Lane Marking Extraction with Combination Strategy and Comparative Evaluation on Synthetic and Camera Images

Evangeline Pollard, Dominique Gruyer, Jean-Philippe Tarel, Sio-Song Ieng, Aurélien Cord

***Abstract*— Lane detections and tracking are crucial stages for a great number of Advance Driving Assistance Systems (ADAS), for instance for road lane following or robust ego localization. In these applications, the most important module is probably the lane marking primitives extraction algorithm. Since several decades, a lot of approaches have been proposed in order to achieve this task. Unfortunately, it is yet difficult to guarantee robust results from these extraction algorithms in case of bad weather conditions, degraded lane markings, or due to intrinsic limitations of cameras. In this paper we propose an approach in order to improve the quality of the lane marking extraction. By extraction, we mean the classification of the image pixels into two classes: marking and non-marking. The extraction is generally the first step of a marking detection system, so its efficiency has a strong impact on the performances of the whole system. The proposed algorithm is based on the combination of two different extraction algorithms. In order to validate the quality of this work, some tests and evaluations are provided and allow to prove the efficiency of such an approach. The evaluation is performed on camera images and then on synthetic images. The results with camera and synthetic images are compared and discussed.**

I. INTRODUCTION

If video-based lane detection is studied for over thirty years [1], [2], [3], it is not only because of the computer vision challenge but also because of many industrial applications for increasing traffic safety and driver comfort. Amongst them, we can quote lane departure warning [4], localization improvement [5], Self Localization And Mapping (SLAM), or trajectory planning [6] for the driving automation.

Lane detection consists in estimating the shape of lanes over time by using white lane markers, including the estimation of the number of lanes. A typical lane detection algorithm is a three step process. The first is the extraction of lane marking points relying on dedicated image features calculation. The second consists on the association of extracted points to lanes, including the number of lanes estimation. And the third ensures the temporal monitoring in order to provide an estimation of lanes shape evolution. Some post-processing steps may additionally be used to insure a cooperation between the three steps. For example, a closed loop feedback can be used, in which the previous tracked lanes are used as an *a priori* position for extracting marking feature points [7].

As pointed out in [8], the evaluation of such lane detection algorithm is seldomly tackled in literature. Most of the papers only evaluate performances on the final lane estimation, considering it is too hard to obtain ground truth for intermediate steps. Obviously, the first step, i.e the lane extraction, is crucial to the whole process: better it is, more accurate the lane description will be. Though, we focus, in this paper, on the lane extraction and on its evaluation. The used feature for the lane extraction is directly the image pixel. Our main contribution in this paper concerns the way to use conjointly two lane marking extraction algorithms. Pixels with high contrast with respect to the background are extracted by the first extraction algorithm in order to limit false alarms and then the pixels around with lower contrast are extracted to reduce false negative.

As appropriate evaluation is required before use in any Intelligent Transportation System (ITS) system, we evaluate the proposed algorithms along two different scenarios:

1) on a reference database (available online in [9]), containing more than one hundred images and combining different difficulty levels (high curves, non-uniform shadows, sun reflections),
2) on synthetic realistic images of high quality coming from the sensor simulator called SiVIC [10],

The comparison of the results obtained with several algorithms on synthetic and camera images allows us to analyze if the produced synthetic images are enough realistic for the evaluation of lane marking extraction algorithms.

Consequently, the paper is structured as follows. First, standard reference extraction algorithms are summarized in Sec. II, as well as the way to combine them. Then, in Sec. III, we report results on our evaluation scenarios experiments. Finally, a conclusion and some future works are provided.

II. EXTRACTION ALGORITHMS

From [8], [11], the best extraction algorithms are all based on local thresholding, where lane feature extraction is based on the assumption that lane pixels form bright regions surrounded by darker regions. Extraction purpose is consequently to elaborate a way to detect these bright regions with a gradient of intensity higher than a threshold T_G with respect to the background. We also suppose that lane marking width is bounded by the interval $[S_m, S_M]$ in pixels in the bottom of the image.

