

Detection of Lane Markings based on Ridgeness and RANSAC

A. López[†], C. Cañero[†], J. Serrat[†], J. Saludes[†], F. Lumbreras[†], T. Graf[‡]

Abstract—Detection of lane markings based on a camera sensor can be a low cost solution to lane departure warning and lateral control. However, reliable detection is difficult due to cast shadows, vehicles occluding the marks, wear, vehicle motion, etc. The contribution of this paper is twofold. Firstly, we propose to explore another low-level image descriptor, namely, the *ridgeness*, instead of the gradient magnitude with the aim of getting a more reliable lane marking detection under adverse circumstances. Besides, the proposed measure comes with an associated orientation which is less noisy than the gradient one. Secondly, we have adapted RANSAC, a generic robust estimation method, to fit a parametric model to the image lane lines using both ridgeness and orientation as input data. In short, in this paper a better feature type and a robust fitting method are proposed, which contribute to improve the lane lines detection reliability, and still achieving real-time.

I. INTRODUCTION

AN important challenge of automotive industry is to develop low cost advanced driver assistance systems (ADAS) able to increase traffic safety. Since vision is the most used human sense for driving, some ADAS features rely on camera based systems [1]. For instance, lane departure warning and lateral control can be addressed by detecting the lane markings of the road by means of a forward-facing camera and computer vision techniques. In this paper we focus on this problem, which, actually, is one of the oldest in ADAS. Many papers have been published on it, since it is a difficult and not yet completely solved problem due to shadows, vehicles occluding the marks, partially erased marks, vehicle ego-motion, etc. Basically, the different proposed algorithms have a first step to collect cues on where the lane markings are, and a second step that uses them to fit a lane model. Some type of tracking is commonly added to minimize disturbances from image clutter and facilitate real-time. Since, ideally, lane markings are white lines over a grey pavement, the first step is usually based on image edges, defined as extrema of the gradient magnitude along the gradient direction. Depending on the posterior processing, it is also possible to work directly with the gradient magnitude as an *edgeness* measure. In all cases, gradient direction can be used to remove edges/edgeness having an orientation outside the expected range of values.

However, the gradient magnitude can be misleadingly high due to the contrast between the asphalt and road elements (*e.g.* vehicles) or be low because of shadows, wearied marks,

etc. Moreover, the gradient orientation tends to be noisy because its very local nature. In fact, these usual circumstances are challenging, since for a road with low traffic, well painted lane markings, shadow free, etc., a well-designed computer vision algorithm may succeed. Therefore, methods based on edge detection algorithms must devise strategies to cope with these problems (*e.g.* local adaptive thresholding, hysteresis thresholding, etc.).

The main contributions of this paper are two. The first one, presented in Sect. II, is to explore a different low-level image descriptor, namely, the *ridgeness*, with the aim of having a more reliable lane marking detection under adverse circumstances. Besides, in the same process we obtain an associated orientation which is less noisy than that of the gradient and will be of high usefulness for fitting a lane model. Secondly, we have adapted RANSAC, a generic robust estimation method, to fit a parametric model to the image lane lines using as input data both ridgeness and orientation (Sect. III). The model consists of a couple of hyperbolas, which are constrained to come from parallel lane markings. We claim that a better feature type (ridges) and a robust fitting method (RANSAC) contribute to improve the lane lines detection reliability. In Sect. IV we show current results of the proposed method and discuss its performance. Finally, Sect. V summarizes the main conclusions and future work.

II. LANE MARKINGS AS RIDGES

By ridges of a grey-level image we refer to the center lines of the elongated bright structures appearing in it. In our case, these bright structures are the lane markings, then, their ridges are the longitudinal center of the painted line (Fig. 1 left). This terminology comes from seeing an image as a landscape, since then these center lines correspond to the landscape's ridges (Fig. 1 right). Ridgeness stands for a measure of how much a pixel neighborhood resembles a ridge. Therefore, a ridgeness measure must have high values along the center lines of the lane markings and low values when approaching the longest boundaries of the paint. A binary ridge image, corresponding to the lane marking centerlines, can be obtained by simple thresholding, provided we have a ridgeness measure with a sufficiently well-contrasted and homogeneous response.

