Real-time detection and classification of road lane markings

Mauricio Braga de Paula*†, Claudio Rosito Jung*
*Institute of Informatics - Federal University of Rio Grande do Sul
Email: mbpaula@inf.ufrgs.br, crjung@inf.ufrgs.br
†Mathematics and Statistics Department - Federal University of Pelotas
Email: maubrapa@ufpel.edu.br

***Abstract*—This paper presents a method for detection and recognition of road lane markings using an uncalibrated onboard camera. Initially, lane boundaries are detected based on a linear-parabolic model. Then, we build a simple model to represent pixels related to the pavement, and explore this model to estimate pixels related to lane markings. A set of features is computed based on the detected lane markings, and a cascade of binary classifiers is adopted to distinguish five types of markings: dashed, dashed-solid, solid-dashed, single-solid and double-solid. Experimental results show that the proposed method presents good classification results under a variety of situations (shadows, varying illumination, etc.).**

***Keywords*-lane detection; lane markings; onboard vehicular cameras; driver assistance system;**

I. Introduction

Traffic safety is a major concern in present days, particularly in underdeveloped and developing countries. According to the World Health Organization (WHO) [1], 90% of the deaths related to traffic accidents occur in low-income and middle-income countries, in a total of more than 1.2 million deaths per year and 50 million injuries every year.

Brazil is considered as an emerging country, presenting vast natural resources and a strong potential for development and industrial production. As stated in the Brazilian National Agency of Land Transport [2], Brazil has approximately 1.7 million kilometers of road network. A study from 2009 [3] reported that more than 33% of Brazilian roadways were considered poor or very poor with respect to their overall condition (signing, geometry and pavement) in 2007, and that the number of traffic-related accidents in 2006 was over 35,000 in Brazil.

The annual road safety report [4], that was presented in 2011 by the International Traffic Safety Data and Analysis Group (IRTAD), shows that road deaths keep decreasing in most IRTAD countries (mostly developed countries), carrying forward the significant reductions in the number of road deaths accomplished in 2008 and 2009. On the other hand, a study by the Brazilian Confederation of Countries (CNM) [5] shows the high mortality rate due to traffic accident in Brazil. Death ended accident ratio in Brazil is 2.5 times more than United States, and 3.7 times more than Europe with respect to population (See Table I).

[1]Mortality rate per 100,000 inhabitants.

TABLE I
Road fatalities in 2008 [5]

Country	Deaths	Population (millions)	Mortality rate[1]
Brazil	57,116	189.6	30.1
United States	37,261	304.0	12.5
European Union	38,876	498.0	7.8

As cited by the IRTAD group, many countries have adopted road safety strategies for reduction of head-on / frontal, rear-end crashes and side collision. These types of vehicle crashes are the main reason of traffic accidents, resulting in at least one death.

In the computer vision research community, there has been significant effort in the last years towards the development of vision-based approaches for intelligent roads and intelligent vehicles [6], [7]. In particular, the development os vision-based lane detection and lane departure warning systems [8] are important to warn the driver when the vehicle begins to move out of its lane, since accidents due to lane crossings are common and may potentially lead to frontal collisions. In that context, the detection and recognition of horizontal lane markings is an important issue, since they indicate road portions were takeovers (voluntary lane changes) are allowed or not.

This paper presents a new approach to detect and classify horizontal lane markings. Our approach employs a lane detection/tracking algorithm, and then applies a cascaded binary classifier to recognize five different lane marking types. The remainder of this paper is organized as follows: Section II revises some existing approaches for detection and recognition of lane markings, focusing on onboard vehicular cameras. The proposed approach is presented in Section II, and results are presented in Section IV. Finally, our conclusions are drawn in Section V.

II. Related Work

Lane change or overtaking and passing another vehicle are one of the most dangerous driving maneuvers and have recently been studied extensively for applications in computer vision in order to help the driver in the driving process. During a journey, the driver might encounter a number of different types of lane boundaries markings, including single (dashed

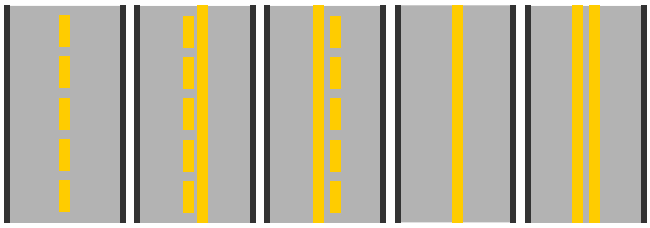


Fig. 1. Types of lane boundaries: (a) Dashed. (b) Dashed-solid. (c) Solid-dashed. (d) Single solid. (e) Double solid.

or solid) or double (dashed and solid, solid and dashed or only solid) painted lines, as illustrated in Fig. 1.

In [9], two types of Advanced Driving Assistance Systems (ADAS) are presented: (i) Lane Change Assistance (LCA) is a kind of system that warns the driver against collisions that may occur due to a lane change maneuver; and (ii) and Lane Departure Warning (LDW) that warns the driver when an unintentional lane departure is about to occur. These systems also monitor lane markings and, consequently, a good lane departure system will be able to detect the type of lane boundaries.