E. Pollard is with the Department of Electrical Engineering, Université de Sherbrooke, Canada evangeline.pollard@usherbrooke.ca

D. Gruyer, A. Cord are with the LIVIC research laboratory, IFSTTAR, 14 route de la Minière, bat 824, 78000 Versailles, France; email: firstname.name@ifsttar.fr

J.-P. Tarel and S.S. Ieng are with the LEPSIS research laboratory, IFSTTAR, 58 Bld Lefebvre, 75015 Paris, France; email : firstname.name@ifsttar.fr

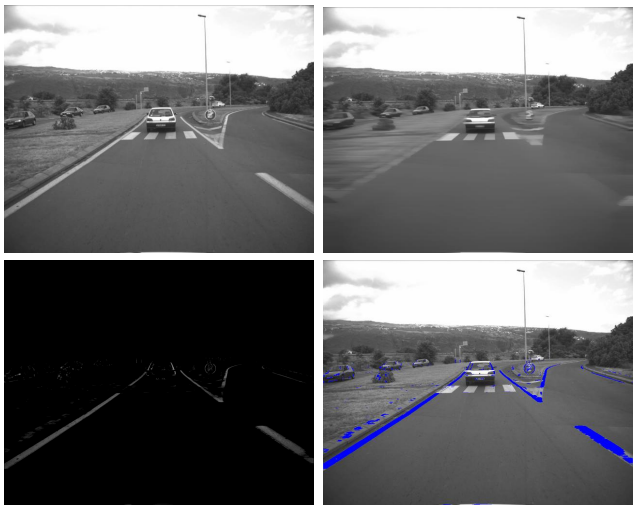


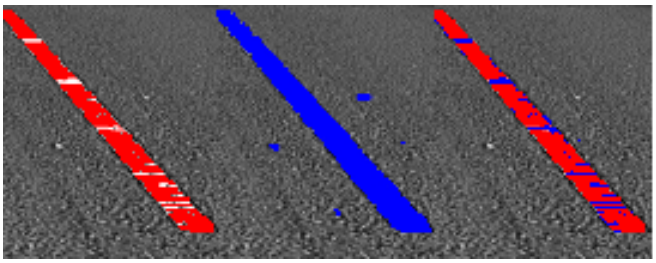
Fig. 1. 1^{st} line: original image and median filtered image taking into account the perspective effect. 2^{nd} line: original image after local threshold $T_G = 20$ using the filtered image and the extraction result (extracted marking pixels are in blue).

A. Single thresholding algorithms

In the local threshold extraction algorithms described in [8], [11], the first step consists in filtering the original image I to remove all the bright objects with an horizontal width small enough to be lane markings. It produces an image of the dark background of the bright objects, we named the filtered image $\bar{I}$. The original image is filtered line-by-line, independently, and the horizontal width S of the filter is linearly modulated for each line to take into account the perspective effect the road is subject to, when observed by a frontal camera. In the second step, the difference between original image I and filtered image $\bar{I}$ is thresholded by a value T_G to produce a binary image of the marking pixels. Last, the third step consists in removing the horizontal segments of marking that are too small to be considered as marking elements. These three steps are illustrated in Fig. 1.

Several different filters might be used in the first step. When the mean filter is used, it is called the mean Local Threshold (LT) algorithm [8]. Taking a median filter leads to the so-called Median Local Threshold (MLT) algorithm [12]. When using the 43rd percentile, instead of the 50th as in the median, it is the 43rd Percentile Local Threshold (PLT) extraction algorithm [11].

In [8], another extraction algorithm was introduced which leads to good performances: the so-called Symmetrical Local Threshold (SLT). Its difference is in the way the threshold of the second step is performed. Rather than testing if $I(y) > \bar{I}(y) + T_G$, where y is the pixel abscissa on a row, the test is symmetrically performed on the neighborhood of the filtered image, *i.e* both $I(y) > \bar{I}(y - S/2) + T_G$ and $I(y) > \bar{I}(y+S/2)+T_G$ must be verified to select a pixel as in lane markings. We recall that S is the width of the horizontal filter. The advantage of the SLT extraction algorithm, over previously described algorithms, is a reduction of false alarms in case of strong step edges.



