsistency-based Spatial-channel Attention Gate for Mirror Segmentation

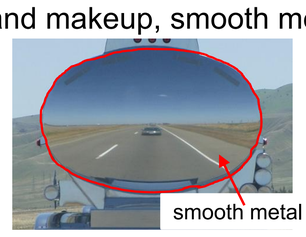
○ Ritsuki Kurohiji Hirotaka Hachiya
Wakayama University, Japan

1. Mirror segmentation problem

Specular objects are present everywhere in daily life.

- Mirrors used for shaving and makeup, smooth metal, and reflective glass, etc.





Specular reflections blur the boundary between real and reflected scenes.

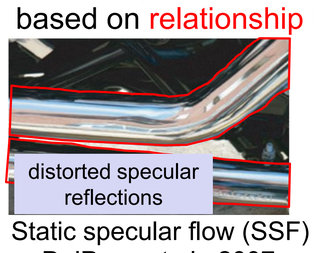
➡ They can severely degrade the performance of SOTA deep learning–based vision systems, e.g., object detection and segmentation.

Accurate mirror segmentation is essential as a preprocessing step.

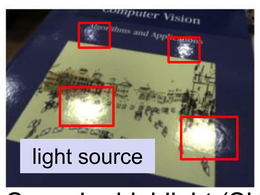
2. Related work

RGB based mirror segmentation

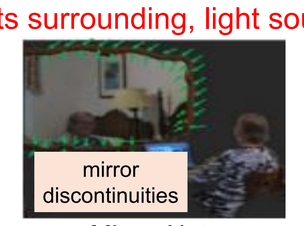
Various specular reflection specific feature extraction methods have been proposed based on relationship between mirror and its surrounding, light sources, and distortion.



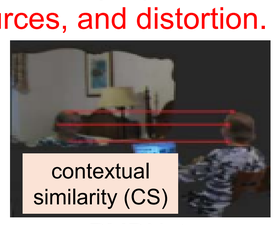
Static specular flow (SSF)
DelPozo et al., 2007



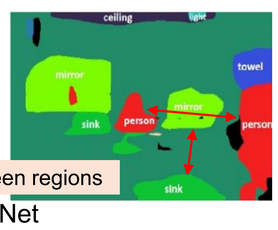
Specular highlight (SH)
Morgand et al.,2019



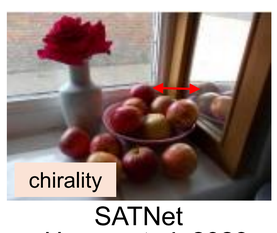
MirrrorNet
Yang et al.,2019



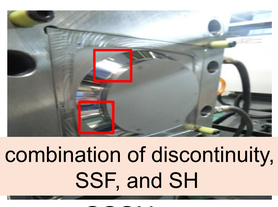
PMDNet
Lin et al.,2020



SANet
Guan et al.,2022



SATNet
Huang et al.,2023



SSSNet
Hachiya et al.,2023

Intrinsically difficult: single RGB image contains limited geometric information

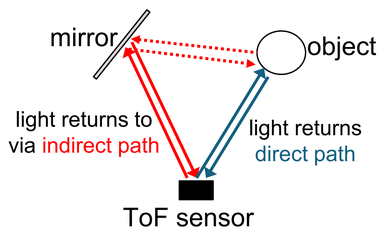
RGB-D based mirror segmentation

Depth measured by ToF sensor is used as geometric information.

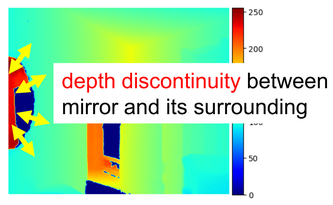
ToF sensor measures depth based on round-trip travel time of reflected infrared light.

Specular objects: the light returns to the sensor via indirect path.

➡ This creates depth discontinuities.



