



Fast detection of multiple 3-D textureless objects

Author: Pavel Kutáč
Supervisor: Jan Gaura

Technical university of Ostrava

Introduction

Main goal of the solution is to detect and localize multiple 3-D object in the given scene, which are trained from multiple templates. Training objects are illustrated from all sides at different angles.

In many cases, and mostly in the industry, are present objects without texture and also color is not important to accomplish correct detection. In those situations could be better and faster to use algorithms, which detections is based only on the shape of the object.

In each step of sliding window in scene, it would be necessary to compare the cutout with all templates. This would be very inefficient and slow and another intermediate step is added, which contains hash table. In the initial phase are saved all templates into the hash table, which key is 7-dimensional vector. Six values are distances and orientations obtained from the triplet, seventh value is index of triplet.

During matching are possible candidates returned from the hash table based on the key. This leads to massive improvement, as only small subset of all templates is returned as candidates.

How it works

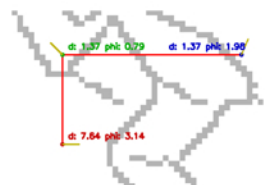
As mentioned, there are no textures present on the objects and color is ignored. In those situations is detection based only on found edges, as shown on the picture below. Comparing template and cutout of sliding window from the scene is not done pixel by pixel, as it would be very slow. Triplets are introduced instead.

Image 1

Source image on the left side and detected edges on the right side



Image 2 - triplet



Triplet over the template with counted distances and orientations of edges in all three points of triplet

Exact amount of triplets are randomly generated on the regular grid over the template. In each point of triplet is counted distance to the closest edge and the orientation of that edge. This is illustrated in the picture above. Comparing in the following steps is done only on those 3 points for each triplet. Also only quantized values of distance and orientation are used to gain some tolerance.

Outputs

Output of the application is list of bounding boxes with information about matched template. Based on the index of matched template, it could be possible to identify exact rotation and the angle. But in present thesis is only marked type of object. In the pictures below are shown 2 scenes with detected objects.

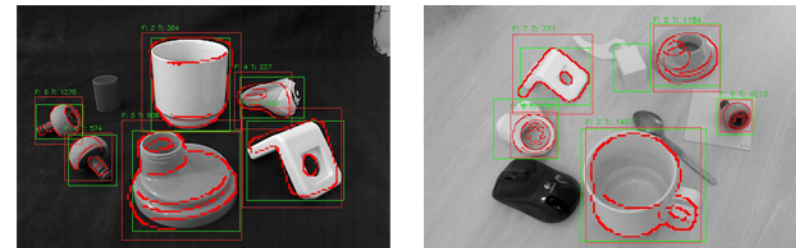


Image 3 - matched objects in scenes

Green rectangles are ground-truth bounding boxes. Red squares are detected locations with illustrated edges of detected template