(a) (b) (c)

Fig. 2. (a) result after extraction algorithm 1 with a large threshold value T_G, (b) after extraction algorithm 2 with a small threshold value, (c) proposed combination.

In case of a color input image, rather that using the corresponding gray level image, it is shown in [13] that it is better to use as input the minimum value over the three color channels, when the white balance is performed on white lane-markings.

B. Double extraction scheme

As each extraction algorithm can be seen as a source of information about lane feature positioning, they can be conjointly used in order to improve performances. As feature marking extraction is a binary classification problem, performances are studied by using the quadruplet $(\mathrm{TP}, \mathrm{FP}, \mathrm{TN}, \mathrm{FN})$ where TP corresponds to the number of *True Positive*, FP to *False Positive*, TN to *True negative* and FN to *False Negative*. Then, the main objective is to reduce FP and to increase TP. Using a simple disjunctive operator would produce limited results. In fact, if $(\mathrm{TP}_i, \mathrm{FN}_i)$ denotes the number of TP and FP from extraction algorithm i $(\forall i \in \{1, 2\})$ and $(\mathrm{TP}^d, \mathrm{FP}^d)$ are for the combination of both, then a disjunctive operator leads to the following performances:

$$\mathrm{TP}^d \leq \min(\mathrm{TP}^1, \mathrm{TP}^2) \qquad \mathrm{FN}^d \geq \max(\mathrm{FN}^1, \mathrm{FN}^2) \quad (1)$$

That is why, we propose to use a more sophisticated technique based on a morphological dilation. Let E the space of 2D-binary images. First, each extraction algorithm i runs independently to each other and generates an extraction, denoted $X^i \subset E$. Let B_x the structuring element centered on x, defined as a square of side S_m:

$$B_x = \{x + b | b \in B\}, \forall x \in X \quad (2)$$

where $B = \{x = (x_1, x_2) \in X | \max(|x_1|, |x_2|) \leq S_m\}$. This structuring element is used to dilate X^1, and the output is intersected with X^2. The result we name $\delta_B(X^1|X^2)$ can thus be written as:

$$\begin{aligned} \delta_B(X^1|X^2) &= (X^1 \oplus B) \cap X^2 \\ &= \left\{x + b | b \in B, x \in X^1, x + b \in X^2\right\} \end{aligned} \quad (3)$$

where $\oplus$ denotes the dilation and $\cap$ denotes the intersection operators.

The first extraction X^1 should use a large threshold value T_G^1 to limit FP. The value of the threshold T_G^2 used for the

second extraction X^2 should be smaller than T_G^1 to increase TP, as illustrated in Fig. 2.

It is important to underline that any kind of pixel based extraction algorithms can be used in the proposed double extraction scheme. The extraction algorithm used with the large threshold value can be the same or different from the extraction algorithm used with the small threshold value. The proposed scheme shares some similarity with the algorithm proposed in [14] where the region of the extracted pixels using a large threshold value is extended using a morphological dilation. The classification is then refined for each pixel in the dilated region using another rule on its intensity. The algorithm proposed in [14] is thus less generic, in the sense that it cannot be applied to any extraction algorithm. Moreover, it is directly based on the intensity, when we usually use the intensity relative to the background which is much more robust in case of shadows.

It will be shown in the following how our strategy leads to improved extraction results in term of false alarm and is able to reduced sensitivity of results to the gradient threshold T_G. This is especially interesting when associated to an adaptive thresholding in order to process image sequences with various lighting conditions.

III. RESULTS

Performances are evaluated on a set of camera images and synthetic sequences. For static images, results are evaluated with respect to the influence of threshold T_G. For video sequences, performances are evaluated over time.

