Model-based (Robustness) Testing of Vision Components of Robotic Systems

Extended Abstract

Wolfgang Herzner, Markus Murschitz, Oliver Zendel

AIT Austrian Institute of Technology GmbH
{wolfgang.herzner, markus.murschitz.fl, oliver.zendel }@ait.ac.at

Abstract. This paper outlines a model-based approach for testing the robustness of computer vision solutions with respect to a given task or application. The outlined approach enables the automatic generation of test data with a measurable coverage of optical situations both typical as well as critical for a given application. In addition correct ground truth data is generated, all with almost no manual effort.[1]

Keywords: computer vision testing, robustness testing, model-based test case generation.

1 Motivation

Public authorities often require a safety certification for commercial application of safety-critical systems. Despite the decades-long use of image processing techniques, still today no general methodology exists for testing their robustness with a measurable coverage of scenes and situational aspects both typical as well as critical for a given application. Therefore, the exploitation of autonomous and robotic systems in human or critical environments that rely on visual perception is hampered by missing means for certifying their visual perception.

Implementation quality assessment, i.e. absence of implementation faults such as access violation or division-by-zero, can be achieved with conventional V&V (validation and verification) methods. One main question remains open: how robust is a solution, i.e. how well does it cope with the huge number of challenges present in the input data, e.g. shadows, reflections, or occlusions? We call these challenges *criticalities*.

For assessing robustness, in most cases recorded test images are used, either from publicly available data sets (e.g. the Middlebury sets[2] of stereo images for disparity

[1] This work is funded by the ARTEMIS-project R3-COP (http://www.r3-cop.eu/) under grant agreement 100233

[2] http://vision.middlebury.edu/stereo/

map testing, or the CAVIAR[3] and "Imageparsing.com" (IP)[4] data for indoor person tracking), or explicitly recorded for a given application. However, both kinds of test data in most cases do not provide an objective measure of covered criticalities.

Furthermore, the expected output (or *ground truth*, GT) needs to be provided for the test data, in order to assess the SUTs response on the test input. This is usually done manually, which is expensive and at least subjective if not error prone.

In order to overcome these deficiencies, the authors developed an approach for generating test data (both stimuli, i.e. images, and expected responses, i.e. GT) from models. This approach is justified by a number of additional observations:

- Computer graphics has reached a maturity status that allows to render realistic images;
- Criticalities can be included explicitly in generated data;
- Scenes can be created which would be too difficult or dangerous to be arranged in reality.

2 The Approach

Following major steps establish model-based test case generation process for computer vision:

1. A *domain analysis* identifies objects that can appear in scenarios of the given application, together with their properties and relationships; all this information is specified in the so-called *domain model.*
2. *Criticalities* that should be included in the test data are either derived from the domain analysis, or selected from a *catalogue of criticalities*. This catalogue is publicly available (2) and open for extension.
3. *3D-scenes* are automatically derived from the domain model such that both the *domain space* and the criticalities are covered. Domain space sampling occurs such that sampling points are distributed with low geometric discrepancy (1).
4. *Images* are rendered from 3D-scenes. Each image is associated with a *trait vector* that characterises the image with respect to the coverage criteria of interest. To minimise redundancy in test data, trait vectors are clustered and cluster representatives selected.
5. Finally, GT is generated for the selected images.

Fig. 1 illustrates this process.

3 Example

For a use case of the R3-COP project, where a robot shall tidy-up a kitchen table, the domain model has been generated. Fig. 2 shows some test image examples, together with their ground truth (centre: depth per pixel, right: segmentation).

[3] http://homepages.inf.ed.ac.uk/rbf/CAVIAR/
[4] http://www.imageparsing.com/frame/MainShow1.html

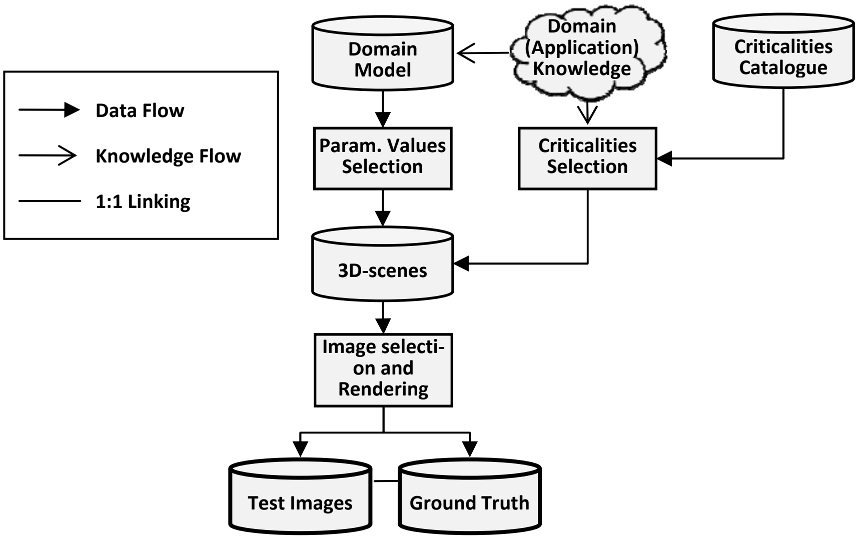

Fig. 1. Test case generation process

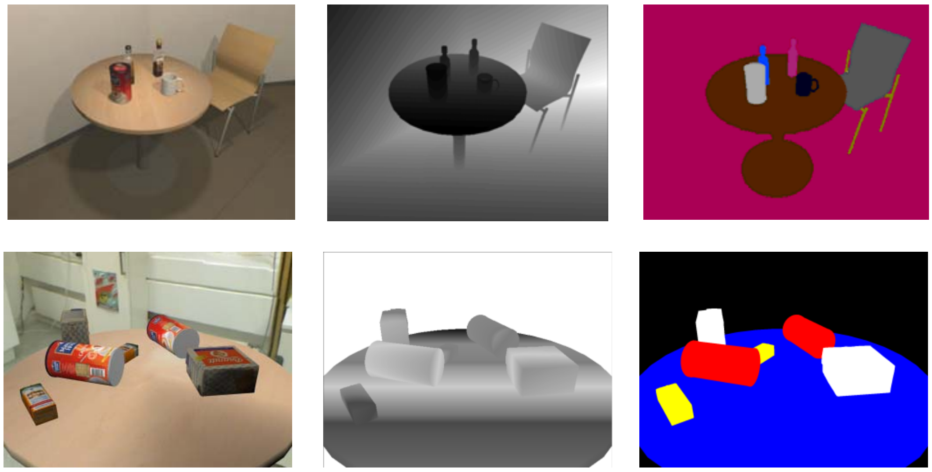

Fig. 2. Test data examples

4 References

1. Jiri Matousek: "Geometric Discrepancy – An Illustrated Guide". Algorithms and Combinatorics Series, Vol.18; Springer, Berlin Heidelberg, 1999. ISBN 3-540-65528-X
2. Oliver Zendel (ed.), Markus Murschitz (ed.): "Criticality Analysis for Computer Vision Algorithms – CV HAZOP". R3-COP deliverable D4.2.3, 2012; available online at http://www.r3-cop.eu/ under "Download".