HAL
open science

Recognition of Supplementary Signs for Correct Interpretation of Traffic Signs

Anne-Sophie Puthon, Fabien Moutarde, Fawzi Nashashibi

To cite this version:

Anne-Sophie Puthon, Fabien Moutarde, Fawzi Nashashibi. Recognition of Supplementary Signs for Correct Interpretation of Traffic Signs. IEEE Symposium on Intelligent Vehicles (IV'2013), Jun 2013, Gold Coast, Australia. hal-00875706

HAL Id: hal-00875706

https://minesparis-psl.hal.science/hal-00875706v1

Submitted on 22 Oct 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Recognition of supplementary signs for correct interpretation of traffic signs

Anne-Sophie Puthon, Fabien Moutarde and Fawzi Nashashibi

***Abstract*— Traffic Sign Recognition (TSR) is now relatively well-handled by several approaches. However, traffic signs are often completed by one (or several) supplementary sign(s) placed below. They are essential for correct interpretation of main sign, as they specify its applicability scope. The main difficulty of supplementary sub-sign recognition is the potentially infinite number of classes, as nearly any information can be written on them. In this paper, we propose and evaluate a hierarchical approach for recognition of supplementary signs, in which the "meta-class" of the sub-sign (Arrow, Pictogram, Text or Mixed) is first determined. The classification is based on the pyramid-HOG feature, completed by dark area proportion measured on the same pyramid. Evaluation on a large database of images with and without supplementary signs shows that the classification accuracy of our approach reaches 95% precision and recall. When used on output of our sub-sign specific detection algorithm, the global correct detection and recognition rate is 91%.**

I. Introduction

Cars are now often equipped with Advanced Driving Assistance Systems (ADAS) for helping the driver and avoiding accidents. Many of those ADAS can benefit from, or even require, vision-based Traffic Sign detection and Recognition (TSR), and a correct interpretation of these signs. For instance, in order to develop a smart Adaptive Cruise Control (ACC) that would automatically tune target cruising speed depending on the current speed limit, one needs reliable detection and recognition of speed-limit signs. This is a mandatory complement to speed-limit information extracted from GPS cartographic data, which is neither always complete nor systematically up-to-date, and furthermore cannot handle temporary speed limits for road works, nor variable speed limits, due to the static nature of pre-defined digital cartographic data.

Until now, most TSR published research and existing systems rely on the same process. At first, a detection stage aims at extracting from the image all the possible locations of speed signs. Two different types of information can be used. Colour information (see e.g. [1], [2], and [3]) makes the detection easier but less robust, and shape-based detection working on greyscale images (e.g. [4], [5], [6], and [7]) increases robustness to variation of illumination conditions and to fading away of colours. Each candidate is then tested through a classification step in order to eliminate false alarms and recognize the sun-sign type. Finally a tracking helps in validating the results from frame to frame.

Figure 1. Examples of traffic signs that cannot be interpreted correctly without taking into account the supplementary sign located below.

However, the exact meaning of traffic signs, in particular speed-limit signs, is often altered by one (or several) supplementary sign(s) placed below (see Figure 1). These supplementary signs are essential for correct interpretation of main signs, as they specify their applicability scope, such

Anne-Sophie Puthon and Fabien Moutarde are with the Robotics Laboratory (CAOR), Mines ParisTech, 60 Bd Saint-Michel, F-75272 Paris cedex 06, FRANCE, Phone: (+33) 1.40.51.92.92, Fax: (+33) 1.43.26.10.51 (e-mail: Anne-Sophie.Puthon@ensmp.fr, Fabien.Moutarde@ensmp.fr)

Fawzi Nashashibi is with IMARA project at INRIA, Domaine de Voluceau – Rocquencourt, B.P. 105 – F-78153 Le Chesnay, FRANCE, Phone: (+33) 1.39.63.52.56 (e-mail: Fawzi.Nashashibi@inria.fr).

as type of vehicle (car, truck, etc.), distance extension, specific lane, particular dates or timeslots, or weather condition (rain, snow, etc.).

II. Related and previous work

The detection and recognition of supplementary signs is a task very similar to usual Traffic Sign detection and Recognition. It might even, wrongly, be thought to be simpler due to the fact that detection can be guided by position and size of the main sign it is related to.

