Supplementary Material for the paper
Learning Descriptors for Object Recognition and 3D Pose Estimation

Paul Wohlhart and Vincent Lepetit
Institute for Computer Vision and Graphics, Graz University of Technology, Austria
{wohlhart,lepetit}@icg.tugraz.at

1. Additional Samples

Figures 1, 2 and 3 show additional examples of templates retrieved for random sets of test samples. The first column shows the test sample. To the right, each row shows the first 10 templates, sorted by descriptor distance. Note how most of the closest templates show very similar views of the correct object that all give a good estimate of the object's pose.

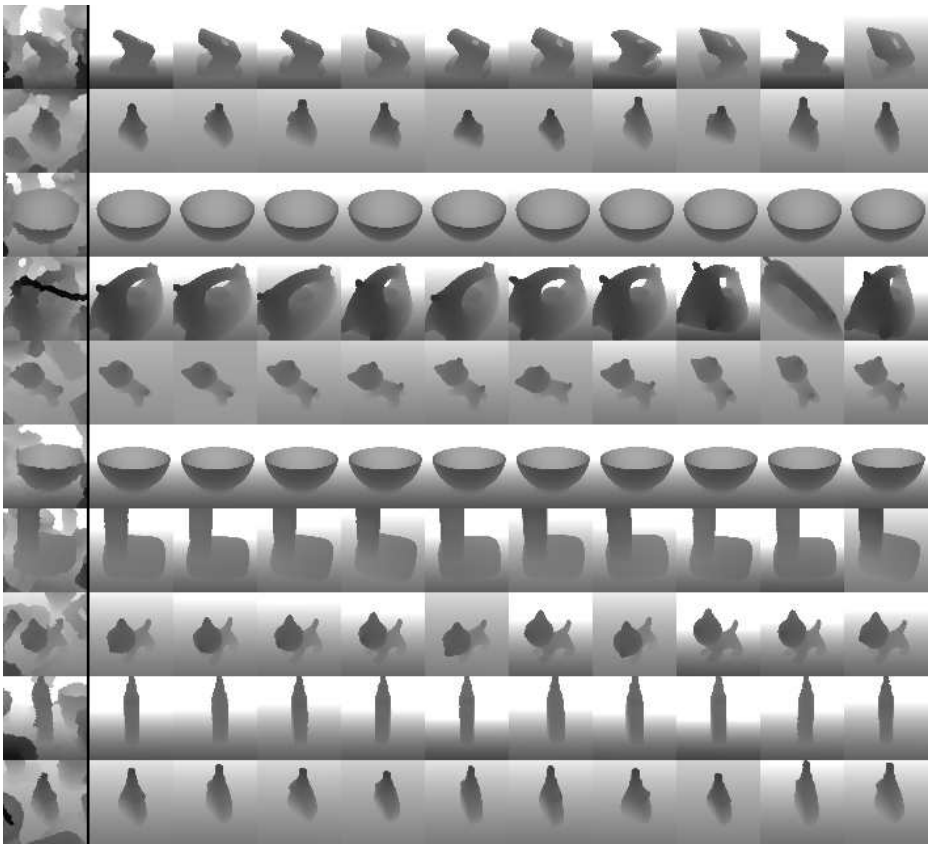


Figure 1. Templates with closest descriptors for samples. Network trained on depth data.

978-1-4673-6964-0/15/$31.00 ©2015 IEEE

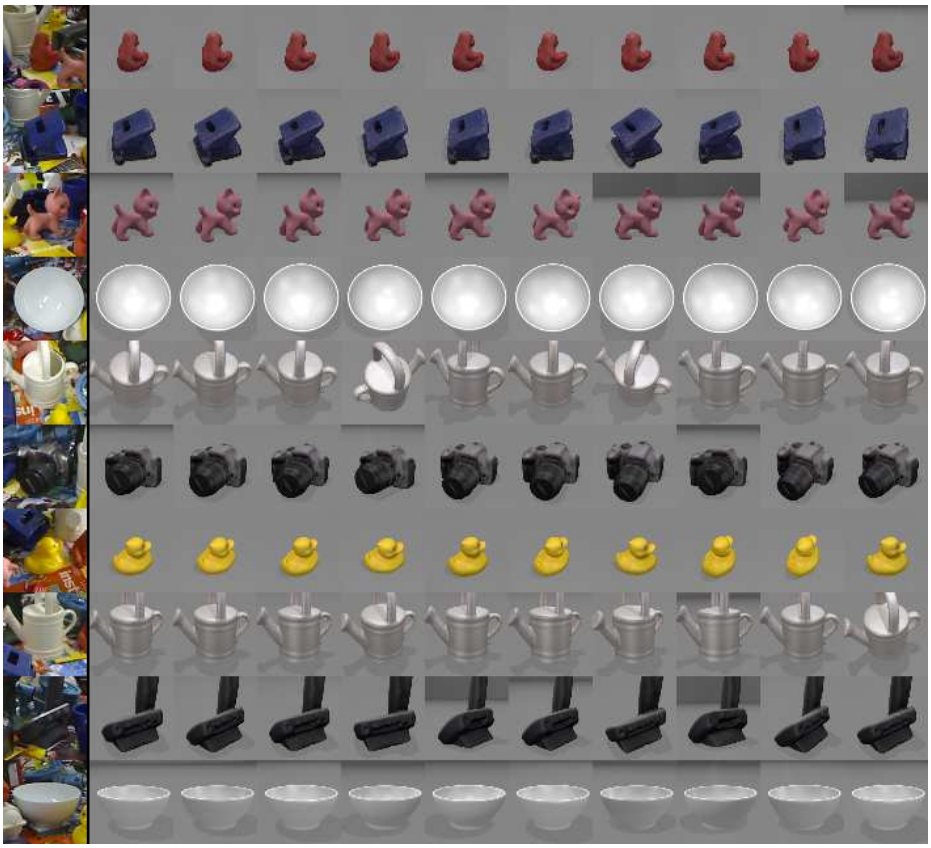


Figure 2. Templates with closest descriptors for samples. Network trained on RGB color data.