There are different mathematical definitions to characterize ridges. However, in [2] we proposed one that compared favorably to others and that we have adapted for the problem at hand. Let $G_{\sigma}(\mathbf{x})$ be a 2D gaussian of standard deviation σ and let $L(\mathbf{x})$ be the grey-level image, being $\mathbf{x} = (u, v)$ where u is a column and v a row. Then, conceptually, our

This research has been partially funded by Spanish MEC project TRA2004-06702/AUT.

[†] Computer Vision Center and Dept. d'Informàtica, Universitat Autònoma de Barcelona. antonio@cvc.uab.es

[‡] Volkswagen AG Group Research, Electronics.

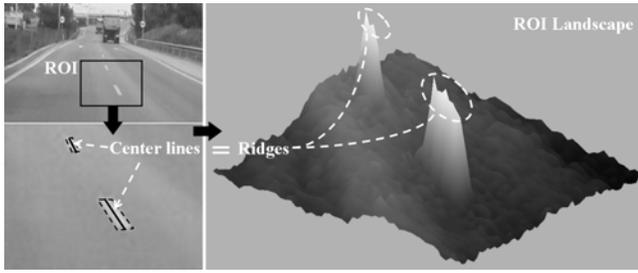


Fig. 1. Left: road image with a region of interest (ROI) outlined. It is shown what we call center lines of the road lane markings. Right: the ROI seen as a landscape. Note that lane markings resemble mountains, and that their ridges correspond to the center lines of the lane markings.

ridge operator is based on the following steps:

- 1) Compute a smoothed version of the image, namely

$$L_{\sigma_d}(\mathbf{x}) = G_{\sigma_d}(\mathbf{x}) * L(\mathbf{x}) ,$$

where ‘*’ denotes convolution.

- 2) Compute the gradient vector field

$$\mathbf{w}_{\sigma_d}(\mathbf{x}) = (\partial_u L_{\sigma_d}(\mathbf{x}), \partial_v L_{\sigma_d}(\mathbf{x}))^T ,$$

- 3) Compute the so-called structure tensor field

$$\mathbf{S}_{\sigma_d, \sigma_i}(\mathbf{x}) = G_{\sigma_i}(\mathbf{x}) * \mathbf{s}_{\sigma_d}(\mathbf{x}) ,$$

being

$$\mathbf{s}_{\sigma_d}(\mathbf{x}) = \mathbf{w}_{\sigma_d}(\mathbf{x}) \cdot \mathbf{w}_{\sigma_d}^T(\mathbf{x}) .$$

- 4) Obtain the eigenvector corresponding to the highest eigenvalue of $\mathbf{S}_{\sigma_d, \sigma_i}(\mathbf{x})$, namely $\mathbf{w}'_{\sigma_d, \sigma_i}(\mathbf{x})$. It is known that $\mathbf{w}'_{\sigma_d, \sigma_i}(\mathbf{x})$ yields the *dominant gradient orientation* of the original image at \mathbf{x} and is perpendicular to the *dominant image orientation* at \mathbf{x} (f.i. if \mathbf{x} is from a lane marking then the dominant image orientation is along it). Therefore, it is a more robust orientation measure than the image gradient itself ($\mathbf{w}_{\sigma_d}(\mathbf{x})$).

It is worth to notice that $\mathbf{w}'_{\sigma_d, \sigma_i}(\mathbf{x})$ defines an orientation field (arrows without head) in the image but for the next step we need a vector field. For this reason we project $\mathbf{w}'_{\sigma_d, \sigma_i}(\mathbf{x})$ into $\mathbf{w}_{\sigma_d}(\mathbf{x})$ as:

$$p_{\sigma_d, \sigma_i}(\mathbf{x}) = \mathbf{w}'_{\sigma_d, \sigma_i}^T(\mathbf{x}) \cdot \mathbf{w}_{\sigma_d}(\mathbf{x}) ,$$

and define the following vector field:

$$\tilde{\mathbf{w}}_{\sigma_d, \sigma_i}(\mathbf{x}) = \text{sign}(p_{\sigma_d, \sigma_i}(\mathbf{x})) \mathbf{w}'_{\sigma_d, \sigma_i}(\mathbf{x}) , \quad (1)$$

where the function $\text{sign}(x)$ returns $+1$ if $x > 0$, -1 if $x < 0$ and 0 if $x = 0$.