However there are several difficulties specific to supplementary signs:

1/ the aspect ratio (height/width) of the rectangle wildly varies between sub-sign types (see Figure 2);

2/ supplementary signs are much less normalized than main signs, and significantly vary from country to country;

3/ the contrast between sub-signs and their surroundings is often low. In greyscale images, the background colour of supplementary signs can easily be confused to sky, and contrary to most main signs, there is no coloured border around them;

4/ supplementary sign recognition is a classification task with a potentially infinite number of classes, as many sub-signs contain rather "free" text.

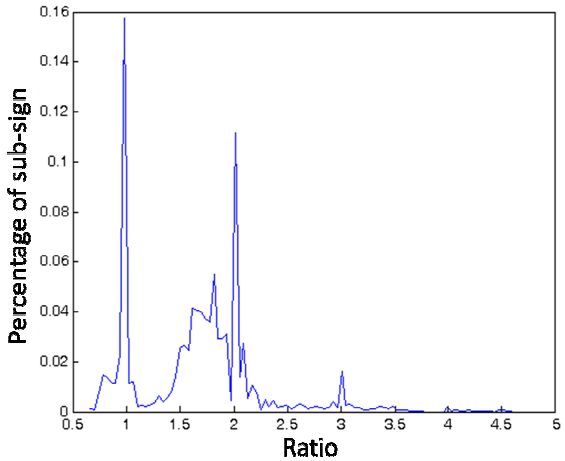

Figure 2. Sub-sign width/height ratio distribution for our database (see Table 1).

Some or all of the above difficulties could explain why, despite its practical importance, supplementary sign detection is not a trendy field and only very few works were published. To our best knowledge, the first paper including some work on supplementary sub-sign detection and recognition is that of Hamdoun *et al.* [8] in 2008, who proposed to use a rectangle edge detection for localizing potential sub-signs in a rather large region below the main signs. However their recognition, using a feed forward neural network applied to rescaled images, was developed and tested only for "Exit Arrow" supplementary sign type (one of the most frequent in France). Nienhüser *et al.* [9] have presented in 2010 a work handling more types of sub-signs: "Exit Arrow", "Truck", "Bei Nässe" ("when raining") and "7.5t". They however had to train a one-against-all Support Vector Machine (SVM) classifier *for each type*, which makes difficult to extend their approach for handling all types of supplementary signs. Liu *et al.* [10] also include supplementary signs detection (based on gradients and Hough voting) and recognition (using a tree of binary SVMs applied to Fourier and wavelet descriptors) in their system presented in 2011. Finally, Puthon *et al.* [11] in 2012 have proposed and tested a new original and performant technique for the detection step, based on region-growing from contrasted seeds.

In the present paper, we propose and evaluate a general framework for *recognition* of supplementary signs.

III. Recognition of traffic supplementary signs

Recognition of supplementary signs is not very different from TSR in general, except for the existence of a potentially very large number of classes, as several sub-sign types consist of "free" text. The undefined and large number of classes pleads for the use of a hierarchical classification approach as adopted by Liu *et al.* in [10]. However, in their approach the addition of a new type of sub-sign would require retraining their whole cascade of classifiers. For this reason, and also for taking into account the specificity of "text" supplementary signs, we have decided to implement a *hand-crafted* hierarchical recognition in which we first determine the "meta-class" of the sub-sign, between "Arrow", "Pictogram", "Text", and "Mixed" (see Figure 3). The idea is that for "text" meta-class, interpretation could be done in a further stage by applying character recognition.

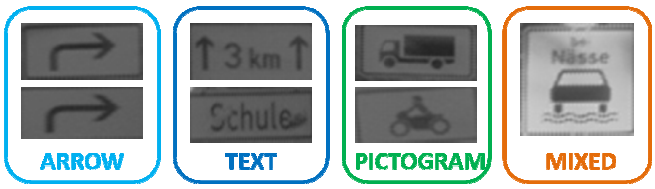

Figure 3. Illustration of the 4 "meta-classes" used for supplementary signs.

The first step is the choice of the features (or "descriptors") that should be computed on the candidate region and fed into classifiers. We focus on global descriptors which describe the entire image contrary to local ones, such as interest points, localizing information around discriminative regions. A solution is to use the pixel values, provided that rescaling to a fixed size is first applied, as was done in [8] and [9]. However the drawback is that any slight offset or "bounding" error in the detection stage can translate into a total change of the resulting descriptor.