Although there many approaches for lane tracking [8], not much work has been devoted to identify lane markings. Collado et al. [10] split the problem into few parts: create a bird-eye view of the road, segment the pixels which belong to longitudinal road markings, extract the lane boundaries by the Hough Transform and realize adjusts in the pitch angle. Then, lane boundaries are classified in continuous (also called solid), broken (known as dashed) or merged (solid-dashed or dashed-solid, with no distinction between them), computed by the power spectrum of the Fast Fourier Transform. This scheme needs a previous camera calibration for the initial procedure that generates the bird-eye view.

A ridge measure was used to detect the lane markings in [11]. The dashed type must have high values along the center line of a region of interest and low values near the boundaries of the road. The related paper present results with the presence of continuous (solid) and discontinuous (dashed) lane markings. However, they do not deal with solid-dashed, dashed-solid or double-solid lane markings, which are common in two-way roads.

Another approach for extracting road markings was proposed by Zhao et al. [12] based on threshold segmentation. The model initially creates a bird-eye view of the road, segments the remapped image through an adaptive threshold, and applies some geometrical constraints to remove spurious responses. The output is a binary image with potential lane markings, but no higher-level features or classification is performed.

Chira et al. [13] present a system for detection, measurement and classification of painted objects in real-time using edge detection and geometric pattern matching. Their focus is mostly on horizontal traffic signs (forward arrow, forward right, forward left, right arrow and dashed lane marking), not being able do discriminate different lane marking types.

In this paper, our goal is to classify each portion of the road monitored by an onboard camera inside a moving vehicle into five possible types of lane markings, as shown in Fig. 1: dashed, dashed-solid, solid-dashed, single-solid and double-solid. As far as we know, this is the first approach that tries to classify lane markings into these five categories (in particular, dashed-solid, solid-dashed, and double-solid markings are common in two-way roads, and indicate when overtaking is possible or not). The proposed approach is presented next.

III. Proposed Approach

The first step of the proposed approach is to detect and follow lane boundaries [14]. Given the detected lane boundaries, we extract some statistical properties of pavement-related pixels based on a rectangular patch placed in between detected lane boundaries, and use such parameters to detect pixels related to lane-markings on a rectangular Region of Interest (ROI) centered at the bottom of the lane boundary. As the vehicle moves, the temporal evolution of extracted lane markings provides cues on the type of lane marking, which are recognized using a cascaded classifier (output classes are ω_1, ω_2,..., ω_5). These steps are explained in details next.

A. Road marking extraction

To estimate pixels related to lane markings, we firstly assume that the intensity of lane markings is larger than the intensity of the road pavement. Next, we estimate the width W of the lane (in pixels) at the bottom of the image by computing the horizontal distance between the detected lane boundaries. To estimate the statistics of pavement-related pixels, we build a rectangular patch r_i with dimensions $0.03W \times 0.20W$, and place it at a distance of $0.25W$ pixels from the lane boundary to the interior of the lane (another possibility would be to place it exactly at the center of the lane, but we decided to place it closer to the lane boundary to avoid the interference of vehicles in front), computing its mean intensity value μ and standard deviation σ.

We also consider an external rectangular region r_e centered at the bottom-most pixel of the lane boundary with the same dimensions as r_i (See Fig. 2), and we check the consistency of each pixel (x, y) in patch r_e with the distribution of pavement-related pixels. More precisely, we build a binary patch $r_b^t(x, y)$ at each frame t defined as

$$r_b^t(x,y) = \begin{cases} 1, & \text{if } r_e(x,y) > \mu + k\sigma \\ 0, & \text{otherwise} \end{cases}, \qquad (1)$$

where $k = 5$ is the tolerance (number of standard deviations, obtained experimentally) used to distinguish pavement and lane marking pixels.

B. Road marking classification

The patches $r_b^t(x, y)$ contain binary data related to the existence of lane markings within the ROI for each frame t. A set of features is extracted from $r_b^t(x, y)$, aiming to discriminate all the five classes in a hierarchical model. In fact, a three-level cascaded classifier was developed, as illustrated in

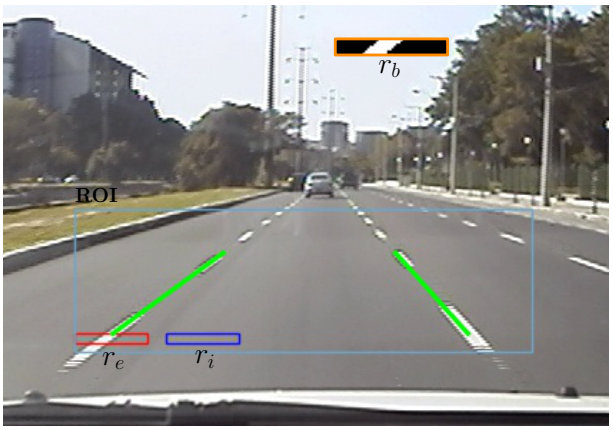


Fig. 2. Typical frame and rectangular regions used to estimate the length of lane markings, along with the detected lane boundaries (green).

Fig. 3, and next we described our choices for the four binary classifiers $\mathcal{C}_1$, $\mathcal{C}_2$, $\mathcal{C}_3$ and $\mathcal{C}_4$, as well as the feature vectors used for each classifier. Also, let Ω_{i1} and Ω_{i2} denote the two possible output classes of classifier $\mathcal{C}_i$, for $i = 1, 2, 3, 4$.