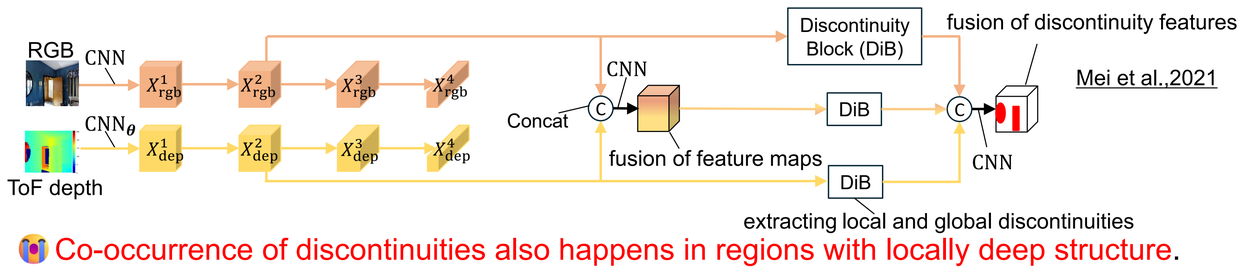
RGB



ToF depth

😄 Depth discontinuities serve as a strong cue for identifying specular reflection regions.

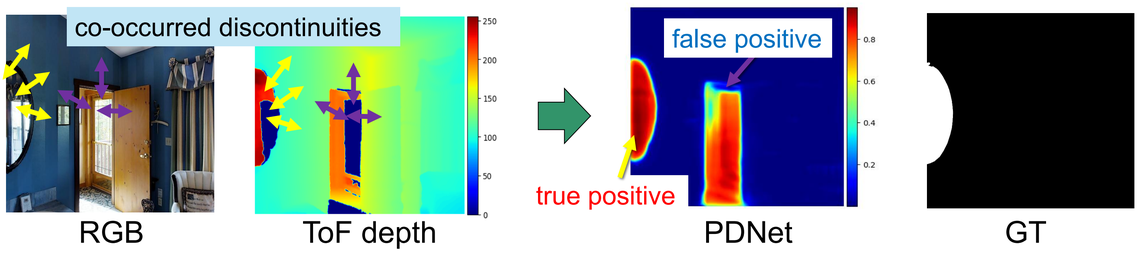
➡ PDNet utilizes co-occurrence of discontinuities in RGB and ToF depth.



Mei et al.,2021

Co-occurrence of discontinuities also happens in regions with locally deep structure.

- For example, opened door and window



➡ PDNet tends to have false positives in such regions.

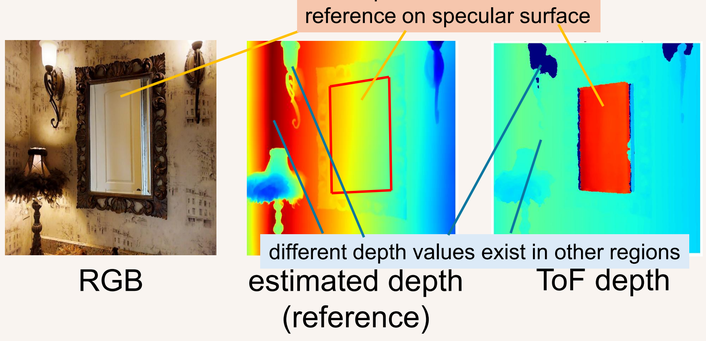
3. Proposed method

ToF depth on specular surface is abnormal

Assumption: regions where ToF depth is abnormally greater than the actual depth are specular surfaces.

Measuring actual depth on specular surface is difficult due to specular reflection.

