Supplementary Material: 6DoF Object Pose Estimation via Differentiable Proxy Voting Regularizer

Xin Yu[1]
xin.yu@uts.edu.au
Zheyu Zhuang[2]
zheyu.zhuang@anu.edu.au
Piotr Koniusz[2,3]
piotr.koniusz@data61.csiro.au
Hongdong Li[2]
hongdong.li@anu.edu.au

[1] University of Technology Sydney
[2] Australian National University
[3] Data61, CSIRO, Australia

1 Derivation of Eqn. (2) from the main submission

In Eqn. (2), the first line indicates that we measure the distance between the keypoint $\boldsymbol{k}$ and its corresponding foot of the perpendicular $\boldsymbol{f}_k(\boldsymbol{p})$. Since there is a close form solution for the distance from a point to a line [1], the distance d is calculated by,

$$d = \frac{|Ax^\star + By^\star + C|}{\sqrt{A^2 + B^2}}, \tag{1}$$

where a line is defined as $Ax + By + C = 0$ and a point is expressed as $(x^\star, y^\star)$.

We first obtain the line equation expressed by the direction $\boldsymbol{v}_k(\boldsymbol{p}) = (v_k^x, v_k^y)$ and a point $\boldsymbol{p} = (p^x, p^y)$ as follows,

$$\frac{x - p^x}{y - p^y} = \frac{v_k^x}{v_k^y}.$$

Then, we have:

$$v_k^y x - v_k^x y + v_k^x p^y - v_k^y p^x = 0. \tag{2}$$

Finally, substituting the keypoint coordinates $\boldsymbol{k} = (k^x, k^y)$ and Eqn. (2) into Eqn. (1), we achieve,

$$d = \frac{|v_k^y k^x - v_k^x k^y + v_k^x p^y - v_k^y p^x|}{\sqrt{(v_k^x)^2 + (v_k^y)^2)}}. \tag{3}$$

The equations above lead to Eqn. (2) in the paper.

2 More qualitative results

In this supplementary material, we provide more visual results of our method.

© 2020. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

Figure 1: Visualization of qualitative results on the **LINEMOD** dataset. Green 3D bounding-boxes represent the ground-truth poses and the blue ones indicate our predictions.

References

[1] https://en.wikipedia.org/wiki/Distance_from_a_point_to_a_line

Figure 2: Visualization of qualitative results on the **LINEMOD** dataset. Green 3D bounding-boxes represent the ground-truth poses and the blue ones indicate our predictions.

Figure 3: Visualization of qualitative results on the **LINEMOD** dataset. Green 3D bounding-boxes represent the ground-truth poses and the blue ones indicate our predictions.

Figure 4: Visualization of qualitative results on the **LINEMOD** dataset. Green 3D bounding-boxes represent the ground-truth poses and the blue ones indicate our predictions.

Figure 5: Visualization of qualitative results on the **Occlusion LINEMOD** dataset. Green 3D bounding-boxes represent the ground-truth poses and the blue ones indicate our predictions.

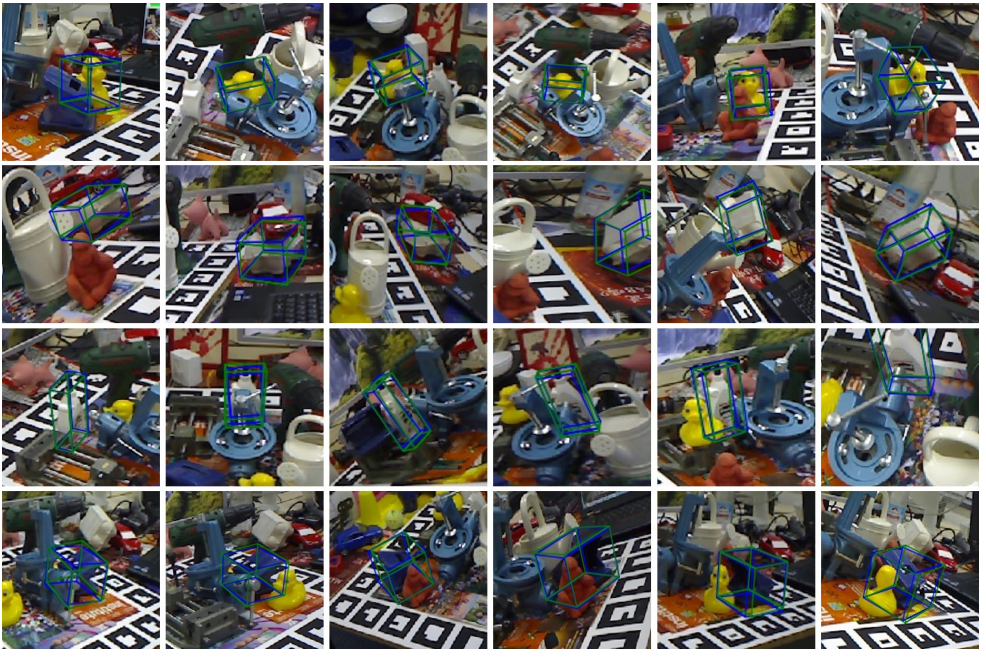

Figure 6: Visualization of qualitative results on the **Occlusion LINEMOD** dataset. Green 3D bounding-boxes represent the ground-truth poses and the blue ones indicate our predictions.