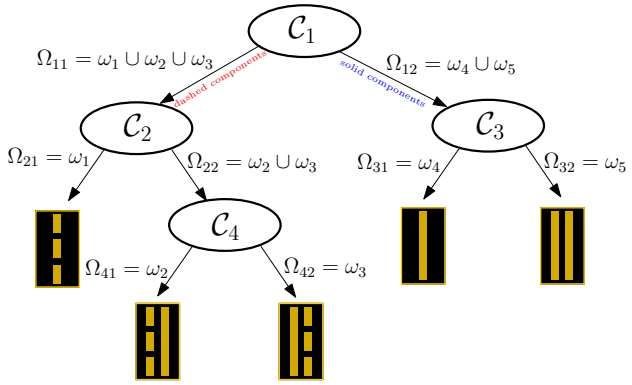


Fig. 3. Schematic illustration of the cascaded classifier used in our approach.

1) Classifier $\mathcal{C}_1$: The first classifier $\mathcal{C}_1$ aims to separate lane markings with dashed components (Ω_{11}) from those with solid components only (Ω_{12}). There are a few parameters that can be used to discriminate these two classes: in road portions with solid components only, $r_b^t(x, y)$ tends to present roughly the same number of pixels related to lane markings; in contrast, the number of such pixels in road portions containing dashed components tends to present more variations (and in a periodic fashion). In this work, $\mathcal{C}_1$ is fed with a three dimensional feature vector $\boldsymbol{f}_1(t) = (f_{11}(t), f_{12}(t), f_{13}(t))^T$, containing elements that quantify the variations mentioned above.

The fraction of marking-related pixels (within the rectangular patch) at each frame t is given by

$$m(t) = \frac{\sum_{x,y} r_b^t(x, y)}{\#r_b^t}, \tag{2}$$

where $\#r_b^t$ is the total number of pixels in r_b^t.

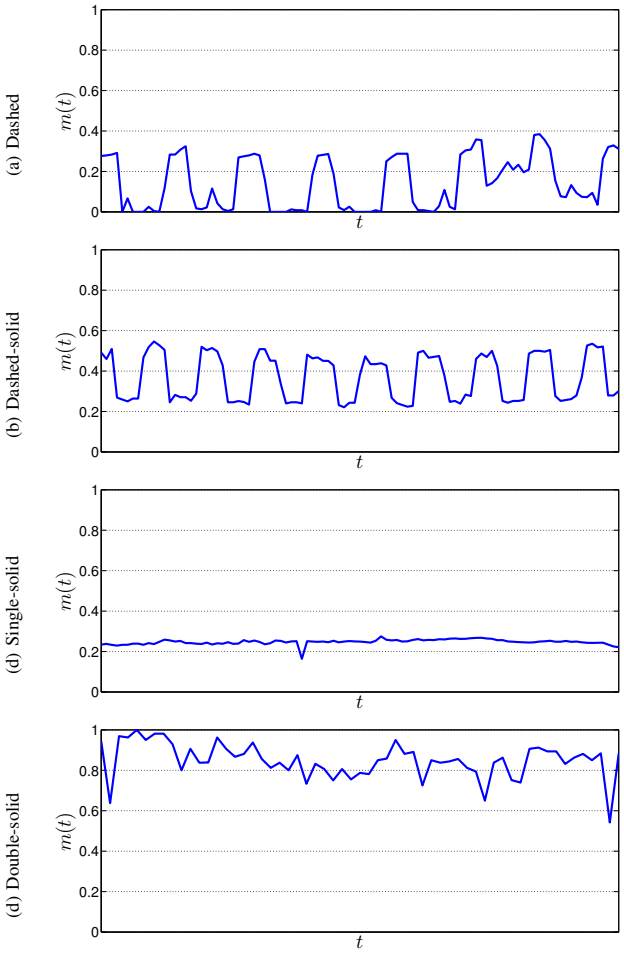


Fig. 4. Plots of the function $m(t)$ for different lane marking types.

Function $m(t)$ should oscillate when dashed components are presents, and remain roughly constant when only solid lane marking components are present, as illustrated in Fig. 4[2]. Based on this fact, for each frame t we evaluate a temporal window $\mathcal{T}(t) = \{t-T+1, t-T+2, \cdots, t-1, t\}$ and compute the weighted amplitude variation of $m(t)$ within the window:

$$f_{11}(t) = \frac{1}{\mu_{\mathcal{T}(t)}\{m\}} \left(\max_{\mathcal{T}(t)}\{m\} - \min_{\mathcal{T}(t)}\{m\} \right), \tag{3}$$

where $\mu_{\mathcal{T}(t)}$, $\max_{\mathcal{T}(t)}$ and $\max_{\mathcal{T}(t)}$ represent the average, maximum and minimum values within $\mathcal{T}(t)$, respectively.

To explore features related to the possible periodicity of $m(t)$, we compute an autocorrelation-like measure of m within the $\mathcal{T}(t)$:

$$R_{mm}^t(\tau) = \frac{1}{T'} \sum_{u=0}^{T'-1} m(t-u)m(t-u-\tau), \tag{4}$$

where $T' < T$ is the overlap window between $m(t-u)$ and $m(t-u-\tau)$, and $\tau = 0, 1, ..., T-T'-1$ is the range of

[2]In Fig. 4, the plot for a solid-dashed marking was omitted, since it is very similar to the dashed-solid case.