For this "offset robustness" issue, it seems better to use descriptors that integrate information on large enough

chunks of the sub-sign. Resizing to very few pixels (e.g. 8x8) might sound tempting, but it would not retain enough information to discriminate categories. We therefore turned to the "pyramid-HoG" (pHoG) descriptor, proposed by Bosch *et al.* [12], which has the advantage of being less sensitive to image offset, while retaining pertinent and discriminating information (see Figure 4).

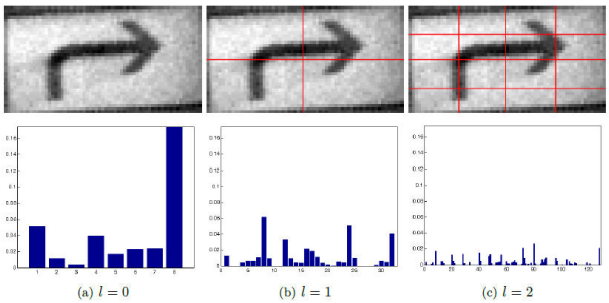

Figure 4. Illustration of pyramid-HoG (pHoG) descriptor used. On the left, with no subdivision, in middle with L=1 level of subdivision, on the right with L=2 levels.

The main parameter for pHoG is the number of levels L of the pyramid. For determining its optimal value, we compared performances for L=0, 1, 2 and 3. As can be seen on Figure 5, increasing L from 0 to 2 provides significant improvement to classification accuracy attained when training Gaussian kernel SVMs but going above L=2 does not bring more recognition performance. **L=2** thus provides the best performance and a relatively small number of values. Another critical parameter for pHoG is the number K of bins for computing the histogram of gradient orientations; similar analysis of the impact of the number K of bins for computing the histogram has led us to fix it to the smallest value **K=8**.

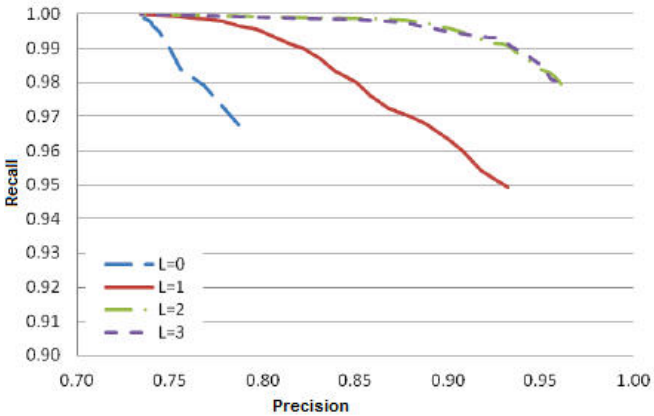

Figure 5 Influence of the number L of levels of the pyramid: L=2 shows much better recall and precision than L=0 and 1, and L=3 brings no further improvement. Evaluation is done on Gaussian kernel SVMs trained on variants of the feature.

By looking closer at frequent confusions that were produced when training SVM classifier on only pHoG descriptor, we realized that most of them seemed due to the fact that pHoG contains only gradient information, while correct discrimination between our meta-classes would clearly be made easier if using also some colorimetric information. We have therefore enhanced pHoG by adding, for each sub-region of the pyramid, the proportion of "dark" pixels (as illustrated on Figure 6). This "dark proportion" is computed using the same binarization technique as that used by us in [11] for extracting contrasted seed from which is grown detected region. A morphological reconstruction algorithm is used to find highly contrasted pixels in the image, then "dark" pixels are selected as pixels with intensity included in the $[t_1;t_2]$ interval, with $t_1=\mu+0.5\sigma$ *and* $t_2=\mu+1.5\sigma$ (where μ is the mean of pixel values, and σ the standard deviation). Examples of pixels considered as "dark" by this method are shown on second column of Figure 6.