- 5) Finally, our *ridgeness* measure is defined as the positive values of

$$\tilde{\kappa}_{\sigma_d, \sigma_i}(\mathbf{x}) = -\text{div}(\tilde{\mathbf{w}}_{\sigma_d, \sigma_i}(\mathbf{x})) , \quad (2)$$

where $\text{div}()$ means divergence of a vector field.

The parameter σ_d is the *differentiation scale* in opposition to σ_i which is the *integration scale*. The former must be tuned to the size of the objects whose orientation has to be determined, while the later must be tuned to the size of the neighborhood we want to use in order to compute dominant orientation.

The positive values of $\tilde{\kappa}_{\sigma_d, \sigma_i}(\mathbf{x})$ give a degree of how much an image pixel resembles a ridge point. In fact, we have shown [3] that these values are in the range $[0, 2]$, where 0 means *it is not a ridge*, about 1 means *it is a ridge*, and 2 means *it is a perfect image maximum*. Besides, we also have shown [3] that these values are homogeneously distributed along the center lines, therefore, facilitating thresholding (if needed). Other properties of $\tilde{\kappa}_{\sigma_d, \sigma_i}(\mathbf{x})$ are invariance under monotonic grey-level transforms and rigid movements of the input image. These properties make ridgeness better suited for lane-marking detection.

We have adapted this operator specifically for the extraction of lane marking points in the following way:

- The ridgeness measure applied to road images must take into account the perspective. In this work we smooth differently depending on the image row, -lower image rows are smoothed more than the upper ones-, in order to avoid smoothing away far line segments.
- Since the dominant orientation of a lane marking is perpendicular to the dominant gradient orientation, and therefore perpendicular to $\tilde{\mathbf{w}}_{\sigma_d, \sigma_i}(\mathbf{x})$, this vector field can be robustly used to discard pixels whose associated orientation is not coherent with the expected by lane markings model.
- The ridgeness measure presents invariance under monotonic grey-level transforms of the input image helps lane detection in presence of shadows, as opposite to what edge based measures do. However, this means that ridgeness also enhances some non-spurious irregularities of the asphalt grey-level. Fortunately, this can be solved up to a large extent by discarding those ridgeness values surrounded by a very low gradient magnitude.

III. DELINEATION OF LANE MARKINGS BY A HYPERBOLIC MODEL FITTED USING RANSAC

A. Lane Markings Model

Horizontal road geometry consists of a combination of straight segments and circular arcs (*i.e.* of constant radius of curvature) connected by transition curves, which usually correspond to clothoids. Mathematically, this means that the curvature (K) is either constant or linearly variable [4]. Lane markings run parallel to road boundaries, therefore following the previous mathematical model. However, since we aim to detect lane markings from images, we must transform the real world model into a projected model. For that purpose we must define the image acquisition geometry. In particular, we use a forward-facing camera placed close to the rear-view mirror of the vehicle. The camera is supposed to be at

- For $i = 1$ to I_{max} , where I_{max} is a predetermined maximum number of iterations (trials):

- Randomly select a subset S_i of N data points from P and estimate the parameters of model M_i .
- Use the instantiated model M_i to determine the subset S_i^* of points in P that are within some error tolerance of M_i . The set S_i^* is called the consensus set of S_i . Therefore, in this step the definition of an error function e_f and an error tolerance e_t is needed.
- If the number of data items of S_i^* is greater than some consensus threshold c_t , use S_i^* to compute (possibly using least squares) a new model M_i^* and terminate with success.

- If no consensus set with c_t or more members has been found, either solve the model with the largest consensus set found, or terminate with failure.

At this point, it is clear that in our problem we have $N = 4$, the P set is composed by the ridgels ($(\mathbf{u}_L, \mathbf{v}_L)$ and $(\mathbf{u}_R, \mathbf{v}_R)$), and model M_i is estimated using equations in (3). Therefore, the main roles to define are I_{max} and e_f .

A selection criterium for I_{max} is also proposed in [7]. Let p_w be the a priori probability that a given data point is within the error tolerance of the correct model (it is not an outlier). If we want to ensure with probability p_z that at least one of our random selections is an error-free set of N data points, then we must expect to make at least T trials (N data points per trial). Therefore, we can set $I_{max} \geq T$, where $T = \log(1 - p_z) / \log(1 - p_w^N)$. However, currently we have not done any process to estimate p_w and we have just experimentally adjusted I_{max} .