A. Evaluation metrics

Following [8], two classical tools are used: Receiver Operating Characteristic (ROC) and Dice Similarity Coefficient (DSC). ROC Curves are obtained by plotting the True Positive Rate (TPR) *vs.* the False Positive Rate (FPR) for different values of the threshold T_G, which are defined as follows:

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \qquad \mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP}+\mathrm{TN}} \tag{4}$$

As the number of pixels corresponding to lane marking is small (about $1-2\%$), only the left part of the ROC curve is meaningful. Moreover, as for any detection problem, the main issue is to find a compromise between the TPR and the FPR. The DSC curve is thus complementary plotted as a function of T_G. It is commonly defined as:

$$\mathrm{DSC} = \frac{2\mathrm{TP}}{(\mathrm{TP}+\mathrm{FP})+(\mathrm{TP}+\mathrm{FN})} \tag{5}$$

The DSC curve quality can be observed through two criteria. First, the maximum value should be highest as possible, because it corresponds to an optimal value of the threshold T_G. Second, the width of the peak should be large, because it informs on the extraction algorithm sensitivity to the threshold value.

B. Quantitative experimental results on static images

Performances of extraction algorithms and double extraction scheme algorithms are first evaluated on a online database called ROMA [9]. This database contains 116 colored images captured by the Network of the French Transportation Department. This image set depicts the variety of common difficulties linked to road scenes which any ITS has to deal with, such as lighting condition changes, shadow, variable road types and signs, as shown on Fig. 3.

As ROMA images are in color, the extraction algorithms are applied on each color channel and then combined, by using a logical conjunctive operator.

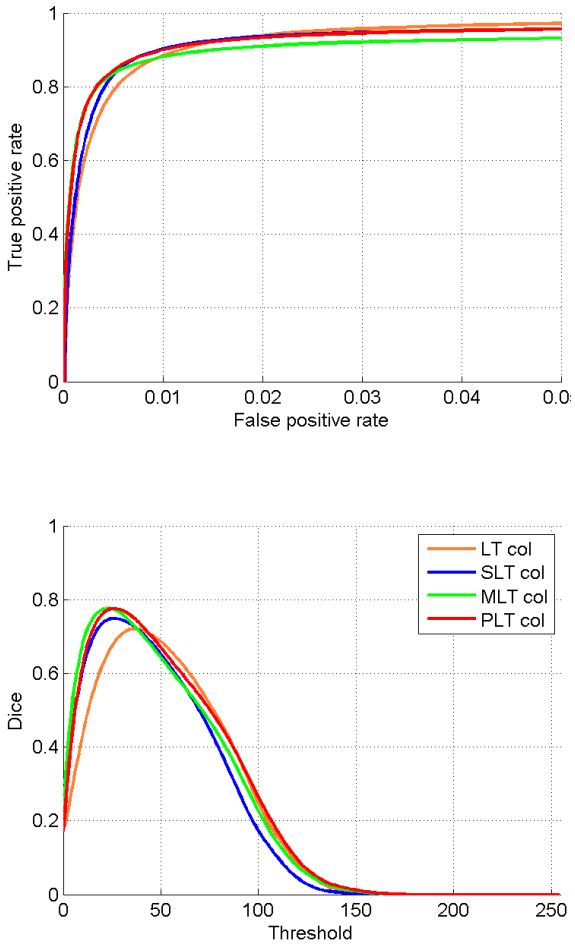


Fig. 4. ROC (top) and DSC (bottom) curves for single extraction algorithms.

Max DICE value	Single [5; 20] pix.	Self Comb. [5; 20] pix.	MLT-X [5; 20] pix.
LT	0.711	0.720	0.786
SLT	0.736	0.749	0.791
MLT	0.775	0.777	
PLT	0.774	0.776	0.778

TABLE I

MAXIMUM VALUE OF THE DICE FOR SINGLE AND DOUBLE EXTRACTION SCHEME ALGORITHMS.

(a) (b) (c) (d) (e)

Fig. 3. Images depicting the variety of encountered difficulties: (a) strong illumination, (b) varying depth lane marking, (c) bad visibility, (d) highly cluttered shadows and dazzling light, (e) faded lane marking and high curvatures.