➡ We use the depth estimated by a deep model, e.g., depth anything, as a reference depth to detect anomaly.
Yang et al.,2024



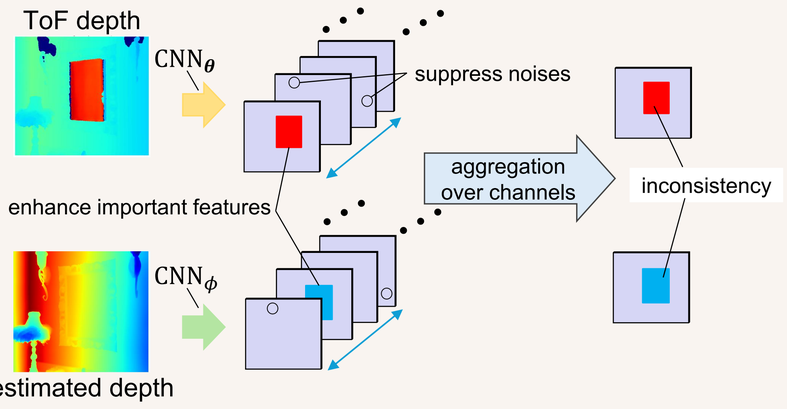
Comparing directly two depth values causes false positives due to intrinsic differences.

- Missing values (out of range and no reflection) in ToF depth
- Different range of values: estimated depth is a relative value

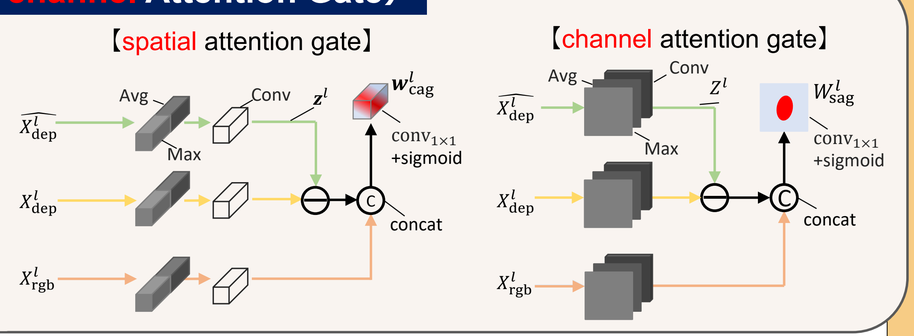
discAG (Depth inconsistency-based spatial and channel Attention Gate)

To find meaningful differences between two depth maps, we propose to:

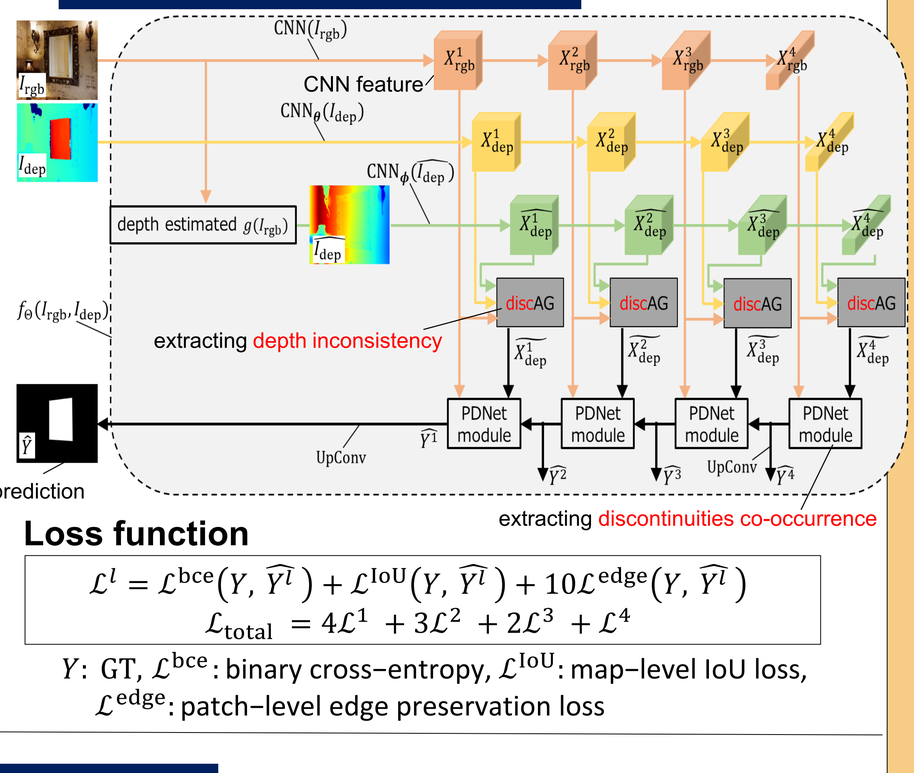
(1) Adaptively suppress/enhance no-related/related features with specular surface using trained CNNs
(2) Align corresponding features between feature maps using statistical and adaptive aggregation along each spatial and channel axis.