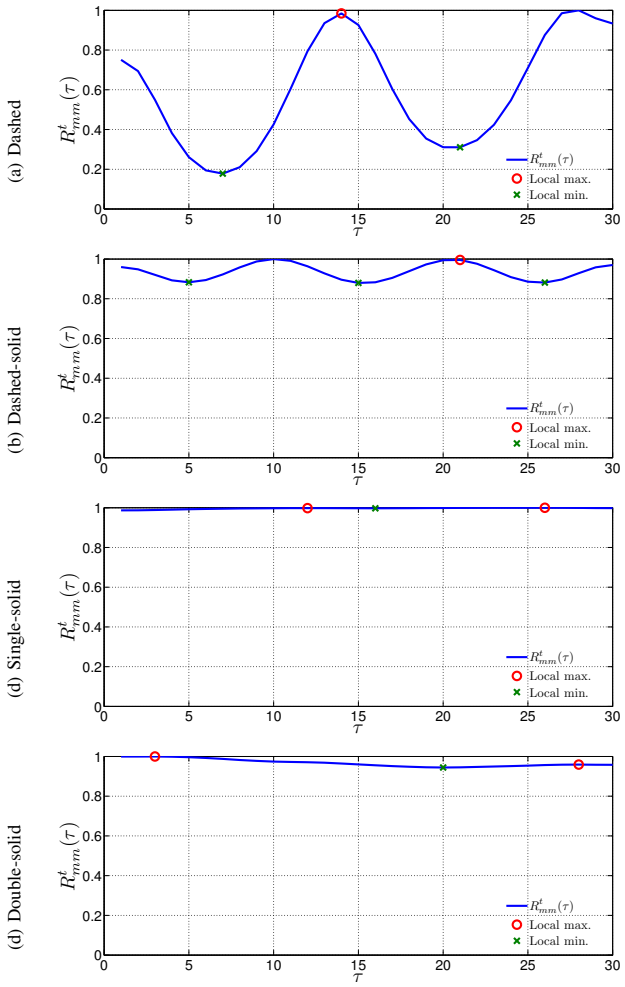


Fig. 5. Plot of autocorrelation function $R^t_{mm}(\tau)$ for different lane marking types (normalized by the largest autocorrelation value).

signal displacements that can be computed with the finite-size signal. If $m(t)$ is periodic, $R^t_{mm}(\tau)$ should produce a sharp peak when τ equals the value of the period. On the other hand, constant functions generate constant autocorrelation functions. As an illustration, the autocorrelation functions related to the plots of $m(t)$ shown in Fig. 4 are depicted in Fig. 5.

Since our estimates $m(t)$ are noisy, larger support values T' to compute the autocorrelation function lead to better noise suppression. On the other hand, if T' (and consequently T) are too large, the values $m(t)$ used in the analysis may correspond to portions of the road with different lane marking types (i.e., the signal $m(t)$ may not be stationary). As a compromise, we used $T = 100$ and $T' = 70$, defined experimentally.

Given a frame t, we compute $R^t_{mm}(\tau)$, and retrieve the first local maximum $R^t_{mm}(\tau_{max})$[3] and the first local minimum $R^t_{mm}(\tau_{min})$, computing the normalized difference

$$f_{12}(t) = \frac{1}{R^t_{mm}(0)} \left(R^t_{mm}(\tau_{max}) - R^t_{mm}(\tau_{min}) \right) \quad (5)$$

[3]To avoid spurious local maxima, we impose that $R^t_{mm}(\tau_{max})$ must be at least half of $R^t_{mm}(0)$.

as the second feature. For lane markings with only solid components, $f_{12}(t)$ is expected to be small, since the autocorrelation function should be (ideally) constant. On the other hand, $f_{12}(t)$ is expected to be larger when dashed components are involved, due to the periodic nature of $m(t)$.

Also, if the distance between dashed lane markings is constant within a portion of the road, the position of the local maximum should occur approximately at the same location, which corresponds to the period of $m(t)$. On the other hand, when only solid markings are present, $R^t_{mm}(\tau)$ is expected to be roughly constant, and the first local maximum may occur at spurious locations due to small fluctuations in $R^t_{mm}(\tau)$. Hence, the third feature used in $\mathcal{C}_1$ explores the variation of the local maximum position within $\mathcal{T}(t)$. Although the standard deviation provides such metric, we decided to use:

$$\begin{aligned} f_{13}(t) &= \underset{u \in \mathcal{T}(t)}{\text{MAD}} \{\tau_{max}(u)\} \\ &= \underset{u \in \mathcal{T}(t)}{\text{median}} \{ |\tau_{max}(u) - \underset{u \in \mathcal{T}(t)}{\text{median}} \{\tau_{max}(u)\}| \} \end{aligned} \quad (6)$$

where $\tau_{max}(u)$ is the location of the first local maxima of $R^u_{mm}(\tau)$, and MAD is the Median Absolute Deviation, which is less affected by outliers than the standard deviation.

As for the classifier $\mathcal{C}_1$ itself, we have chosen to use a Support Vector Machine (SVM), and our experimental results indicated that the Radial-Based-Function (RBF) kernel presented the best results. More details on training and test data will be presented in Section IV.

2) Classifier $\mathcal{C}_2$: As shown in Fig. 3, classifier $\mathcal{C}_2$ is fed with samples that have dashed components according to $\mathcal{C}_1$ (classes ω_1, ω_2 and ω_3), and classify them as dashed ($\Omega_{21} = \omega_1$) or solid-dashed/dashed-solid ($\Omega_{22} = \omega_2 \cup \omega_3$).