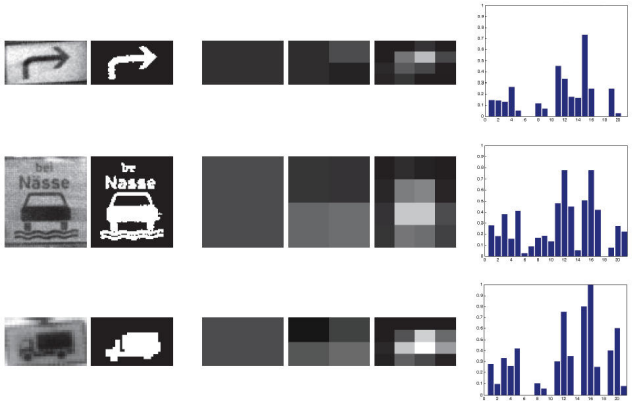

Figure 6. Illustration of the « darkness proportion » computed on same pyramid as pHoG: left, the original image, and its inverse binarized version; middle, visualization of the darkness proportions in the various sub-regions for the 3 levels of the pyramid; right, the resulting vector.

We nickname pHoG-Dp this new proposed descriptor concatenating the standard pHoG with the "darkness proportion" information computed on the same pyramid. The new pHoG-Dp descriptor allows significant classification improvement, as can be seen on precision-recall curves of figure 7.

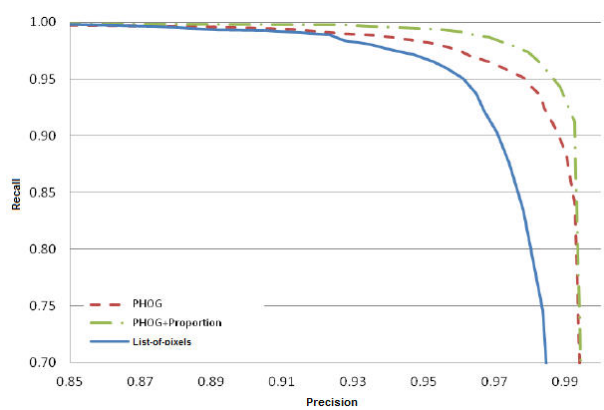

Figure 7. Comparison of classification performance of 3 descriptors: pixels of rescaled image (bottom curve, in blue), pHoG (middle in dashed red), and pHoG-Dp (top curve, in green dash-point).

IV. Experiments and Results

In order to evaluate the classification accuracy of our approach, we use a large database of sub-signs as described in Table 1. As mentioned in previous section, supplementary signs were grouped in 4 "meta-classes": Text, Arrow, Pictogram, Mixed.

Table 1. Content of the supplementary signs database used for training and testing.

Category	Examples	Training	Test
Text		*1693*	*697*
Arrow		*267*	*99*
Pictogram		*2030*	*473*
Mixed		*1293*	*521*
Negative		*13856*	*4944*

After training Gaussian kernel SVMs, the classification accuracy is satisfactory, as can be seen on Figure 8, which shows precision-recall curves computed separately for each meta-class. Note that this first evaluation is for recognition alone, as it is done on test images extracted after manual labelling, which emulates a "perfect" detection step.

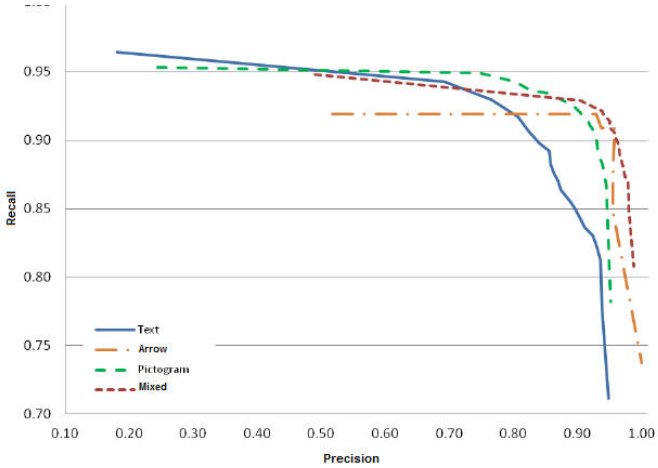

Figure 8. Precision-recall curves separately for each meta-class, computed on test database (from manually-labeled bounding-boxes).

More details on the classification accuracy reached can be found on the confusion matrix computed on test examples of our database, shown in Table 2. The Text category is the one with the lowest precision, which can be explained by some confusions, in particular with guard rails (see top line of Figure 9). The slightly lower precision for "Pictogram" type can be explained by some confusions arising in particular with pictograms contained in main signs lying in the searched sub-region (see bottom line of Figure 9).

Table 2. Confusion matrix.

