Automatic bus line number localization and recognition on mobile phones: A computer vision aid for the visually impaired

Claudio Guida
Dario Comanducci
Carlo Colombo

Dipartimento di Sistemi ed Informatica
University of Florence, Italy

Computational Vision Group
http://cvg.dsi.unifi.it

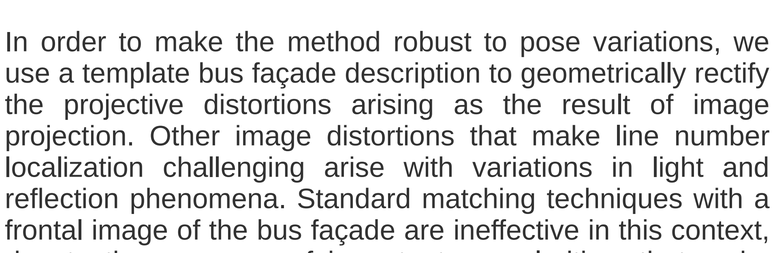
Summary

Automatically reading bus line numbers with a smart phone can greatly improve the autonomy of visually impaired people in urban scenarios. The problem is a challenging one, since standard methods fail due to the abundance of distractors, occlusions, illumination changes, highlights and specularities, shadows, and perspective distortions. The proposed solution exploits a careful mix of model based machine learning and geometric computer vision.

The method

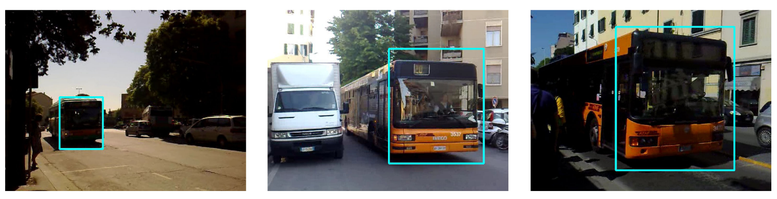
Our method is divided into three main steps:

- Bus detection
- Localization of the bus number region
- Bus number recognition



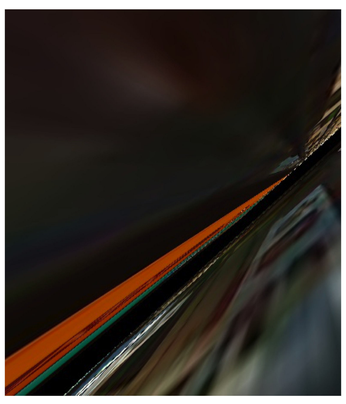
Bus detection

In this is preliminary phase live images are analyzed to detect the bus presence. For this purpose, the Viola and Jones cascade classification algorithm [1] is used.

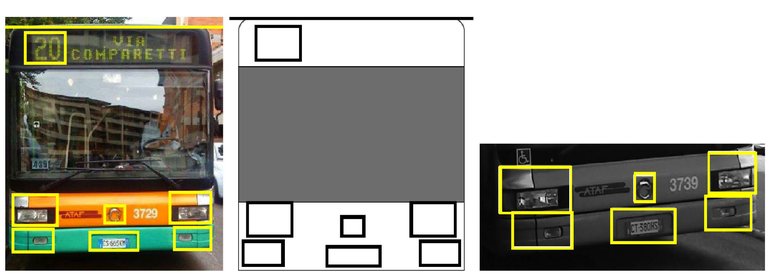


Localization of the bus number region

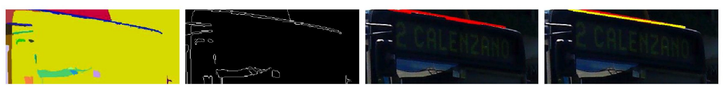
In order to make the method robust to pose variations, we use a template bus façade description to geometrically rectify the projective distortions arising as the result of image projection. Other image distortions that make line number localization challenging arise with variations in light and reflection phenomena. Standard matching techniques with a frontal image of the bus façade are ineffective in this context, due to the presence of important specularities, that make feature extraction and image-to-template alignment prohibitive tasks.