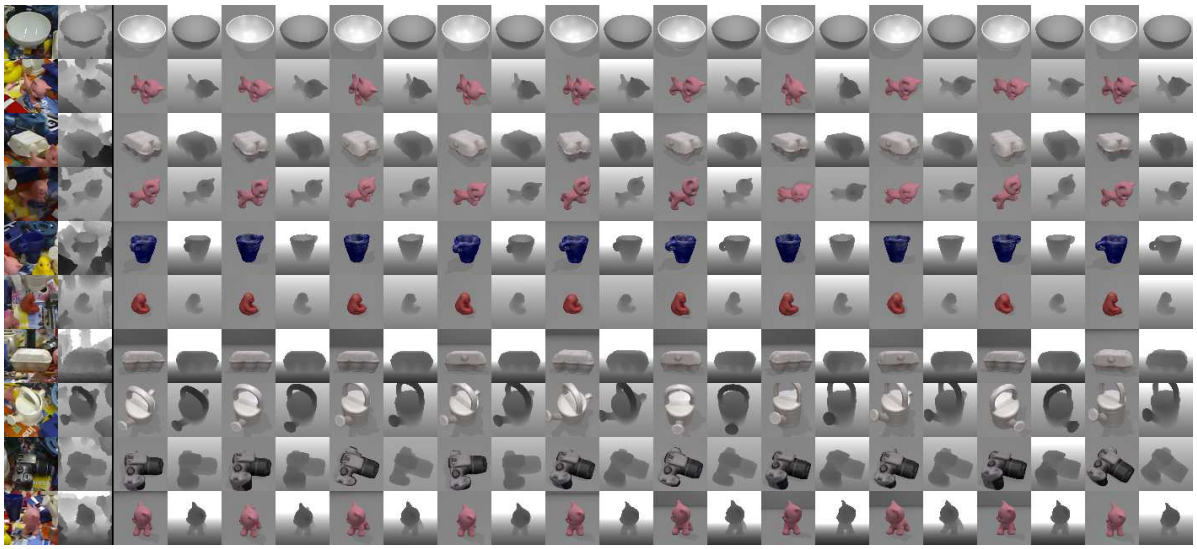
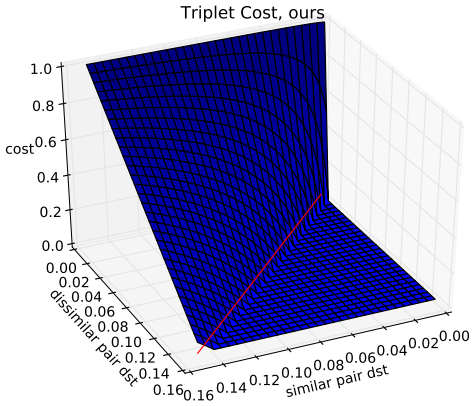


Figure 3. Templates with closest descriptors for samples. Network trained on RGBD data.

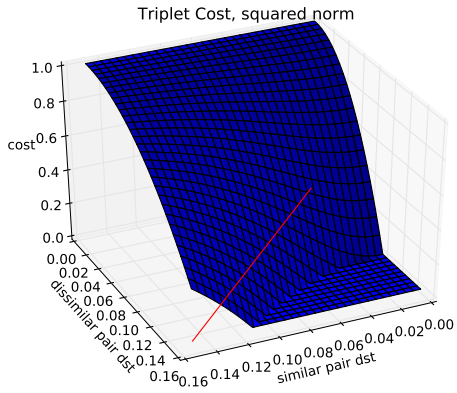
2. Triplet Cost

In Sections 3.2.1 and 3.3 we discuss our definition and implementation of the cost of a triplet in contrast to the definition in related work. Figure 4 shows the value of the cost of one triplet given the distances between similar and dissimilar samples on the x- and y-axis, respectively. On top is our definition, on the lower left our definition, but with the distance squared and on the lower right the definition of Wang *et al*. in CVPR'14. As can be seen, in our definition the value of the cost does not depend on the total scale of the triplet. This allows us to define triplets over arbitrary ranges. A triplet reaching across the whole template dome does not dominate small local triplets and does not contract the similar pair more than it pushes apart the dissimilar one.

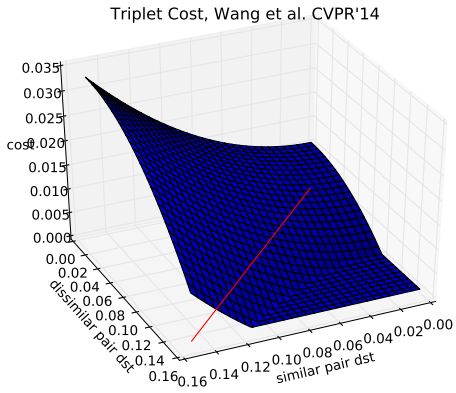
Additionally, like our definition, the other two versions correctly assign high cost to triplets that have a very low distance between the descriptors of the dissimilar samples. However, since the square of the distances is taken, when the distance of the dissimilar pair approaches zero, the derivative w.r.t the distance of the dissimilar pair goes to zero, thus, not pushing apart dissimilar pairs when they are violating the constraints the most.



(a) our definition



(b) ours with squared distances



(c) Wang et al. CVPR'14

Figure 4. Cost of a triplet for different definitions.