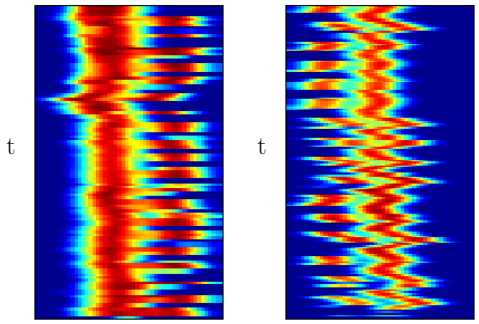
For that purpose, we first estimate the amount of lane markings detected along the horizontal direction. More precisely, for each frame t we compute

$$\rho^t_b(x) = \sum_y r^t_b(x, y), \quad (7)$$

which provides the accumulated lane marking evidence in the vertical direction y within the rectangular ROI. For dashed lane markings, the shape of $\rho^t_b(x)$ will alternate from a flat (low) plateau when no lane marking is locally present at frame t, to a single-peaked plot when a lane marking is present. On the other hand, in dashed-solid or solid-dashed markings, $\rho^t_b(x)$ will alternate from a single-peaked plot (solid component only) to a double-peaked plot (solid and dashed component). As an illustration, Fig. 6 shows image plots corresponding to the surfaces $\rho^t_b(x)$ as a function of t and x for solid-dashed (Fig. 6(a)) and dashed-solid (Fig. 6(b)) portions of the road.

We then compute the number of peaks (local maxima) $n_p(t)$ present in $\rho^t_b(x)$ along the x axis, and define as feature f_2 for classifier $\mathcal{C}_2$ the average number of peaks within a temporal window T_2:

$$f_2(t) = \frac{1}{T_2} \sum_{u=0}^{T_2-1} n_p(u), \quad (8)$$



(a) (b)

Fig. 6. Example of accumulated lane marking evidence $\rho_b^t(x)$ for (a) solid-dashed and (b) dashed-solid lane marking. High values in red, low values in blue.

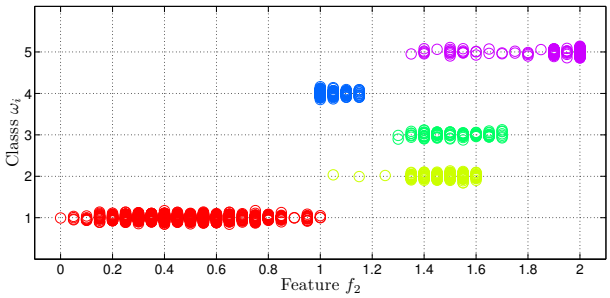


Fig. 7. Distribution of feature f_2 for samples belonging to all five classes ω_i. Samples from different classes are shown in different horizontal lines.

where $T_2 = 20$ frames was defined experimentally. Clearly, for samples belonging to dashed markings, $f_2(t)$ should be close to 0.5, whereas samples related to dashed-solid or solid-dashed markings should present values $f_2(t)$ close to 1.5. Hence, a simple rule is adopted for classified $\mathcal{C}_2$: if $f_2(t) < 1$, the sample is assigned to class ω_1; otherwise, it is assigned to class Ω_{22} (classifier $\mathcal{C}_4$ will discriminate ω_2 and ω_3).

3) Classifier $\mathcal{C}_3$: The goal of $\mathcal{C}_3$ is to discriminate between single-solid ($\Omega_{31} = \omega_4$) and double-solid lane markings ($\Omega_{32} = \omega_5$). In single-solid lane markings, the plot of $\rho_b^t(x)$ at all frames t is expected to produce a single peak related to the lane marking, while two distinct peaks are expected when double-solid markings are present. Hence, we use $f_3(t) = f_2(t)$ for classifier $\mathcal{C}_3$, as defined in Equation (8). The decision rule is also very similar to $\mathcal{C}_2$: if $f_3(t) < 1.5$, the sample is assigned to class ω_4; otherwise, it is assigned to class ω_5.

It is important to point out that the same feature $f_2(t)$ is used in both classifiers $\mathcal{C}_2$ and $\mathcal{C}_3$, and its expected value for classes Ω_{21}, Ω_{22}, Ω_{31} and Ω_{32} are, respectively, 0.5, 1.5, 1 and 2. Hence, the reader might wander if a single multi-class classifier based on $f_2(t)$ could replace $\mathcal{C}_1$, $\mathcal{C}_2$ and $\mathcal{C}_3$. Although this is in fact possible, results are worse than using the proposed cascade: classes Ω_{21} - Ω_{22} and Ω_{31} - Ω_{32} presents larger pairwise separation (which is important for the chosen two-class classifiers) than all the four classes together in a single multi-class classifier. In fact, Fig. 7 shows the distribution of f_2 for different classes, and it can be observed that classes Ω_{22} and ω_5 present considerable overlap, and classes Ω_{21} - Ω_{22} and Ω_{31} - Ω_{32} presents significant pairwise separation.

4) Classifier $\mathcal{C}_4$: The last classifier in the cascade aims to distinguish dashed-solid ($\Omega_{41} = \omega_2$) and solid-dashed markings ($\Omega_{42} = \omega_3$). In both cases, the number of peaks $n_p(t)$ along the x axis in $\rho_b^t(x)$ alternates between one and two as a function of t. However, in the former case the position of the first peak is consistent for all frames t, and the second peak arises only at some frames (please see Fig. 6(a)); in the latter, the opposite behavior is expected: the second peak is consistent, and the first one appears at some frames only (please see Fig. 6(b)).