A different route is taken here. Specific classifiers [1] are trained so as to locate and delimit with boxes the most salient parts of the bus façade in the live image. Each box center **x** and its template counterpart **x'** are put in correspondence.



The top border line of the bus façade is also used as an element of the template. This line is located in a geometric way using Canny edge detection and the Hough transform. To cope with the high image noise, a graph-based region segmentation algorithm [2] is used.



The image-to-template façade homography is then estimated in a robust way via least squares and LMEDS from the box center (x,x') and top line (**l**,**l'**) correspondences:

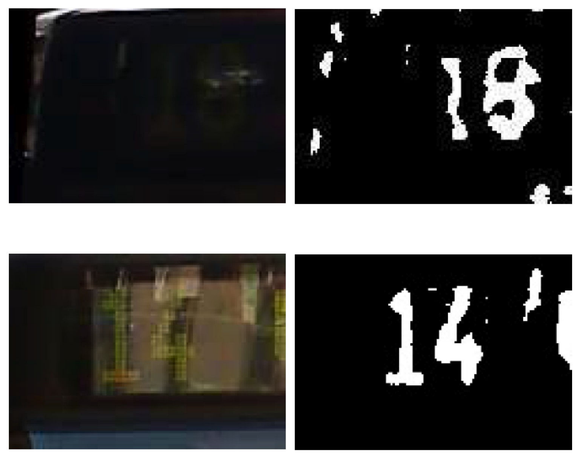
$$\boldsymbol{l}' = H^{-T}\boldsymbol{l}$$
$$\boldsymbol{x}_i' = H\,\boldsymbol{x}_i \quad i=1,\ldots,6$$

We can now safely locate the bus number ROI by simply selecting the corresponding area in the template, and mapping it onto the live image using the façade homography.

Bus number recognition

Due to severe shadowing and mirroring phenomena, classic OCR binarization methods [3] are also ineffective.

We propose here an alternative binarization method based on complementary colors, that exploits the knowledge of the bus number color. Working in the HSV color space, we find the complementary of the bus number hue, and use it as a threshold value for binarization.



To make the process invariant to light variations, we actually use an adaptive thresholding method, by which the optimal threshold value is selected by iteratively binarizing the image and counting the number of points that are above threshold.