The four extraction algorithms described in Sec. II (LT, MLT, PLT, SLT) were applied on the whole database for range $[S_m, S_M] = [5cm, 20cm]$. For each algorithm, the performance figure consists in the ROC curves and then the DSC curve. Then, we experiment with different combinations of double extraction scheme algorithms. Fig. 4 recalls performances of LT, MLT, PLT, SLT extraction algorithms, and confirms their good performances, the MLT and PLT being better than the two others. This can be seen in the first column of Tab. I with the achieved maximum values of the Dice.

Each extraction algorithm is first combined with himself for different threshold values. The second threshold T_G^2 is chosen as the value which gives the best Dice value of the extraction algorithm alone. The obtained ROC and DICE curves are slightly better, but the improvement is not really significant. This is summarized in the second column of Tab. I.

In Fig. 5, combinations of different extraction algorithms are evaluated. MLT extraction algorithm, which has the higher maximum value of the Dice as a single extraction algorithm, is combined with the other extraction algorithms: LT, PLT and SLT. We obtained improved results with double extraction scheme algorithms. At first, ROC and Dice curves are slightly higher. From the third column of Tab. I, the combination MLT-SLT increases the maximum value of the Dice by more than 2% in comparison with single extraction algorithms. It appears, that it is more efficient to combine algorithms which are relatively different (such as MLT and LT or MLT and SLT) than algorithms which are similar (such as MLT and PLT which are quite similar) in order to increase TP, without increasing FP or FN. The reason is that only false positives not located at the same place in both extraction results are removed by the combination.

We now concentrate on performances of the double MLT-SLT extraction algorithm which gives the best performances with MLT-LT. The obtained performance improvement is related to the choice of the value T_g^2. SLT extraction algorithm has its maximum Dice value for $T_G^2 = 26$. On Fig.6, ROC and Dice curves are also plotted for T_G^2 values between 5 to 50 (in olive or cyan). The best maximum Dice value is founded for $T_G^2 = 27$ (Dice=0.791). It corresponds to an improvement of more than 2% in comparison to the dice of MLT single extraction algorithm. With smaller or equal value than 26, the ROC curve is permanently higher than for the

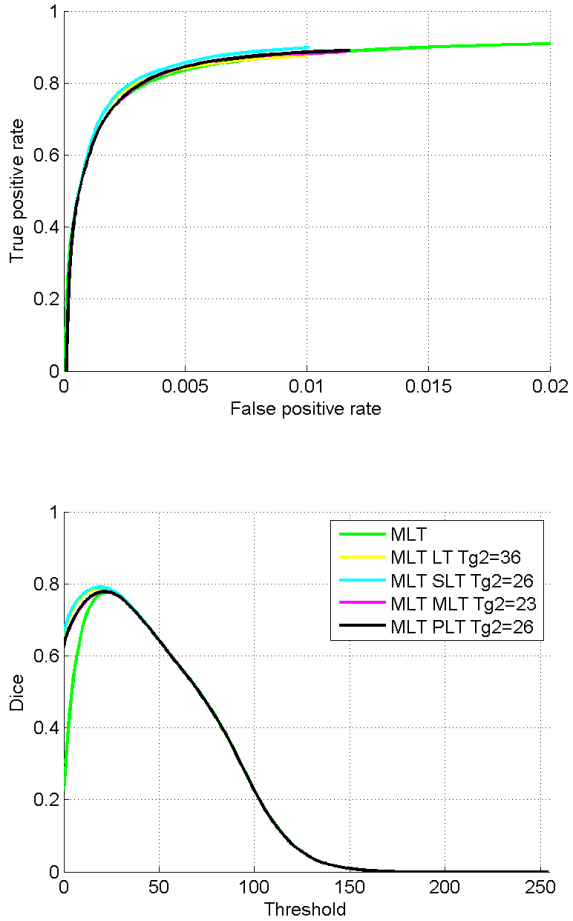


Fig. 5. ROC (top) and Dice (bottom) curves for double MLT-X extraction algorithms.