😄 Depth inconsistency regarding specular surfaces can be adaptively attended along each spatial and channel axis.



Entire network architecture



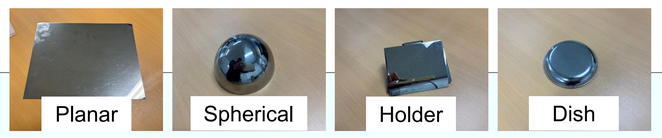
Loss function

$$\mathcal{L}^l = \mathcal{L}^{\mathrm{bce}}\left(Y, \widehat{Y^l}\right) + \mathcal{L}^{\mathrm{IoU}}\left(Y, \widehat{Y^l}\right) + 10\mathcal{L}^{\mathrm{edge}}\left(Y, \widehat{Y^l}\right)$$
$$\mathcal{L}_{\mathrm{total}} = 4\mathcal{L}^1 + 3\mathcal{L}^2 + 2\mathcal{L}^3 + \mathcal{L}^4$$

Y: GT, $\mathcal{L}^{\mathrm{bce}}$: binary cross-entropy, $\mathcal{L}^{\mathrm{IoU}}$: map-level IoU loss, $\mathcal{L}^{\mathrm{edge}}$: patch-level edge preservation loss

4. Experimental evaluation

Dataset



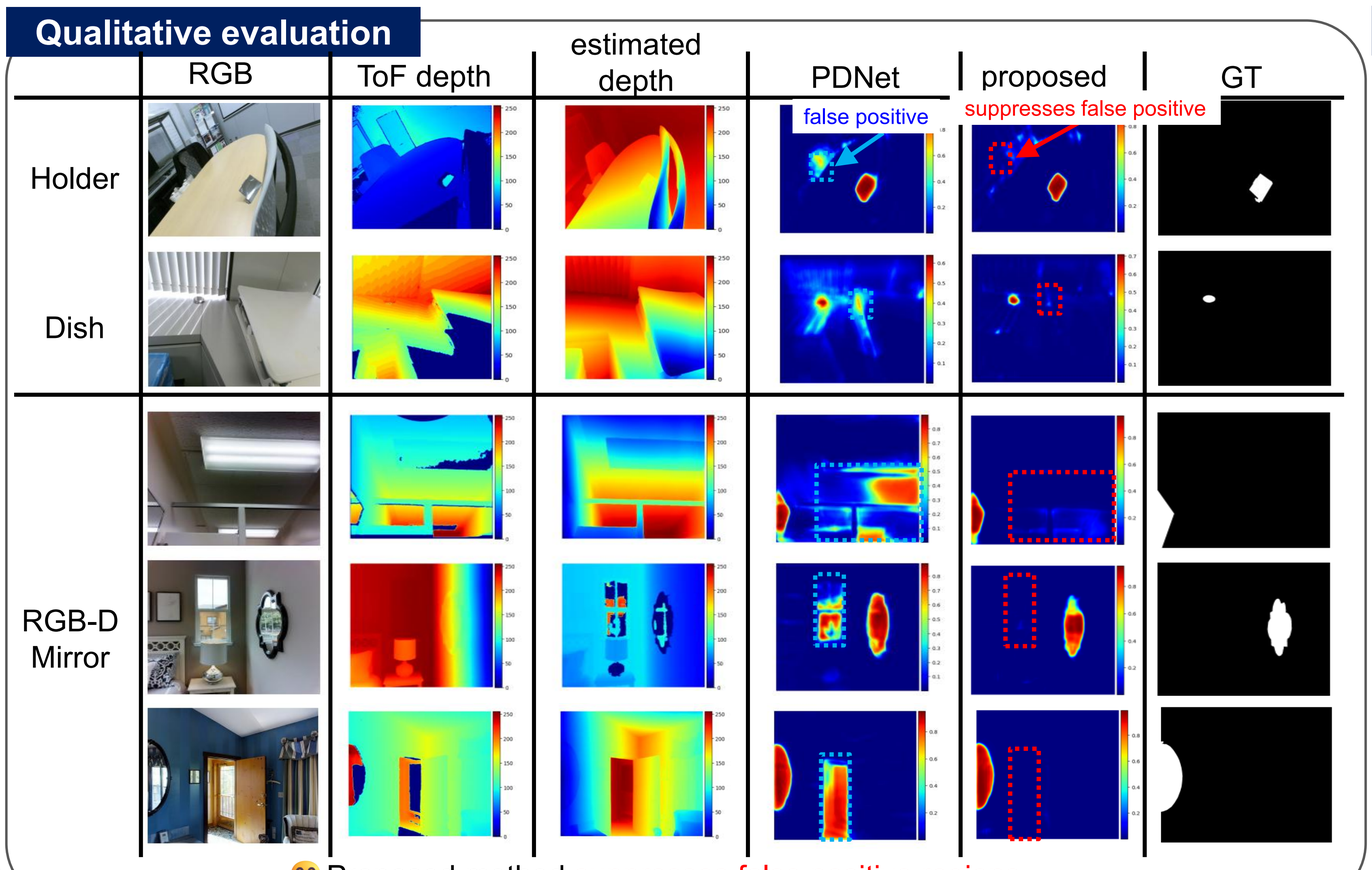
Small specular object dataset (original)

- Originally captured using Kinect V2 in labs and house rooms
- Only single object can appear in a single image

RGB-D mirror segmentation dataset Mei et al.,2021

- Publicly available mirror dataset containing various types of mirrors, e.g., sink, studio, and stand mirrors

Qualitative evaluation

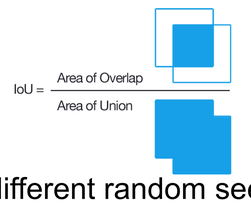


😄 Proposed method suppresses false positive regions.

