A novel approach for real-time semantic context labeling

Citation for published version (APA):

Document status and date:
Published: 01/01/2015

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.

Download date: 31. Aug. 2023
A Novel Approach for Real-Time Semantic Context Labeling

Fons van der Sommen, Martin Pieck, Svitlana Zinger, Peter H.N. de With
Eindhoven University of Technology
P.O. Box 513, 5600 MB Eindhoven

Abstract

The use of context information is an important aid for full semantic scene understanding in security and surveillance applications. This paper presents an innovative semantic context-labeling algorithm for three context classes – water, sky and vegetation – while trading-off quality and real-time execution. Our system consists of three consecutive stages: image segmentation, region-based feature extraction and classification. We propose the joint use of color features in HSV space, texture from Gabor filters and spatial context, in combination with the Directional Nearest Neighbor (DNN) method for segmentation. Compared to recent literature, this combination is over 35 times faster and achieves a coverability rate that is 65% higher [6].

1. Introduction

Semantic interpretation of image content is a challenging problem in computer vision. An algorithm for this type of image analysis is of high importance to the security and surveillance industry, because the manual analysis of video feeds is a tedious and error-prone process. Early approaches by Bao et al. [1] and Rankin et al. [7] focus on one context class and are extended too multi-class by Javanbakhti [2]. Based on the framework of Javanbakhti, we propose a near real-time context detection algorithm with an improved detection quality. The contributions of this paper are three-fold. First, innovative methods are proposed for segmentation, texture features and classification, to optimize for both quality and execution time. Second, an improved quality performance evaluation is proposed and third, an efficient C++ implementation using multi-core functionality is presented that achieves near real-time execution.

2. Proposed method

The proposed framework is inspired by a recent region-based approach [2] and consists of three major stages: (1) segmentation of the image into regions, (2) the computation of region features and (3) the classification of each region.

2.1. Image segmentation

For segmentation of the image, variations of the graph-based segmentation algorithm by Felzenswalb et al. [4] are applied of which later a specific version will be chosen to embed in the framework. For construction of the graph we explore three options: Grid-based (8N), Neighbor-based (24N), as employed by Javanbakhti [2] and Directional Nearest Neighbor (DNN), as proposed by Liu et al. [5].

2.2. Feature extraction

In the feature extraction stage, the HSV color intensity, Gabor-based texture information and Spatial Context (SC) are extracted from the regions resulting from the segmentation. Compared to recent literature, this combination is over 35 times faster and achieves a coverability rate that is 65% higher [6].

2.3. Region Classification

In the last stage, every region is classified and labeled by a Support Vector Machine (SVM). The SVM is trained offline using a radial basis function kernel with 10-fold cross-validation combined with a grid search for optimal parameters C and γ. This search also incorporates the SVM model size, as fewer support vectors result in faster classification.

3. Experimental results

We evaluate the proposed semantic context-labeling algorithm on images of natural scenes, selected from the dataset used in [2], so that all images contain the semantic classes water, sky and vegetation. This set contains 54 images of which 49 have a 481 × 321 pixels resolution and 5 have a 320 × 240 resolution. The results regarding execution time are computed using only the 481 × 321 images to obtain comparable uniform results. For each image, the manually
Table 1. Comparison of quality performance for water detection.

<table>
<thead>
<tr>
<th>Features</th>
<th>Segm.</th>
<th>CR (%)</th>
<th>ER (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I HSV</td>
<td>DNN</td>
<td>59.6</td>
<td>36.3</td>
</tr>
<tr>
<td>II HSV+Gabor</td>
<td>DNN</td>
<td>79.3</td>
<td>19.1</td>
</tr>
<tr>
<td>III HSV+Gabor+SC</td>
<td>DNN</td>
<td>92.2</td>
<td>12.9</td>
</tr>
<tr>
<td>IV HSV+Gabor+SC</td>
<td>8N</td>
<td>84.0</td>
<td>14.0</td>
</tr>
</tbody>
</table>

Table 2. Time performance comparison for water detection in milliseconds for each algorithm stage: Pre-Processing (PP), Segmentation (SEG), Gabor Filtering (GF) and Classification (CL).

<table>
<thead>
<tr>
<th>PP SEG GF CL Total time</th>
</tr>
</thead>
<tbody>
<tr>
<td>I 1 142 113 15 272</td>
</tr>
<tr>
<td>II 2 143 (parallel) 15 161</td>
</tr>
<tr>
<td>II 3 109 (parallel) 15 127</td>
</tr>
<tr>
<td>IV 1 106 (parallel) 18 125</td>
</tr>
<tr>
<td>V 1 587 58 3,754 4,400</td>
</tr>
</tbody>
</table>

annotated ground truth is available. Figure 1 shows an example image, including the corresponding ground truth and algorithm output. The SVM is trained on 24 images and the algorithm is tested using the remaining 30 images, where we measure the quality by the performance of Coverability Rate (CR) and Error Rate (ER) as proposed by Schmitt et al. [8]. The algorithm is implemented in C++, using OpenCV for image processing functions and the OpenMP API [3] for parallel functionality. The application is executed on an i7 quadcore @ 2.4-3.4 GHz processor with 8 GB DDR@ 1600 MHz RAM on 64-bit Windows 8.1 Pro.

Table 1 compares the quality performance for different segmentation and feature configurations. From the table, it can be observed that adding extra features for the same segmentation method (DNN) leads to higher performance. Namely, the CR increases while the ER decreases. The implementation with the highest performance, HSV+GABOR+SC in combination with DNN, outperforms the method of Javanbakhti by more than 65%.

Table 2 shows the results of timing analysis for each step in the algorithm. From the table it can be observed that by exploiting the parallelism, the total computation time can be reduced from 272 ms to 161 ms. Applying extra parallel functionalities, make our non-parallel approach more than 2.2 times faster, leading to overall 125-ms execution time.

4. Conclusion

In this paper, we propose a novel semantic context labeling algorithm that exploits the joint use of the HSV color features, Gabor filters and spatial context in combination with a DNN segmentation method. Experiments show that it is over 35 times faster and achieves a CR that is 65% higher, compared to state-of-the-art alternatives.

References