To detect such behavior, we maintain two "buffers" $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ with the position(s) of the peak(s) in the previous frames. When only one peak is detected at a given frame t, its position is compared to the last values stored in $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$, respectively, and assigned to the buffer that presents the smallest difference (in the other buffer, the value -1 is stored to indicate the absence of a peak at that frame).

At each frame t, we count the number of valid peak positions (i.e. values different than -1) at each buffer in the last T frames: the buffer $\boldsymbol{p}_i$ with the largest number of valid peaks relates to the solid marking, while buffer $\boldsymbol{p}_j$ relates to dashed markings. We then compute the average peak position $P_{p_i}(t)$ and $P_{p_j}(t)$ of each buffer (considering only valid peaks), and decide for class ω_2 (dashed-solid) if $P_{p_i}(t) < P_{p_j}(t)$, and for class ω_3 (solid-dashed) otherwise.

IV. Experimental Results

In our experiments, we have used 10 video sequences acquired with two different cameras, 2 sequences from the publicly available Environment Perception and Driver Assistance dataset[4] - set3 (Suburban Bridge and Trailer Follow) and one video from the Cambridge-driving Labeled Video Database[5] (CamVid) (See Table II).

For the first classifier $\mathcal{C}_1$, we use a Gaussian Radial Basis Function (RBF) kernel, defined by:

$$K_\gamma(x, y) = e^{-\gamma \|x-y\|^2}, \tag{9}$$

where $\gamma > 0$ is the width parameter. We have used a SVM with flexible margin, controlled by parameter $C > 0$. We have experimentally evaluated different values for both C and γ with the holdout cross-validation randomly selecting a fraction $g < 1$ for training (the training set is indicated in Table III) the model[6], and a fraction $1-g$ for validation. The achieved cross validation accuracy was 100.0% and the estimated parameters were: $C = 1$ and $\gamma = 4$.

[4] http://www.mi.auckland.ac.nz/

[5] http://mi.eng.cam.ac.uk/research/projects/VideoRec/CamVid/

[6] We used $g = 0.5$ in all experiments.

TABLE II
VIDEO SEQUENCES USED

Video sequence	Acquired	Public	Resolution
Assis Brasil	x		480×640
BR-116 1: Pelotas - POA	x		480×640
BR-116 6: Pelotas - POA	x		480×720
BR-290	x		480×640
Cambará	x		240×320
Cambridge		x	600×800
Castelo Branco	x		480×640
Germany	x		240×320
Ipiranga	x		480×640
RS-040	x		480×640
RS-287: São Francisco - Canela	x		240×320
Suburban Bridge		x	640×800
Trailer Follow		x	640×800

The training set was composed by 527 instances containing the three dimensional feature vector $\boldsymbol{f}_1(t)$. Its important to note that the video sequences used in the training process were split in multiple clips containing the same type of the lane marking. For test set 1, we used the remaining 527 samples including other video clips (126 frames), each one with a single type of lane marking. To evaluate the transition between different lane marking types, another test set (called number 2) containing longer video sequences interspersing different kinds of markings was used. The characteristics of the clips used in the two datasets are summarized in Table III.

TABLE III
VIDEO CLIPS

Excerpt	Type	Frames	Train	Test 1	Test 2
Assis Brasil	dashed	86	x	x	
BR-116 (1)	dashed	86	x	x	
BR-116 (6)	double solid	86	x	x	
BR-290	dashed	86	x	x	
Cambará (1)	dashed/solid	36	x	x	
Cambará (3)	dashed/solid	66	x	x	
Cambará (5)	solid/dashed	55	x	x	
Cambridge (1)	single solid	46	x	x	
Cambridge (2)	dashed	46	x	x	
Castelo Branco	dashed	86	x	x	
Germany	dashed	86	x	x	
Ipiranga (1)	dashed	86	x	x	
Ipiranga (2)	dashed	36	x	x	
RS-040 (1)	more than one	10797			x
RS-287 (1)	solid/dashed	35	x	x	
RS-287 (2)	more than one	706			x
Suburban (1)	dashed	126		x	
Suburban (2)	single solid	46	x	x	
Suburban (4)	more than one	717			x
Trailer Follow	dashed	86	x	x	

The confusion matrix related to classification results for test set 1 are shown in Table IV. The global accuracy for test set 1 was 99.32%, with just a samples from class ω_1 being classified as ω_2 or ω_4.

As mentioned before, test set 2 contains longer video sequences with different lane marking types. Since all the features used in the proposed classification scheme explore a temporal window, transitions among the different lane types are problematic (the temporal window contains samples from different lane markings). The plots in Fig. 9 illustrate the

TABLE IV
CONFUSION MATRIX FOR TEST SET 1

Output Class \ Target Class	ω_1	ω_2	ω_3	ω_4	ω_5	
ω_1	**802**	0	0	0	0	100%
ω_2	7	**102**	0	0	0	93.58%
ω_3	0	0	**90**	0	0	100%
ω_4	1	0	0	**92**	0	98.92%
ω_5	0	0	0	0	**86**	100%
	99.01%	100%	100%	100%	100%	**99.32%**

ground truth values (over time) for the three video sequences used in the test set, as well as the classification results produced by the proposed approach.

The first plot shows that there are a few fluctuations around the classes, but results are mostly coherent. We can also observe that there is a lag when there are changes in the lane marking type. In fact, this behavior is expected, since we use a temporal window of $T = 100$ frames to compute our features (the lag is approximately 50 frames, which is half of the temporal window size). The full confusion matrix is shown in Table V, and the worst results were the misclassification of several samples from class ω_5 as ω_2, related to the change of lane marking type around frame 550. Nevertheless, the overall classification rate was 71.53%.