	Classifier					
Ground Truth	*Neg.*	*Text*	*Arrow*	*Pict.*	*Mixed*	**Recall**
Negative	***4766***	*107*	*0*	*49*	*22*	**96%**
Text	*48*	***632***	*2*	*13*	*2*	**91%**
Arrow	*2*	*4*	***90***	*2*	*1*	**91%**
Pictogram	*23*	*8*	*0*	***440***	*2*	**93%**
Mixed	*24*	*14*	*0*	*9*	***474***	**91%**
Precision	**98%**	**83%**	**98%**	**86%**	**95%**	**95%**

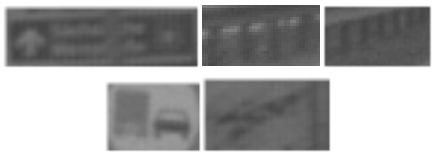

Figure 9. Some examples of negatives confused with "Text" type supplementary sign (top line), and of negatives confused with "pictogram" type (bottom line).

All the above evaluations were conducted on a database in which the bounding boxes around supplementary signs stems from manually-made "ground truth". From the application point of view, it is essential to also test performance of our recognition approach when applied on the candidate sub-sign regions found by our sub-sign detection technique. As described in more details in [11], our detection method is based on region-growing applied from highly contrasted seeds (hopefully lying on border of black pictogram or text of the sub-sign).

Figure 10 shows, for our global chain consisting of detection by region-growing from contrasted seeds followed by recognition with SVM applied on pHoG-Dp, the recall = f(FPPP) curves separately for each meta-class obtained (FPPP = False Positive Per Positive). One can see that maximum recall is rather low (~45%) for Arrow category, essentially due to lower *detection* rate for this type of sub-signs. Maximum recall for Pictogram and Mixed categories are respectively ~65% and ~72%. The Text meta-class has the best maximum recall ~92%, but also the highest false

positive rate, which was expected given that this class had the lowest precision in table 2. The recall rates may seem low, but these are frame-by-frame evaluations, including first "human-detectable" appearance of sub-sign which has often too few pixels to be easily classified by recognition step. In practice however, one does not need to detect and recognize on every frame, but only on enough frames to be confident in output.

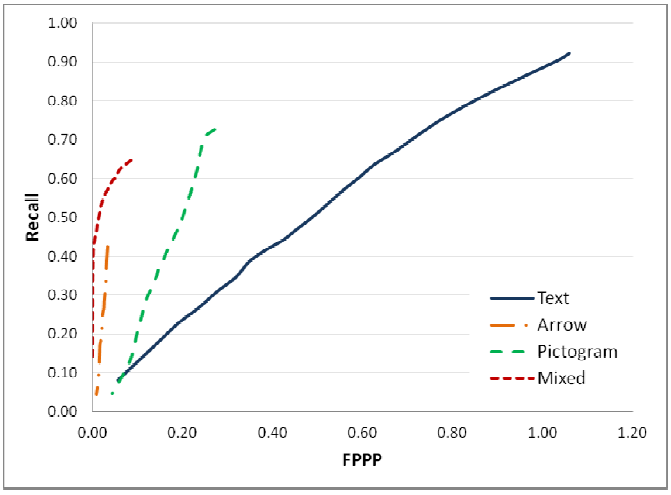

Figure 10. Frame-by-frame evaluation of our complete detection + recognition chain: Recall=f(False Positive Per Positive) curves separately for each meta-class.

In the final application, performance should only count the percentage of physical sub-signs that are "validated" during their appearance. Our validation criterion adopted is that same sub-sign is detected and recognized (with same class!) at least 3 times during the total time-window of sub-sign visibility. The resulting correct detection + recognition + validation rates are reported in Table 3. As one can see the global recall is from 70% for the "Arrow" meta-class to 99% for the "Text" meta-class, with an average recall of 91%. For the same setting, the false alarm rate for signs without any sub-sign placed below is ~12%.

Table 3. Correct detection + recognition + validation rates for each sub-sign meta-class, and false alarm rate on signs without sub-signs.

