Mapping and Assessing the Traffic Infrastructure

Croatian Science Foundation, University of Zagreb development fund
in partnership with Inst. of Traffic and Communications and Graz University of Technology

Siniša Šegvić[1], Zoran Kalafatić[1], Karla Brkić[1], Axel Pinz[2]

[1] University of Zagreb Faculty of Electrical Engineering and Computing
[2] TU Graz, Institute of Electrical Measurement and Measurement Signal Processing

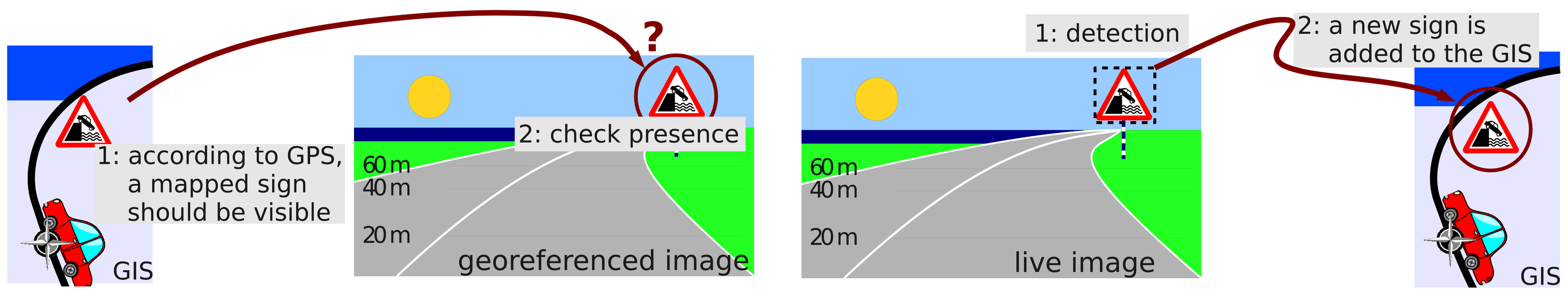
1. Introduction

This project addreses inovations in the field of exploitation and maintenance of geoinformation inventories for storing elements of traffic infrastructure. The main motivation for developing such systems is to provide a comprehensive insight into the prescribed state of the road. This capability is subsequently applied for streamlining various safety inspections of a public road network in operation.

The most important objective of road safety inspection concerns the compliance of traffic control infrastructure. The inspections are designed to detect broken, covered, worn-out or stolen traffic signs, and erased or incorrectly painted road surface markings. In current practice, this task is carried out by tedious, subjective and error-prone manual comparisons of the recorded video against the reference state.

Problem description

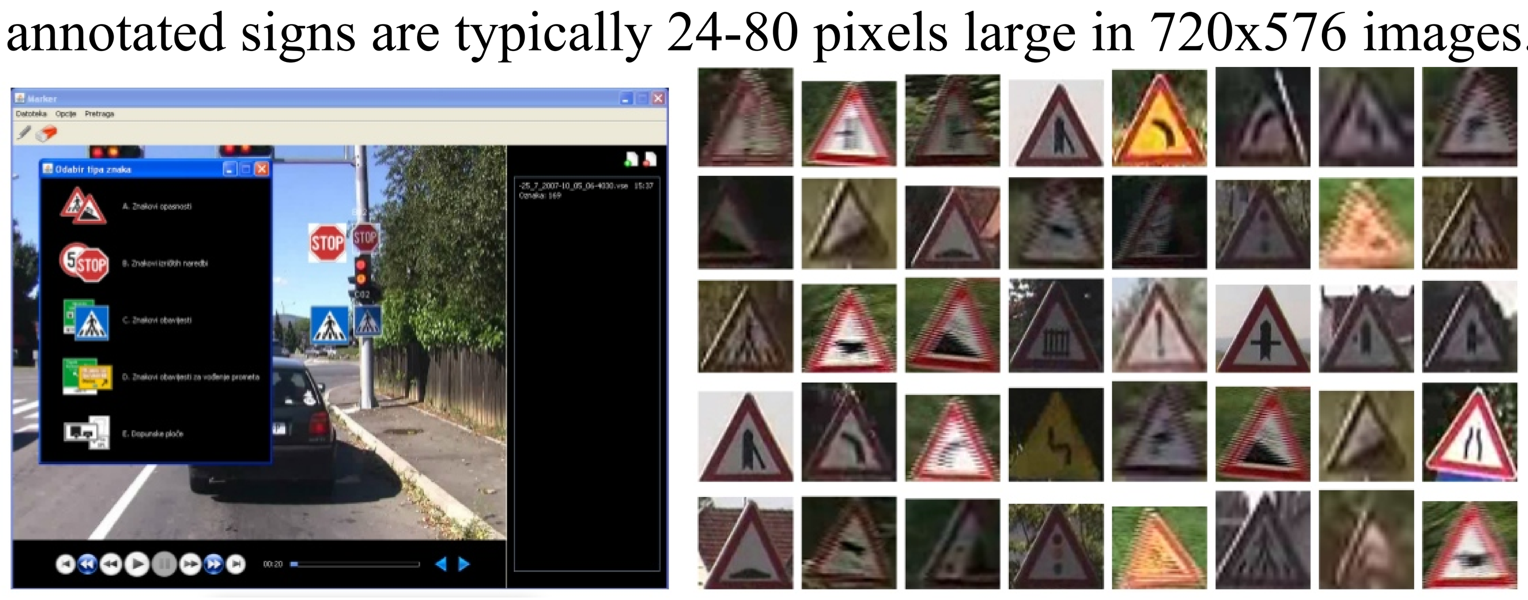
We aim at improving the quality of road safety inspection by relaxing the requirements for expensive human experts. We therefore research opportunities for automating two important operations on geo-referenced video from the driver's perspective.