TABLE V
CONFUSION MATRIX OF *RS-287 (2)*

Output Class \ Target Class	ω_1	ω_2	ω_3	ω_4	ω_5	
ω_1	**199**	2	0	0	0	99.00%
ω_2	13	**160**	0	0	79	63.49%
ω_3	15	22	**100**	0	0	72.99%
ω_4	0	2	22	**0**	21	0.00%
ω_5	0	0	25	0	**46**	64.79%
	87.67%	86.02%	68.03%	—	31.51%	**71.53%**

In the second video, *Suburban (4)*, we can also observe the lag during changes in the lane marking types (around frames 280 and 750). Table VI shows the confusion matrix for this experiment, and the classification rate was 77.27%.

The last plot shown in Fig. 9(c) relates to the longest video sequence (*RS-040 (1)*, with over 10,000 frames). As in the two previous examples, it also presents some spurious fluctuations around the classes, particulary around frames 4000-5000 and 8700-10000. Part of these fluctuations can be explained by the presence of horizonal traffic signs (as illustrated in Fig. 11), so that the training patch used for extracting lane markings does not fall on unmarked pavement regions. The overall classification rate for this long video sequence was 85.42%, and the full confusion matrix is shown in Fig. VII. It is also important to point out that none of the frames of this video sequence were used to train the SVM in classifier $\mathcal{C}_1$.

The number of correctly classified objects is highly dependent on the quality of the road markings.

The proposed system was implemented in C++ (GCC 4.2.1 compiler), using the **Open** source **C**omputer **V**ision

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

Fig. 8. Examples of different lane marking types contained in dataset 1.

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

Fig. 10. Frames of sequence RS-287 (2).

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

Fig. 11. Frames of sequence RS-040 (1).

TABLE VI
CONFUSION MATRIX OF *Suburban (4)*

		Target Class					
		ω_1	ω_2	ω_3	ω_4	ω_5	
Output Class	ω_1	**196**	0	0	12	0	94.23%
	ω_2	0	**0**	0	0	0	—
	ω_3	0	0	**0**	116	0	—
	ω_4	35	0	0	**358**	0	91.09%
	ω_5	0	0	0	0	**0**	—
		84.85%	—	—	73.66%	—	**77.27%**

TABLE VII
CONFUSION MATRIX OF *RS-040 (1)*

		Target Class					
		ω_1	ω_2	ω_3	ω_4	ω_5	
Output Class	ω_1	**1156**	6	319	0	47	75.65%
	ω_2	0	**374**	76	0	623	34.86%
	ω_3	8	44	**5296**	0	26	98.55%
	ω_4	0	37	95	**0**	253	0.00%
	ω_5	0	2	38	0	**2397**	98.36%
		99.31%	80.78%	90.93%	—	71.64%	**85.42%**

library (OpenCV) Version 2.4.4[7] to implement the lane detection/tracking algorithm and to extract the required features, and the libSVM library [15] for the SVM classification. All experiments were conducted on a 2.53GHz Core 2 Duo laptop computer with 8GB RAM memory. The average execution time for each frame, for different resolutions, was described in Table VIII. As it can be observed, the proposed system runs at 30 FPS for resolutions up to VGA (480×640), and code can be further optimized to reduce running time.

[7]http://opencv.willowgarage.com/wiki/

TABLE VIII
AVERAGE EXECUTION TIME.

Resolution	Time (ms)
240×320	7.82 ms
480×640	25.10 ms
480×720	94.53 ms
600×800	165.65 ms

The video results of the identification of the type of the lane markings can be viewed online at http://www.inf.ufrgs.br/~mbpaula/publications/.

V. CONCLUSIONS

This paper presented a real-time algorithm to detect and identify different types of lane markins using an onboard vehicular camera in a fully automatic manner. In the proposed approach, a simple statistical model is used to represent pixels related to the pavement, which is then used to extract lane markings. A set of features is computed based on the temporal evolution of the detected lane markings, and a cascaded

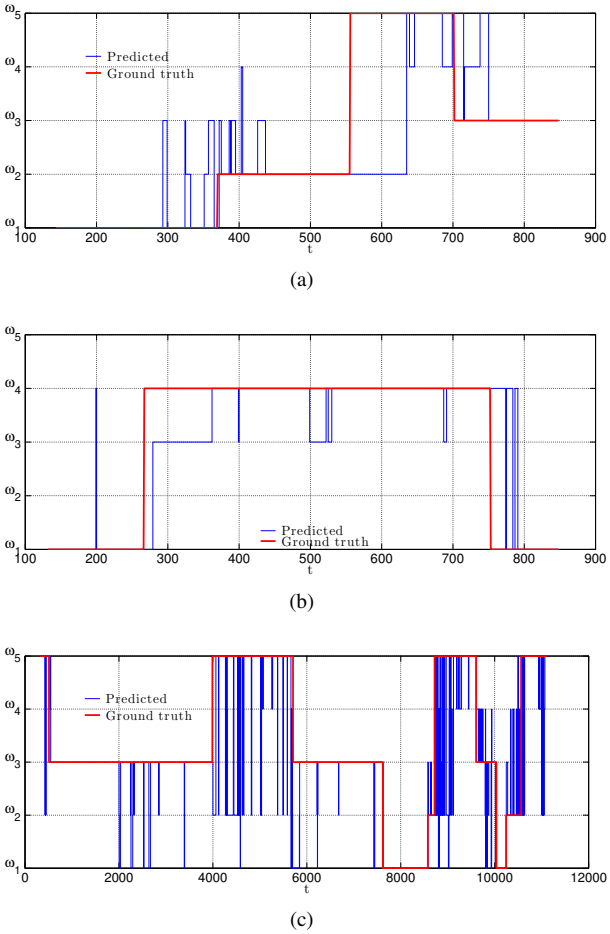


Fig. 9. Ground truth (red) and our results (blue) for three video sequences used in test set 2.

classifier is used to recognize five types of markings (shown in Fig. 1).