Quantitative evaluation

Metrics: IoU (Intersection over Union) with threshold for predicted score of 0.5

$$\mathrm{IoU} = \frac{|\mathrm{Predicted} \cap \mathrm{Ground\ Truth}|}{|\mathrm{Predicted} \cup \mathrm{Ground\ Truth}|}$$



Average (standard deviation) of the five trials with different random seeds.

Small specular object

object	PDNet	Proposed
Planar	9.39 (1.37)	12.66 (1.69)
Spherical	56.97 (0.36)	60.66 (2.23)
Holder	44.15 (1.24)	48.40 (0.98)
Dish	57.35 (3.10)	63.07 (2.46)

RGB-D mirror

area rate	PDNet	Proposed
all	68.85 (1.13)	70.94 (0.88)
50%~	65.88 (3.17)	68.01 (3.46)
10%~50%	77.62 (0.79)	78.41 (1.25)
~10%	62.93 (1.39)	66.01 (0.77)

😄 Proposed method achieves higher performance than existing method and the performance improvement is prominent for the small objects

5. Summary and future work

- ToF value is anomaly greater than actual depth for specular surface.
- To find meaningful inconsistency between ToF and estimated depth values, we proposed to suppresses false positive regions.
- Experimental evaluation shows the proposed method is effective.
- We will continue to refine to further reduce false positive.