With respect to the error function e_f , which defines the selection criteria of the consensus set for each trial, we propose the use of two factors. The former consists of a distance between the candidate ridgel and the estimated model (curve) for the current trial. The latter takes into account the orientation of the model curve and the dominant orientation of the image at the ridgel \mathbf{x}_r . Therefore, the value of $\tilde{\mathbf{w}}_{\sigma_u, \sigma_v}(\mathbf{x}_r)$, computed using Eq. (1), plays a relevant role in our fitting algorithm. Besides, and before checking these factors, we test if the distance between the curves under consideration makes sense according to the expected range of lane widths, and if it is not the case we just discard the current trial and go further to a next one.

Since the curves we search are hyperbolas, as first factor of e_f we propose Sampson's distance between a point and a conic [8]. Given the curves defined by the point subset S_i following equations in (3), Sampson's distance to that curves

from a candidate ridgel $\mathbf{x}_r = (u_r, v_r)$ is given by:

$$d_s(u_r, v_r) = \frac{(k_2 + k_4(-2k_3 + k_1k_4))^2}{4(k_4^2 + (k_3 - k_1k_4)^2)}, \quad (5)$$

where

$$\begin{aligned} k_1 &= \frac{E_u(\mu L - 2x_0)}{E_v z_0}, \\ k_2 &= E_u E_v K z_0, \\ k_3 &= E_v \theta - v_r, \\ k_4 &= u_r + E_u \psi, \end{aligned}$$

with $\mu = +1$ for the left curves, and $\mu = -1$ for the right ones.

For those ridgels fulfilling the Sampson's distance criterium, the second factor of the error function is computed. Given a ridgel \mathbf{x}_r , the point of the conic, say $\mathbf{x}_c = (u_c, v_c)$, that is closest to \mathbf{x}_r is assumed to be at the same image row ($v_r = v_c$), which is a fair approximation. Under this assumption the orientation of the conic (tangent) at \mathbf{x}_c can be computed as:

$$\alpha(u_c, v_c) = \arctan\left(\frac{E_v \theta - v_c}{u_c + c_1 + (E_v \theta - v_c)c_2}\right),$$

where

$$c_1 = E_u \psi \text{ and } c_2 = -\mu \frac{E_v L}{E_v z_0} + \frac{2E_u x_0}{E_v z_0}.$$

Thus, we understand that the orientation of the candidate ridgel $\mathbf{x}_c = (u_c, v_c)$ matches the dominant image orientation (the one expected for the lane markings) at \mathbf{x}_r if it is perpendicular to $\tilde{\mathbf{w}}_{\sigma_u, \sigma_v}(\mathbf{x}_r)$.

Finally, it is also worth to comment that, in practice, the pitch angle θ is actually unknown since this parameter varies when the dampers actuate, which happens because of a pothole on the road, a sudden slope change, or with acceleration and braking. Hence, the angle obtained from a calibration of the camera would not be valid for every image. This means that if we do not have an additional sensor able to measure θ on-line then we must estimate it from the own image. A possibility consists of testing different hypothesis for θ when doing the RANSAC based fitting. Another possibility is to use some estimation method to chose the optimum value for θ according to the goodness of the obtained fittings. Currently, we have adopted the former option, while the latter is under research.

IV. RESULTS

We have tested the proposed method in sequences acquired by different CMOS based cameras. In all cases the images are of 640×480 pixels, but we sub-sample them to the half in order to speed up the process. Experimentally we have checked that this sub-sampling does not affect the results.

Figure 4 shows the ridgeness response in different adverse circumstances: deteriorated lane markings, presence of white vehicles, shadows, high curvature curves, and night. We think that these results are sufficiently good as to use ridgeness as evidence of the presence of lane markings. In fact, in order to speed up the processing and thanks to the well-behaved dynamic range of this measure, we are considering

as ridgels just those pixels \mathbf{x} with $\tilde{\kappa}_{\sigma_d, \sigma_i}(\mathbf{x}) > 0.25$ (we remind that the maximum value for $\tilde{\kappa}_{\sigma_d, \sigma_i}(\mathbf{x})$ is 2 for 2D images). Also, to reduce clutter and speed up the process, as general rule we discard a ridgel \mathbf{x}_r if $\tilde{\mathbf{w}}_{\sigma_d, \sigma_i}(\mathbf{x}_r)$ is in $[3\pi/4, 5\pi/4]$, which happens for lane markings with a large horizontal component. Of course, this could be a problem for curves with a very high curvature. However, we have checked experimentally that this criterium performs well for curves that can be taken at a speed of more than 40 km/h.