Our experimental results show that classification results are very good when using video clips with the same lane marking at all frames, but the performance decreases when longer video sequences with several transitions are present (in particular, there is a lag around the transitions of lane marking types). Nevertheless, we believe that the overall results of around 78.07% for a five-class problem is promising. It is important to mention that classification results are affected by errors along the whole pipeline: strong shadows and/or illumination changes may affect the lane tracker (as well the quality of extracted features), particularly when the training patch and the test patch present different illuminations (See Fig. 10(b) and Fig. 10(c)). Obviously, poorly painted lane markings pose an additional challenge.

Our main goal as future work is to improve classification results when the type of lane marking changes. We think that the key for that is to adjust the size T of the temporal window used to compute the features. At the moment, temporal coherence is induced at the feature level (when more samples of the same lane marking type are present, the extracted features are less noisy). One possibility would be to reduce T and impose temporal coherence after the classification procedure using, for instance, a Hidden Markov Model (HMM).

ACKNOWLEDGMENT

This work was partially supported by Brazilian agencies Capes and CNPq.

REFERENCES

[1] W. H. Organization, *Global status report on road safety*. World Health Organization, 2009. [Online]. Available: http://whqlibdoc.who.int/publications/2009/9789241563840_eng.pdf

[2] ANTT, "Transporte de passageiros," 2013. [Online]. Available: http://www.antt.gov.br/passageiro/apresentacaopas.asp

[3] E. A. Vasconcellos and M. Sivak, "Road safety in Brazil: Challenges and opportunities," The University of Michigan, Tech. Rep. UMTRI-2009-29, 2009.

[4] I. R. Traffic and A. Database, "Road safety annual report 2011," International Road Traffic and Accident Database, Tech. Rep., 2011. [Online]. Available: http://www.internationaltransportforum.org/irtadpublic/pdf/11IrtadReport.pdf

[5] C. N. dos Municípios, "Mapeamento das mortes por acidentes de trânsito mapeamento das mortes por acidentes de trânsito no brasil," Confederação Nacional dos Municípios, Tech. Rep., 2009. [Online]. Available: http://portal.cnm.org.br/sites/9000/9070/Estudos/Transito/EstudoTransito-versaoconcurso.pdf

[6] M. Bertozzi, A. Broggi, M. Cellario, A. Fascioli, P. Lombardi, and M. Porta, "Artificial vision in road vehicles," *Proceedings of the IEEE*, vol. 90, no. 7, pp. 1258–1271, July 2002.

[7] N. Buch, S. Velastin, and J. Orwell, "A review of computer vision techniques for the analysis of urban traffic," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 3, pp. 920–939, 2011.

[8] A. B. Hillel, R. Lerner, D. Levi, and G. Raz, "Recent progress in road and lane detection: a survey," *Machine Vision and Applications*, pp. 1–19, 2012.

[9] C. Visvikis, T. L. Smith, M. Pitcher, and R. Smith, "Study on lane departure warning and lane change assistant systems," Transpor Research Laboratory, Tech. Rep., November 2008. [Online]. Available: http://ec.europa.eu/enterprise/sectors/automotive/files/projects/report_ldwlca_en.pdf

[10] J. M. Collado, C. Hilario, A. de la Escalera, and J. M. Armingol, "Adaptative road lanes detection and classification," in *Proceedings of the 8th international conference on Advanced Concepts For Intelligent Vision Systems*, ser. ACIVS'06. Berlin, Heidelberg: Springer-Verlag, 2006, pp. 1151–1162. [Online]. Available: http://dx.doi.org/10.1007/11864349_105

[11] A. Lopez, C. Canero, J. Serrat, J. Saludes, F. Lumbreras, and T. Graf, "Detection of lane markings based on ridgeness and ransac," in *Intelligent Transportation Systems, 2005. Proceedings. 2005 IEEE*, 2005, pp. 254–259.

[12] Z. Li, Z. xing Cai, J. Xie, and X. ping Ren, "Road markings extraction based on threshold segmentation," in *Fuzzy Systems and Knowledge Discovery (FSKD), 2012 9th International Conference on*, 2012, pp. 1924–1928.

[13] I. Chira, A. Chibulcutean, and R. Danescu, "Real-time detection of road markings for driving assistance applications," in *Computer Engineering and Systems (ICCES), 2010 International Conference on*, 2010, pp. 158–163.

[14] C. R. Jung and C. R. Kelber, "An improved linear-parabolic model for lane following and curve detection," in *Proceedings of SIBGRAPI*. Natal, RN: IEEE Press, October 2005, pp. 131–138.

[15] C.-C. Chang and C.-J. Lin, "Libsvm: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 27:1–27:27, May 2011. [Online]. Available: http://doi.acm.org/10.1145/1961189.1961199