The first key-operation corresponds to the verification of traffic control devices by automatic comparison of a recent georeferenced video with previously constructed geoinformation inventory (cf. fig. above-left). The second key-operation is the construction of a geoinformation inventory for previously uncharted road network, again starting from georeferenced video (cf. fig. above-right). Both cases require advanced computer vision techniques for detecting and recognizing traffic signs and road surface markings.

3. Traffic sign datasets

In order to be able to train and evaluate detection and recognition algorithms, we have manually annotated about 7500 traffic signs. The annotated signs are typically 24-80 pixels large in 720x576 images.



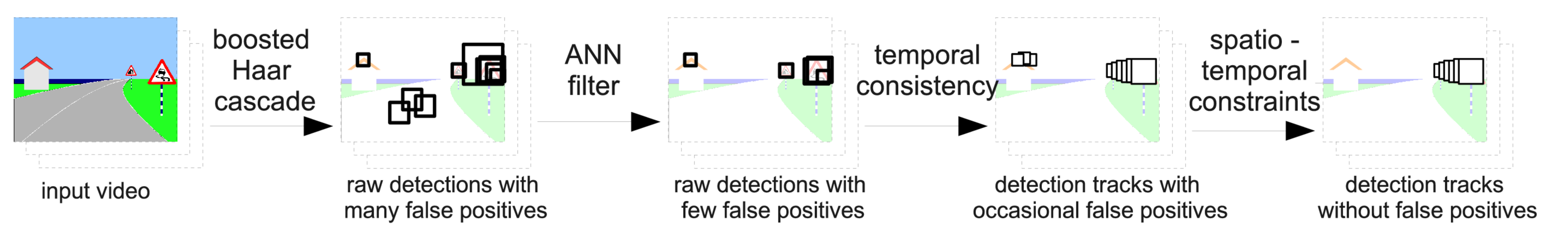
In experiments we focus on triangular warning signs, since 50% of signs on Croatian local roads belong to that class. We organize our corpus of about 3000 triangular signs into two datasets:

T2009: ~2000 signs from images acquired with an interlaced camera.
T2010: ~1000 signs from images acquired with a progressive camera
We use T2009 (left) for training and T2010 (right) for testing.

[1] Siniša Šegvić, Karla Brkić, Zoran Kalafatić, Axel Pinz. Exploiting temporal and spatial constraints in traffic sign detection from a moving vehicle. Machine Vision and Applications, accepted for publication.
[2] Karla Brkić, Axel Pinz, Siniša Šegvić, Zoran Kalafatić. Histogram-Based Description of Local State-Time Appearance", SCIA, Ystad Saltsjöbad, Sweden May 2011.
[3] Siniša Šegvić, Karla Brkić, Zoran Kalafatić, Vladimir Stanisavljević, Marko Ševrović, Damir Budimir and Ivan Dadić. A computer vision assisted geoinformation inventory for traffic infrastructure. ITSC, Madeira, Portugal, September 2010.

4. Methodology

Best detection has been achieved by combining boosted Haar cascades (great sensitivity), bootstrap trained conventional classifiers such as HOG+ANN or HOG+SVM (better specificity), differential tracking (temporal consistency), and spatio-temporal features (contextual constraints). These algorithms are organized in a heterogeneous attentional cascade, as illustrated in the figure below.