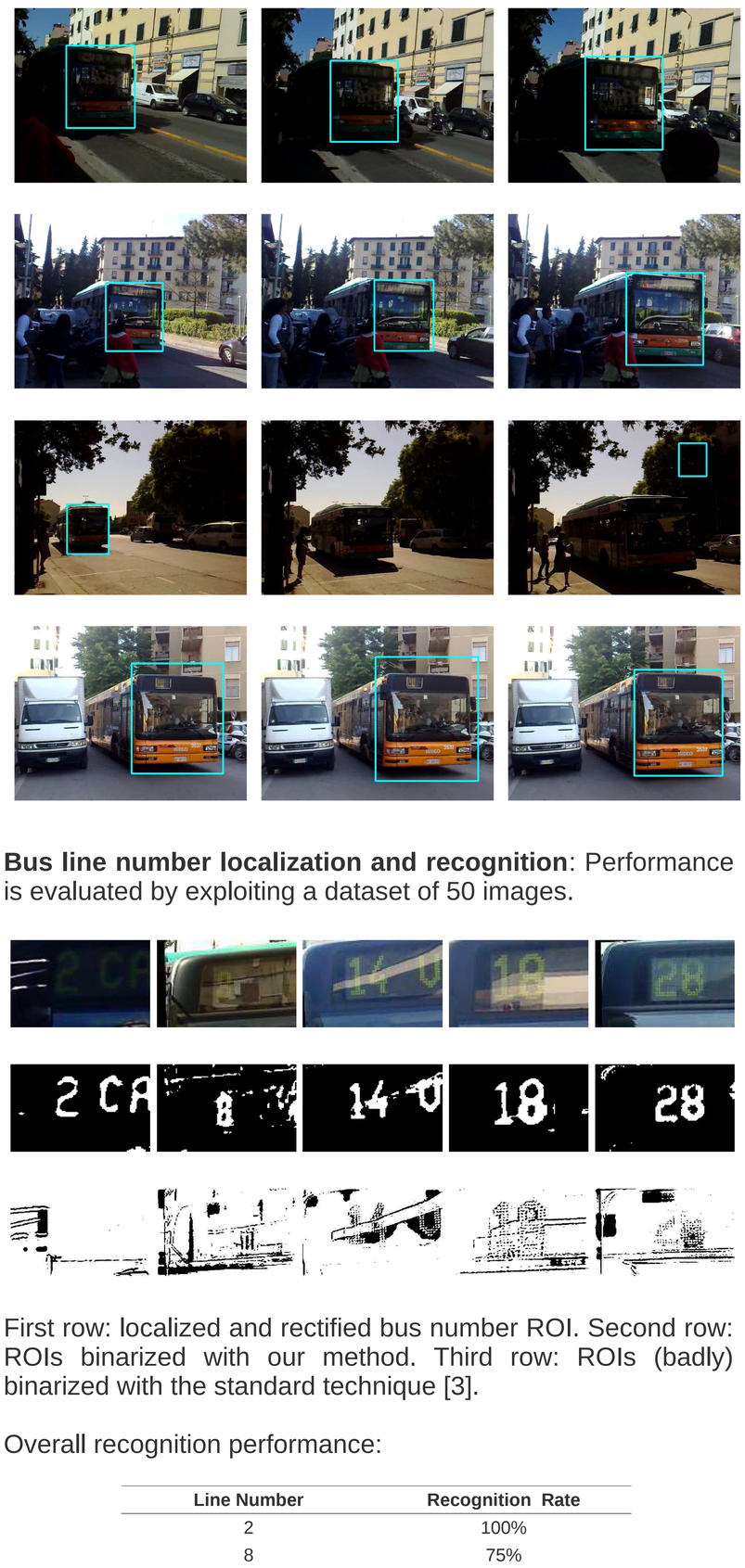
To recognize the bus number we trained a classifier for each possible candidate. Classifiers are put in a chain of ascending precision of classification, and a first-win-takes-all schema is employed.

Experimental results

Bus detection: Some video sequences were taken at different bus stops, in order to test the bus detection accuracy.

Sequence	Frame	Detection	False Positive	Rate
No. 1	17	11	1	64.7%
No. 2	40	40	N/A	100%
No. 3	32	6	2	18%
No. 4	24	23	N/A	95.8%
No. 5	46	33	3	72%
No. 6	48	27	1	56.2%

Bus detection rate is quite high in all sequences save No. 3, where errors occur due to direct sunlight overexposure.



Bus line number localization and recognition: Performance is evaluated by exploiting a dataset of 50 images.

First row: localized and rectified bus number ROI. Second row: ROIs binarized with our method. Third row: ROIs (badly) binarized with the standard technique [3].

Overall recognition performance:

Line Number	Recognition Rate
2	100%
8	75%
14	100%
18	100%
28	100%

Conclusions and future work

Our method combines geometric computer vision with machine learning so as to achieve robustness with respect to highlights, specularities, shadows, occlusions, and so on. Future work will address using multiple bus templates, and dealing with several oncoming buses simultaneously. Bus localization will be performed within a pyramidal framework, and faster alternatives to the Viola and Jones approach will be investigated for the detection of template elements and the localization of the number region.

References

[1] Viola, P., Jones, M.: Robust real-time face detection. International Journal of Computer Vision 57(2), 137–154 (2004)

[2] Felzenszwalb, P., Huttenlocher, D.: Efficient graph-based image segmentation. International Journal of Computer Vision 59, 167–181 (2004)

[3] Seeger, M., Dance, C.: Binarizing camera images for OCR. Proc.6th International Conference on Document Analysis and Recognition. pp. 54–58 (2001)