MLT algorithm. When T_G^2 is close to 0, combined extraction algorithm performances become similar to the one of the MLT With T_G^2 values greater than optimum, performances are damaged only after $T_G^2 = 45$, where the maximum Dice value is lower than for the MLT and the TPR is limited to 0.78 (*vs.* 0.84).

C. Quantitative simulated results

Synthetic images come from a virtual embedded frontal camera simulated with SiVIC$^{\text{TM}}$ platform. SiVIC$^{\text{TM}}$ models a virtual road environment including vehicles, obstacles, infrastructures and sensors. It enables a virtual prototyp-

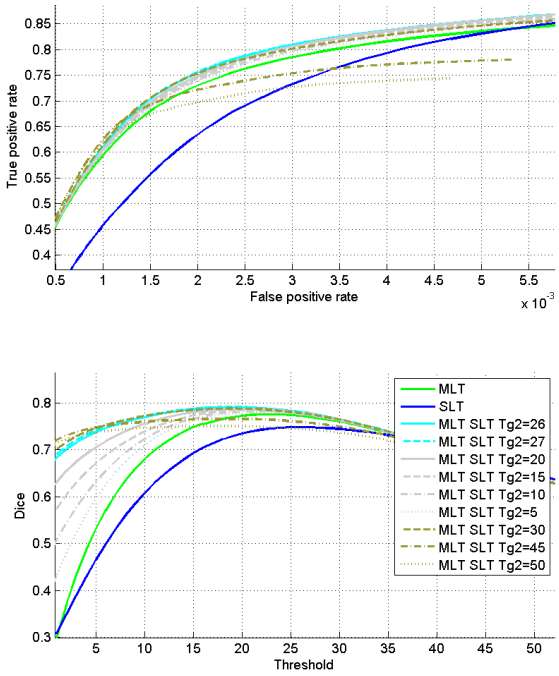


Fig. 6. Zoom on ROC (top) and Dice (bottom) curves for double MLT-SLT extraction algorithm for different values of T_G^2.

ing and testing of a wide range of sensors (exteroceptive, proprioceptive) [15]. Concerning images sequences, they are visually realistic as shown on Fig. 7. This platform is used in order to build a realistic environment (Satory's track in Versailles, France) with specific test scenarios.

Max DICE value	**Single** $[5; 20]$ pix.	**Self Comb.** $[5; 20]$ pix.	**MLT-X** $[5; 20]$ pix.
LT	0.670	0.671	0.740
SLT	0.708	0.711	0.741
MLT	0.721	0.728	
PLT	0.707	0.711	0.727

TABLE II

MAXIMUM DICE VALUE FOR SINGLE OR DOUBLE EXTRACTION SCHEME ALGORITHMS.

Goals of this simulation are double. First, we want to show the similarity of results with camera and synthetic images to ensure that the realism is enough for testing such kind of extraction algorithms. Second, we want to confirm the interest of using double MLT-SLT and MLT-LT extraction algorithms on image sequences.

Fig. 8 and 9 present ROC and Dice curves for Sivic sequences for range $[S_m, S_M] = [5pi., 40pi.]$. Curve shapes are quite similar as for the ROMA database. MLT extraction algorithms better perform in both cases. LT extraction algorithm is classified fourth in both cases. The difference is that performances of PLT and SLT algorithms are quite similar on synthetic images when performances are rather different on camera images. We again obtained that the best combinations are MLT-LT and MLT-SLT on synthetic images. This similarity of results between the camera and synthetic images give

Fig. 7. Actual image and corresponding synthetic image on the Satory test track, France.

