Detecting Scene Elements Using Maximally Stable Colour Regions

David Obdržálek, Stanislav Basovník, Lukáš Mach, and Andrej Mikulík
Charles University in Prague, Faculty of Mathematics and Physics
Malostranské náměstí 25, 118 00 Praha 1, Czech Republic
david.obdrzalek@mff.cuni.cz, sbasovnik@gmail.com,
lukas.mach@gmail.com, andrej.mikulik@gmail.com

Abstract. Image processing for autonomous robots is nowadays very popular. In our paper, we show a method how to extract information from a camera attached on a robot to acquire locations of targets the robot is looking for. We apply maximally stable colour regions (a method originally used for image matching) to obtain an initial set of candidate regions. This set is then filtered using application specific filters to find only the regions that correspond to scene elements of interest. The presented method has been applied in practice and performs well even under varying illumination conditions since it does not rely heavily on manually specified colour thresholds. Furthermore, no colour calibration is needed.

Keywords: Autonomous robot, Maximally Stable Colour Regions.

1 Introduction

Autonomous robots often use cameras as their primary source of information about their surroundings. In this work, we describe a computer vision system capable of detecting scene elements of interest, which we used for an autonomous robot. The main goals we try to achieve are robustness of the method and computational efficiency.

The core of our vision system are Maximally Stable Extremal Regions, or MSERs, introduced by Matas et al (see [1]) for gray-scale images and later extended to colour as Maximally Stable Colour Regions, or MSCR (see [2]). Details about MSER and MSCR principles are given in Sections 2 and 3 respectively.

The main usage of MSER detection is for wide-baseline image matching mainly because of its affine covariance and high repeatability. To match two images of the same scene (taken from different viewpoints), MSERs are extracted from both images and then appropriately described using (usually affinely invariant) descriptor (see [3,4]). Because MSER extraction is highly repeatable, the majority of the regions should be detected in both images. If the descriptor is truly affinely invariant, identical regions should have the same (or similar) descriptors even though they are seen from different viewpoints (assuming the regions correspond to small planar patches in the scene). Then, the matching can be done using nearest neighbour search of the descriptors.
In our system, MSCRs are not used for matching but for object detection. The system operates in the following steps:

- Detect large number of contrasting regions in the image.
- Classify detected regions and decide which correspond to elements of interest.
- Localize detected elements and pass this information to other components of robot’s software.

MSER and MSCR algorithms often return large number of (possibly overlapping) regions. We therefore introduce our classification algorithm which rejects regions with small probability of corresponding to a scene element of interest. The relative position of the scene element is then determined using standard algorithms from computer vision and projective geometry [5].

Typical input image and output in the form of list of detected objects can be seen in Figure 1.

<table>
<thead>
<tr>
<th>Localization results</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>object 1</td>
<td>[-0.74, 0.72, 0.01]</td>
</tr>
<tr>
<td>object 2</td>
<td>[0.84, 1.97, 0.05]</td>
</tr>
<tr>
<td>object 3</td>
<td>[-0.81, 2.20, -0.01]</td>
</tr>
<tr>
<td>object 4</td>
<td>[0.23, 3.21, 0.04]</td>
</tr>
</tbody>
</table>

Fig. 1. Input image, detected regions, and final output table – triangulated coordinates of detected objects

The following text is structured as follows: We first briefly describe the MSER (Section 2) and MSCR (Section 3) algorithms. In Section 4 we present our filtering system which processes the regions and outputs locations of detected objects. Section 5 discusses the overall efficiency of the proposed algorithm.

2 MSER

In this section, we describe the MSER algorithm as a basis for our region detection.