We remark that these rules for selecting relevant ridgels have been used in all our sequences with the same parameter settings, even though the camera, optics and vehicles were not the same for all the sequences. From sequence to sequence we tried to have a similar vanishing point location looking at a straight flat road, but the adjustment was done manually without trying to be too precise.

For fitting the hyperbolas to the ridgels we also set the same parameter values for all sequences. The more relevant are $I_{max} = 1000$ for the number of trials of the RANSAC, and $c_t = 4$ for the consensus threshold (which is not a too low value provided the Sampson's distance and orientation criteria have been set more demanding). Moreover, we explore the goodness of the fittings under 3 pitch angles (θ): the expected one (fixed), one below and another above. Besides, we also demand to the method a minimum number of trials, in particular, 25. This means that the RANSAC will test the goodness of at least 25 models and a maximum of 1000.

Fitting results under different circumstances are shown in Fig. 5: presence of continuous and discontinuous lane markings, cars almost occluding all the lane markings, going out and entering tunnels, disturbances from non-lane markings marks and roadworks, presence of shadows, and night images of low contrast and with large reflections. Figure 6 shows several representative frames of a sequence where the ego-vehicle runs on a curved road that bifurcates in two.

It is worth to notice that in our experiments we found very relevant the use of the $\tilde{\mathbf{w}}_{\sigma_d, \sigma_i}(\mathbf{x})$ vector field in the RANSAC based fitting.

Finally, we point out that the proposed method runs in about 40 ms/frame in a 2.0 GHz Pentium IV.

V. CONCLUSIONS

We have introduced ridgeness as low-level image descriptor for detecting lane markings using a forward-facing camera. We have illustrated how this measure is able to output high values indicating the presence of lane markings even in challenging situations. Besides, the ridgeness computation process also outputs a vector field perpendicular to the dominant orientation of the image at each pixel.

In addition, we have also proposed as novelty the adaptation of RANSAC, a general robust estimation method, to simultaneously fit in the image plane the couple of hyperbolas which approximately correspond to the projection of the actual lane markings. The main points of the adaptation consists of using the ridgeness measure as input data for fitting and then Sampson's distance as well as the above mentioned associated vector field to asses the goodness

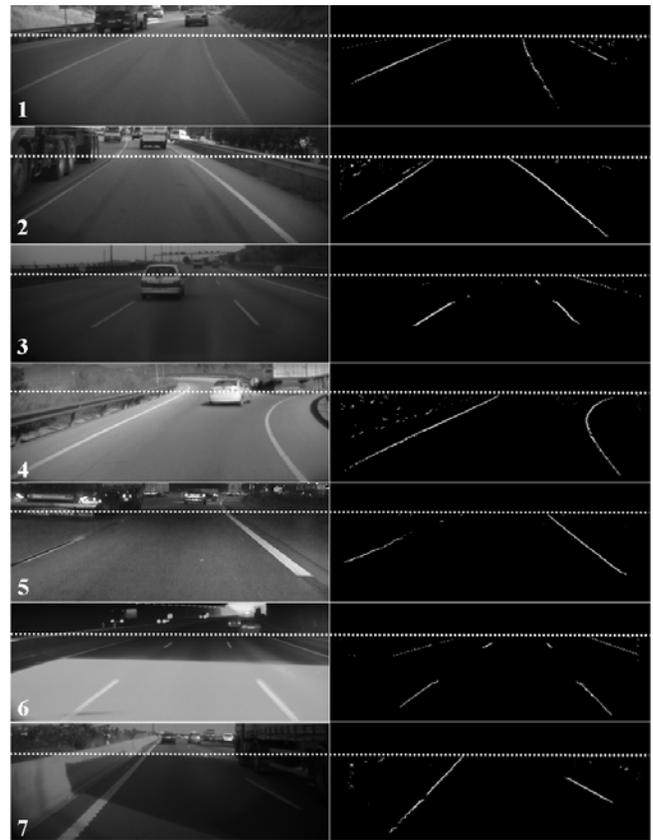


Fig. 4. Right column: ridgeness of lane markings. The dotted line is the fixed initial row below of which ridgeness is computed (v_{min} in Fig. 3). In the original images we see deteriorated lane markings: (1) the right one; (2) the left one (partially black because of wheel residues after some hard brake). (3) White vehicles in the scene. (4) High curvature right curve. (5) Night image with a dirty left discontinuous line. Shadows: (6) entering a tunnel; (7) the left line is partially in shadow and the right is completely.

of the fitting. While Sampson's distance can probably be replaced by some geometric distance, taking into account the dominant orientation of the image appears as very relevant in our experiments. Therefore, it becomes more critical the fact that this orientation is more robust if it is based on the vector field associated to the ridgeness measure than if it is based on the image gradient vector field, which has a more local nature, *i.e.* the idea of *dominant* orientation does not appear.