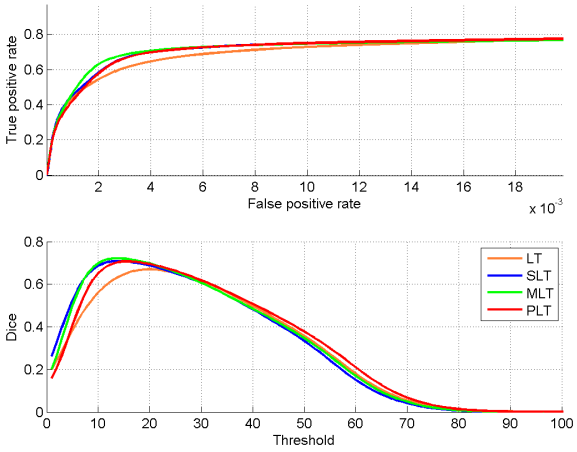


Fig. 8. ROC (top) and Dice (bottom) curves for single extraction algorithms on Sivic sequences.

support that the realism of the synthetic images is enough for testing lane marking extraction algorithms. The combinations MLT-LT and MLT-SLT provides an improvement of 2.7%, see Table II, in comparison to single extraction algorithms.

In Fig. 10, the maximum dice value is show as a function of time along a synthetic sequence. It is clear from this figure, that the use of the double MLT-SLT extraction algorithm provides a significant improvement comparatively to single extraction algorithms.

IV. CONCLUSIONS AND FUTURE WORKS

In this paper, we proposed to combine several extraction algorithms to improve performances in lane feature extraction. To rigorously validate the proposed scheme and algorithms, we proposed an evaluation on both sets of images: a set of camera images with a hand-made ground-truth and a

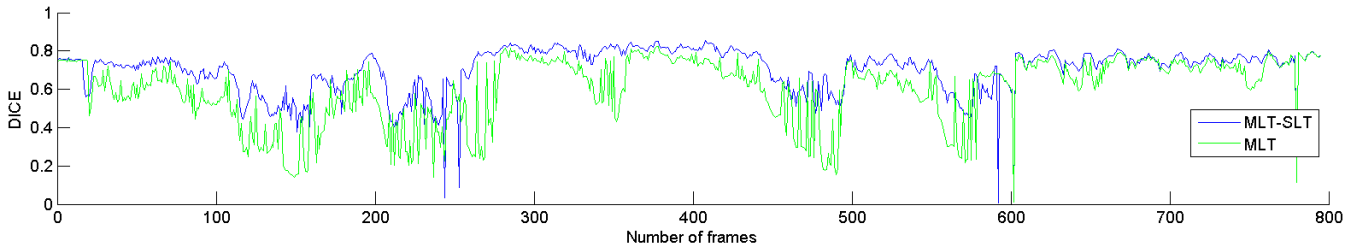


Fig. 10. Variation of the maximum Dice value along a Sivic sequence with double MLT-SLT extraction algorithm.

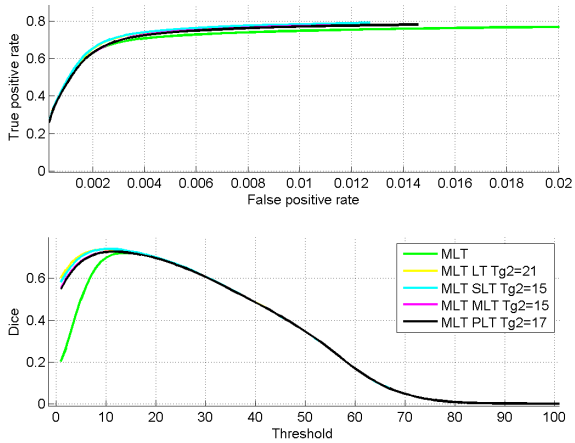


Fig. 9. ROC (top) and Dice (bottom) curves for double extraction algorithms on Sivic sequences.

sequence of synthetic images with a semi-automatic ground-truth. We show that the proposed scheme is valuable only when the two combined extraction algorithms are relatively different. The combination consists in a dilation on the extraction output with a large gray level threshold intersected with the output with a small gray level threshold. In this way, false alarms are partially removed on a wide threshold range. The consistency between the results obtained with camera images and with synthetic images suggests that the generated synthetic sequences are enough realistic to be used in lane marking extraction evaluations.