Meta-class	**Text**	**Arrow**	**Pictogram**	**Mixed**	**ALL SUB-SIGNS**	**No sub-sign**
Total	102	23	55	49	229	278
detection + recognition	**101** ***99%***	**16** ***70%***	**53** ***95%***	**39** ***80%***	**209** **91%**	**34** ***12%***

V. Conclusions, Discussion and Perspectives

We have presented in this paper an original framework for hierarchical recognition of supplementary signs. We used as feature pHoG-Dp, a new proposed original variant of "pyramid HoG", in which we have added the "proportion of dark pixels" evaluated on the same pyramid of sub-regions as the histograms of gradient. Furthermore, to cope with the potentially very large number of supplementary sign types, we proposed to apply a hierarchical classification in which the meta-class of sub-sign (Arrow, Pictogram, Text, Mixed) is first determined.

An experimental evaluation conducted on a large database of sub-signs showed that accuracy of our proposed framework can reach 95% recall and precision on perfectly cropped sub-signs. Furthermore, the global performance when chaining our recognition after our detection was 91% on average on all supplementary signs types (with a false alarm rate of 12% on signs *without* sub-signs).

Future works include training and evaluation of second stage of the hierarchical classifier (in order to recognize precise type of supplementary sign), and experiment on applying optical character recognition (OCR) on the sub-signs identified as "text" type.

Acknowledgements

This work was conducted in the framework of the SpeedCam project (ANR-09-VTT-11), funded by the French ANR (Agence Nationale de la Recherche).

References

[1] de la Escalera A., Armingol J.M. and Mata M., "*Traffic sign recognition and analysis for intelligent vehicles*", Image and Vision Computing, 21:247-258, 2003.

[2] Bahlmann C., Zhu Y., Ramesh V., Pellkofer M. and Koehler T., "*A System for Traffic Sign Detection, Tracking, and Recognition Using Color, Shape, and Motion Information*", IEEE Intelligent Vehicles Symposium (IV'2005), Las Vegas (USA), June 2005.

[3] Broggi A., Cerri P., Medici P., Porta P. and Ghisio G., "*Real Time Road Signs Recognition*", IEEE-IV2007 , Istanbul, Turkey, 13-15 June, 2007

[4] Gavrila D.M., "*Traffic sign recognition revisited*", Proc of 21st DAGM symposium fur Musterekennung, pp. 86-93, Springer-Verlag, 1999.

[5] Barnes N. and Zelinsky A., "*Real-time radial symmetry for speed sign detection*", Proc. IEEE Intelligent Vehicle Symposium, pages 566-571, Parma (Italy), 2004.

[6] García-Garrido M. A., Sotelo M. A., and Martín-Gorostiza E., "*Fast traffic sign detection and recognition under changing lighting conditions*", Proc. IEEE Intelligent Transportation Systems Conference, pages 811-816, Toronto (Canada), 2006.

[7] Moutarde F., Bargeton B., Herbin A. and Chanussot C., "*Robust on-vehicle real-time visual detection of American and European speed limit signs, with a modular Traffic Signs Recognition system*", proc. of IEEE Intelligent Vehicles Symposium, Istanbul, 13-15 juin 2007.

[8] Hamdoun O., Bargeton A., Moutarde F., Bradai B. and Chanussot L., "*Detection and Recognition of End-of-Speed-Limit and Supplementary Signs for Improved European Speed Limit Support*", proc. of 15th World Congress on Intelligent Transport Systems (ITSwc'2008), New-York (USA), nov. 2008.

[9] Nienhüser D., Gumpp T., Zöllner J. M. and Natroshvili K., "*Fast and Reliable Recognition of Supplementary Traffic Signs*", proc. IEEE Intelligent Vehicles Symposium (IV'2010), San Diego (USA), june 2010.

[10] Liu W., Lv J., Gao H., Duan B., Yuan H. and Zhao H., "*An Efficient Real-Time Speed Limit Signs Recognition Based on Rotation Invariant Feature*", proc. of IEEE Intelligent Vehicles Symposium (IV'2011), Baden-Baden (Germany), june 2011.

[11] Puthon A.-S., Moutarde F. and Nashashibi F., "*Subsign detection with region-growing from contrasted seeds*", proc. 15th IEEE Intelligent Transportation Systems Conference (ITSC'2012), Anchorage (USA), sept. 2012

[12] Bosch A., Zisserman A. and Munoz X., "*Representing shape with a spatial pyramid kernel*", Proceedings of the 6th ACM international conference on Image and video retrieval (CIVR'07), Amsterdam (Netherlands), July 2007.