Computational model of driver gaze direction

Peyman Moghadam*, Janusz A. Starzyk^ and Wijerupage Sardha Wijesoma*

*Department of Electrical & Electronic Engineering,
Nanyang Technological University,
50 Nanyang Avenue, Singapore 639798

peym.mo@gmail.com, eswwijesoma@ntu.edu.sg Phone: +65-91160043

^ School of Electrical Engineering and Computer Science
Russ College of Engineering and Technology
Ohio University, U.S.A.
starzyk@bobcat.ent.ohiou.edu Phone: (740)593-1580, Fax: (740)593-0007

Content types: Computational

Abstract:

A wide variety of computational models for human visual attention have been presented during the past decades, although they focus mostly on single tasks modeling of visual attention such as object recognition in cortex. However, when it comes to more complex scenarios related to active vision such as the visual behavior of drivers, none of these models provides a satisfactory solution. Driving is primarily a visual task strongly link to motor control, yet the visual behavior of drivers is still poorly understood. Over the decades, studies on driver eye fixations try to understand where and how long drivers look to obtain visual information either in natural scenes or in driving simulators [1-4].

They are many theories about what features a driver might use while driving [5]. Many algorithms such as road following methods [6] have been developed to find the lane markings or the road edges as the main cues of the road while studies on the visual scanning pattern of experienced drivers revealed that only 1% of all fixations in the open-road driving scenarios fixated on the lane markings or road edges [1]. On the other hand, several road detection algorithms [5] look for distinct appearance characteristics from the road regions to segment the scene to road and non-road sections though Rockwell et al. [7] reported that at distance 75 to 250 feet in front of the car drivers look at the road less than 21%.

According to the results of recent researches, drivers prefer to look close to the end of the road (64% of all fixations [8]) where the road edges converge on the horizon (vanishing point) and drivers can obtain information with maximal lead-time [2-3]. Others also have indicated that the gaze distribution tends to be increasingly constrained to the vanishing point with an increase in driving speed [4, 9]. Hence, it is clear that vanishing point plays an important role as a global constraint for detecting roads for human drivers and if any computational model wants to mimic the driver's behaviors it should start with estimating the vanishing point locations.

Our goal in this work is to propose a computational model for the visual gaze behavior of drivers by starting with estimating the vanishing point location in image space based on the biologically inspired model of the joint activity of simple cells and complex cells populations in the primary visual cortex.

The receptive fields of orientation-selective simple cells in the primary visual cortex (V1) are modeled by two-dimensional Gabor filters. Next, the activities of a pair of simple cells with a phase difference of $\frac{\pi}{2}$ are used to provide a model of cortical complex cell receptive fields, which are phase insensitive.

Since all parallel border road lines, road edges, and even ruts and tire tracks left by previous vehicles on the road can potentially point to a vanishing point in the image space, one should first estimate the dominant orientation by centering a family of complex cells tuned to a full range of orientations at each possible position in the image space. The maximum response is selected as the dominant orientation at each position. Then, the vanishing point can be determined by making the pixels vote for the location of the candidate vanishing points in the direction of their dominant orientations. To achieve a precise angular resolution, filters with a large number of orientations are generally used ($n = 72$ in [10] and $n = 36$ in [11]). However, psychophysical evidence [12] suggests that information in the cortex is more widely represented by the joint activity of neuronal populations rather than by a single neuron activity. Therefore, we present a biologically inspired model, meant to mimic the orientation-selective properties of complex cells in the cortex, based on the joint activity of only 4 main oriented filters $\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ at each possible location in the image space.

The summary of the proposed biologically inspired model is as follow. To estimate dominant orientation at each pixel location, a gray-value input image is convolved with four oriented Gabor filters, $\theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$. Next, these Gabor responses are reduced to two dominant filter activation strengths with an opponent offset. The first two maximum Gabor responses are reduced by their orthogonal direction activations and then these two resultant filter activation strengths are used to represent the direction/strength of the dominant orientation of that pixel location.

In order to assess the accuracy of the proposed method a set of synthetic images with associated ground truth are used. The angular estimated error is symmetric and it is bounded. The proposed technique has a finite angular resolution of 1.4(± 0.75) degrees with only 4 oriented Gabor filters compared to the 2.5 degrees angular resolution with 36 oriented Gabor filters in [11] and 1.25 degrees with 72 oriented Gabor filters in [10]. Even a higher resolution can be obtained in two-stage process where the initial estimate is used to refine fitting to most likely directions, and then using Gabor filter activations to obtain a better estimate.

After the dominant orientation is estimated, in order to estimate the vanishing point location a ray is drawn along each dominant orientation and the pixel with the maximum number of supporting rays is declared the candidate vanishing point. However, in such method all the dominant orientations would have the same influence on the voting scheme. To address this problem we considered assigning weights to each dominate orientation based on different weighting methods. Later, a series of quantitative and qualitative analyses are presented using real data sets from the DARPA Grand Challenge projects to demonstrate the effectiveness of the proposed methodology.

Reference:

[1] Mourant, R.R., & Rockwell, T.H. (1970). Mapping eye-movement pattern to the visual scene in driving: An exploratory study. Human Factors, 12, 81-87.

[2] Land, M. F., & Lee, D. N. (1994). Where we look when we steer. Nature, 369, 742–744.

[3] Serafin, C. (1993). Preliminary examination of driver eye fixations on rural roads: Insight into safe driving behavior (Rep. No. UMTRI 85843). Ann Arbor: University of Michigan, Transport Research Institute.

[4] Land, M. F., Horwood, J., & Cornwell, T. S. (1994). Fast driving reduces eye movement amplitude and frequency. Investigative Ophthalmology & Visual Science, 35, 2033.

[5] Alon Y., Ferencz A., and Shashua A. (2006). Off-road path following using region classification and geometric projection constraints, in IEEE Computer Society Conference on Computer Vision and Pattern Recognition.

[6] Bertozzi M. and Broggi A. (1998). GOLD: a parallel real-time stereo vision system for generic obstacle and lane detection, IEEE Transactions on Image Processing, 7(1): p. 62-81.

[7] Rockwell, T.H., Emst, R.L., and Rulon, M.J. (1970). Visual information acquisition by the night driver. In Visual Requirement in Night Driving (pp. 23-31). Washington, D.C. National Research Board.

[8] Rackoff, N.J., and Rockwell, T.H. (1975). Driver search and scan patterns in night driving. Transportation Research Board Special Report 156,53-63.

[9] Rogers, S. D., E. E. Kadar, and A. Costall (2005). Gaze Patterns n the Visual Control of Straight-Road Driving and Braking as a Function of Speed and Expertise, Ecological Psychology 17(1):19-38

[10] Rasmussen C. (2004). Grouping dominant orientations for ill-structured road following, in Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition.

[11] Kong H., Audibert J., and Ponce J. (2010). General Road Detection from a Single Image, Image Processing, IEEE Transactions on, PP(99): p. 1-1.

[12] Singh, G., Memoli, F., Ishkhanov, T., Sapiro, G., Carlsson, G., & Ringach, D. L. (2008). Topological analysis of population activity in visual cortex. Journal of Vision, 8(8), 11–18.