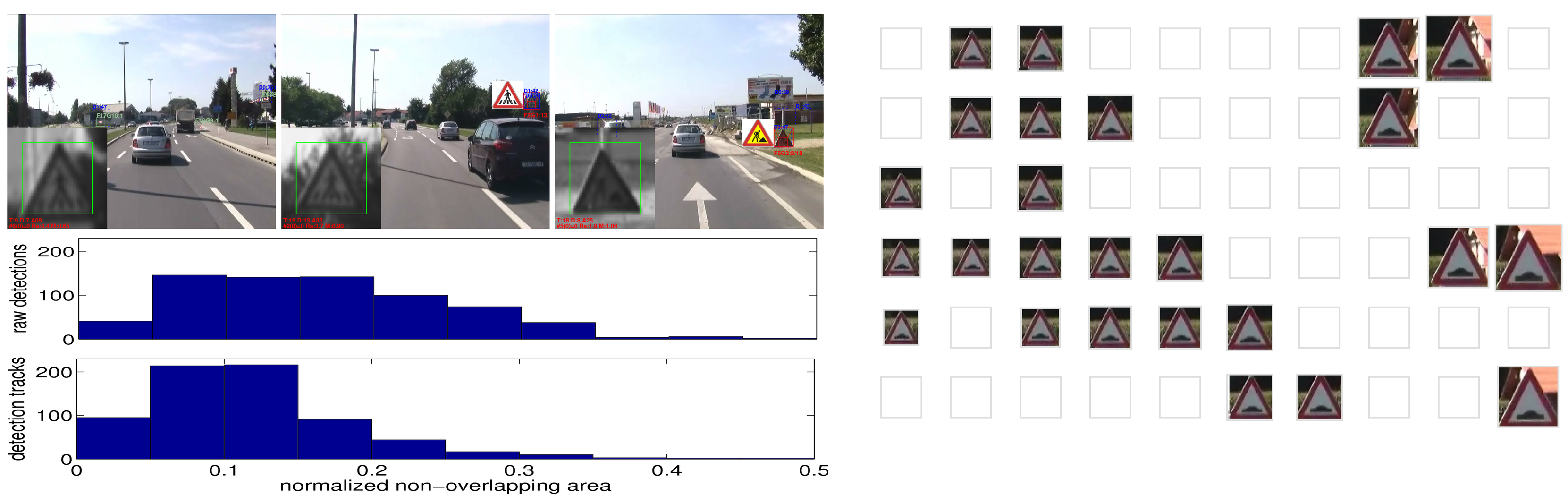


We have achieved best recognition results by machine learned decision trees of binary SVM classifiers operating on HOG features. We obtained best performance after conditioning the training set according to the empirically determined localization inaccuracy of the detection responses.

5. Detection experiments

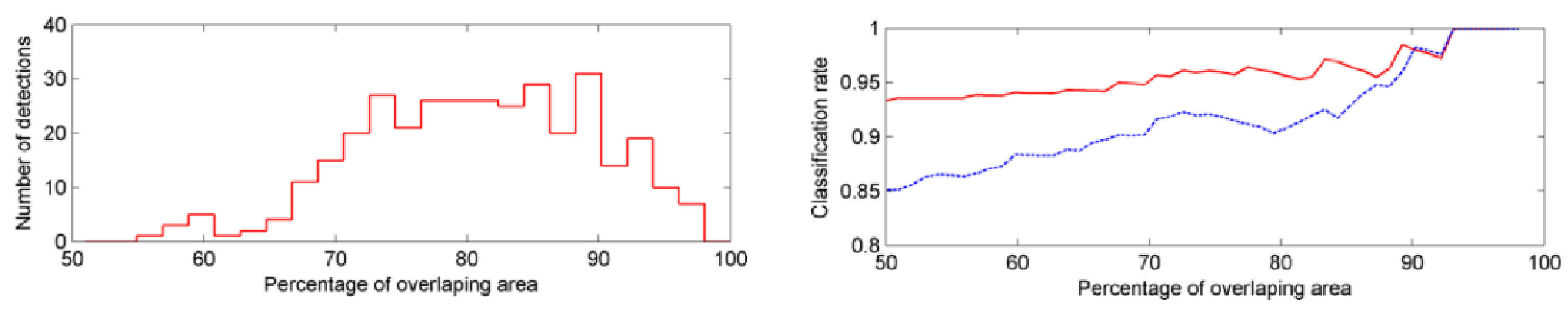
Our basic detection algorithm applies a boosted Haar classifier in a detection window. This typically achieves 95% recall and about 50% precision. After additional filtering with HOG+ANN (or HOG+SVM) the precision grows to about 90%, at the price of a lower recall for smaller traffic signs (total filtered recall is also about 90%).

At the system level, the recall grows to near 100% (fig. below-left) since all signs get detected when they come close to the camera. However, some false positives remain. When a widely used detection chaining criterion is used, the false positive incidence is about 2 per minute. By requiring temporal consistency of the detection appearance (fig. below-right), this improves to about 0.3 false positives per minute while the location accuracy increases (fig. bottom-left).



6. Recognition experiments

The achieved accuracy in the recognition experiments depends on traffic sign size, localization inaccuracy of the detection window and the number of the training samples. The achieved results range between 94% and 98% in individual images, and about 99% in video. Effects of the training set conditioning are shown in the figure below.



7. Conclusions

We have developed vision techniques with an exceptional potential in numerous traffic applications. Even though the presented results exclusively deal with triangular warning signs, similar results have been obtained on other classes of ideogram-based signs. Due to spatial limitations, we could not present research on detection and recognition of road surface markings where we also achieved encouraging results.

First Croatian Computer Vision Workshop, Zagreb 20-21 September 2012