Future works on the extraction aim at the automation of the selection of the threshold value. More generally, this work is the first part of the evaluation of lane marking detection system. Our idea is to evaluate this process from the extraction step to the lane marking tracking step. Moreover, the proposed extraction algorithm can be used for all road markings by increasing S_M similarly to [13].

ACKNOWLEDGEMENT

This work is part of CooPerCom, a 3-year international research project (Canada-France) and of ABV, a French research project for low speed automation. The authors would like to thank the National Science and Engineering Research Council (NSERC) of Canada and the Agence nationale de la recherche (ANR) in France for supporting these projects.

REFERENCES

[1] K. Kluge and S. Lakshmanan, "A deformable-template approach to lane detection," in *Proceedings of the Intelligent Vehicles Symposium.*, sep 1995, pp. 54 –59.

[2] J. McCall and M. Trivedi, "Video based lane estimation and tracking for driver assistance: Survey, system, and evaluation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 7, no. 1, pp. 20–37, 2006.

[3] Z. Kim, "Robust lane detection and tracking in challenging scenarios," *IEEE Transactions on Intelligent Transportation Systems*, vol. 9, no. 1, pp. 16 –26, march 2008.

[4] J. W. Lee, "A Machine Vision System for Lane-Departure Detection," *Computer Vision and Image Understanding*, vol. 86, no. 1, pp. 52 – 78, 2002. [Online]. Available: http://www.sciencedirect.com/science/article/B6WCX-473VKG0-3/2/8a98760b2e9588b90fefc359dcd9ae27

[5] M. Bertozzi and A. Broggi, "Gold: a parallel real-time stereo vision system for generic obstacle and lane detection," *IEEE Transactions on Image Processing*, vol. 7, no. 1, pp. 62 –81, jan 1998.

[6] S. Glaser, B. Vanholme, S. Mammar, D. Gruyer, and L. Nouvelie andre, "Maneuver-based trajectory planning for highly autonomous vehicles on real road with traffic and driver interaction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 3, pp. 589 –606, sept. 2010.

[7] S. Ieng, J. Vrignon, D. Gruyer, and D. Aubert, "A new multi-lanes detection using multi-camera for robust vehicle location," in *IEEE Proc. on Intelligent Vehicles Symposium*. IEEE, 2005, pp. 700–705.

[8] T. Veit, J. Tarel, P. Nicolle, and P. Charbonnier, "Evaluation of road marking feature extraction," in *IEEE 11th International Conference on Intelligent Transportation Systems*. IEEE, 2008, pp. 174–181.

[9] [Online]. Available: http://www.lcpc.fr/english/products/image-databases/article/roma-road-markings-1817

[10] D. Gruyer, C. Royere, N. du Lac, G. Michel, and J. Blosseville, "SiVIC and RTMaps, interconnected platforms for the conception and the evaluation of driving assistance systems," in *ITS World Congress, London, United Kingdom*, 2006.

[11] Y. Sebsadji, J.-P. Tarel, P. Foucher, and P. Charbonnier, "Robust road marking extraction in urban environments using stereo images," in *Proceedings of IEEE Intelligent Vehicle Symposium (IV'2010)*, San Diego, California, USA, 2010, pp. 394–400.

[12] J. Ninot, J.-P. Tarel, T. Gavrilovic, L. Smadja, and K. Heggarty, "Amélioration de la reconnaissance des marquages routiers par l'optimisation d'algorithmes d'extraction," in *Proceedings of colloque COGIST'09*, St Quay Portrieux, France, 2009.

[13] P. Foucher, Y. Sebsadji, J.-P. Tarel, P. Charbonnier, and P. Nicolle, "Detection and recognition of urban road markings using images," in *14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, Oct. 2011.

[14] R. Danescu and S. Nedevschi, "Detection and classification of painted road objects for intersection assistance applications," in *13th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, sept. 2010, pp. 433 –438.

[15] D. Gruyer, N. Hiblot, P. Desouza, H. Sauer, and B. Monnier, "A new generic virtual platform for cameras modeling," in *Proceeding of International Conference VISION*, 2010.