Another point to remark is the fact that to fit the two hyperbolas only 4 parameters must be estimated (assuming a fixed pitch angle) which together with the simplicity of the goodness assessment makes possible the use of RANSAC in the required real-time.

It is also worth to notice that, like in other works, our approach can output lane curvature as well as vehicle position and orientation inside the lane in real-time. Of course, for that we need first to calibrate the camera.

Examples illustrating that the proposed method is a promising framework for lane marking detection have been presented too.

Note that this paper has focused the lane marking detection



Fig. 5. Curves on images acquired using different CMOS cameras and vehicles. In all cases, since the horizon is similarly located, we have used the same parameters. We can see: (1,2,3) Curves in presence of continuous and discontinuous lane markings. (4) With cars almost occluding all the lane markings at the left of our lane. (5,6) Going out and entering tunnels. (7,8) Disturbances from non-lane markings marks and roadworks, resp. (9,10) Special discontinuous lane markings. (11,12) Presence of shadows. (13,14) Night images of low contrast and with large reflections, resp.

in a single frame. Our next step is to add temporal coherence by tracking either the fitted model parameters or the physical ones along time (f.i. using a Kalman-like filter). Using tracking techniques, the system shall be more robust under spurious errors, and image processing will speed-up, since the parameters of the lane markings detected in a frame can be used to process less image pixels in the following frame. Another point under research is the estimation of the best pitch angle in a single-frame basis, for its posterior incorporation to the tracking.

REFERENCES

[1] M. Bertozzi, A. Broggi, A. Fascioli. *Vision-based Intelligent Vehicles: State of the Art and Perspectives*, Robotics and Autonomous Systems, vol. 32, pg. 1–16, 2000.



Fig. 6. Several images of a sequence taken at about 10 fps. After going out of a highway ($f_1, f_{23}, f_{33}, f_{44}$), the ego-vehicle finds a bifurcation (f_{68}, f_{75}, f_{91}) in two different roads taking the right one (f_{115}). Notice the high curvature of the lane markings and how the method is able to detect them. At frames f_{68} and f_{75} the hyperbola branch at the right indicates that the continuation of our road is to the left, and this implies that, due to the bifurcation to the right, the farther away part of the hyperbola branches at the right do not fit actual lane markings. However, when we decide to go to the right, then the fitting is correct again (f_{91}, f_{115}).

[2] A.M. López, F. Lumberras, J. Serrat, J.J. Villanueva. *Evaluation of Methods for Ridge and Valley Detection*, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 21, pg. 327–335, 1999.

[3] A.M. López, D. Lloret, J. Serrat, J.J. Villanueva. *Multilocal Creaseness Based on the Level-Set Extrinsic Curvature*, Computer Vision and Image Understanding, vol. 77, pg. 111–144, 2000.

[4] E.D. Dickmanns and B.D. Mysliwetz. *Recursive 3D road and relative ego-state recognition*, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 14, pg. 199–213, 1992.

[5] A. Guiducci. *Parametric Model of the Perspective Projection of a Road with Application to Lane Keeping and 3D Road Reconstruction*, Computer Vision and Image Understanding, vol. 73, pg. 414–427, 1999.

[6] R. Aufrère and R. Chapuis and F. Chausse. *A Model-Driven Approach for Real-Time Road Recognition*, Machine Vision and Applications, vol. 13, pg. 95–107, 2001.

[7] M.A. Fischler and R.C. Bolles. *Random Sample Consensus: a Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography*, Commun. ACM, vol. 24, pg. 381–395, 1981.

[8] P.D. Sampson. *Fitting conic sections to very scattered data: an iterative refinement of the Bookstein algorithm*, Computer Graphics and Image Processing, vol. 18, pg. 97–108, 1982.