ROAD DETECTION IN AN OUTDOOR ENVIRONMENT

Ondřej Mikšík
Master Degree Programme (1), FEEC BUT
E-mail: xmiksi01@stud.feec.vutbr.cz

Supervised by: Luděk Žalud
E-mail: zalud@feec.vutbr.cz

Abstract: This paper deals with road detection in an outdoor environment. By comparison with preceding approaches, which use various sensors, our system uses only a monocular camera. In this paper, we propose a novel approach - a fusion of frequency based vanishing point estimation and probabilistically based color segmentation.

Keywords: Road detection, Vanishing point estimation, Mixture of Gaussians, Gabor wavelets

1 INTRODUCTION

The mobile robotics community made an enormous effort to build robots with elements of autonomous behavior during the past two decades. The field of possible objectives is various from maze solving robots, letter–carrier robots to fully autonomous vehicles which can operate in an unknown environment. Many successful projects has proven in the past that the idea of fully autonomous vehicle is not utopia (e.g. DARPA Grand Challenges).

Many papers about vision-based road segmentation have been published - such approaches usually employ many various and expensive sensors, including Radar, Lidar or stereo vision. There exist many situations, when it is inconvenient or even impossible to use additional sensors instead of a monocular camera (robots are usually already equipped).

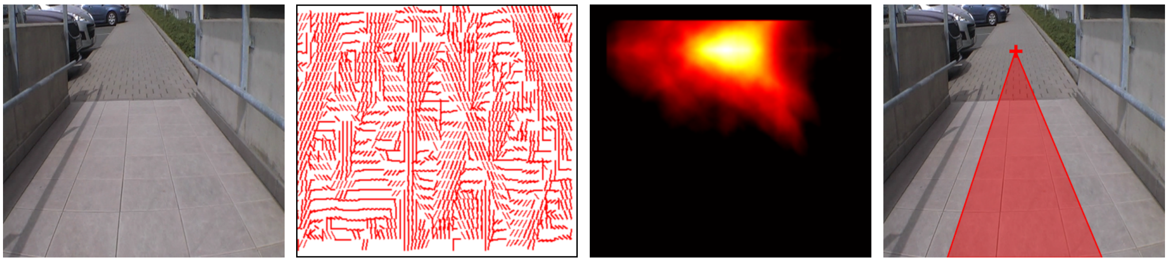
2 SYSTEM DESIGN

Let us define some important features and demands of the system: the system has to reliably find the way in diverse light conditions (shadows, overexposed highlights, . . .), work reliably on both, structured and ill-structured roads (sand, concrete, tarmac, . . .) and use a minimum number of sensors.

We fulfill these demands by a fusion of the frequency based estimation of so called **vanishing point** and probabilistically based **texture segmentation**. A combination of two different approaches, allows us to solve difficult situations without any a priori knowledge of robot's environment. The basic idea of our solution is estimation of the vanishing point, which determines the training area for texture segmentation. Next, road color models are constructed from sample pixels defined by the training area. These models are associated with previously learned models, which are stored in a memory. Further, learned models are adaptively updated. Therefore, the models include both the road colors' history and the current road appearance.

3 VANISHING POINT ESTIMATION

Parallel lines in the real world, do not look like parallel lines under the perspective projection. Therefore, borders of each straight road in an image plane, converge at some point, the so called **vanishing point**. The first step of a vanishing point estimation algorithm is, the estimation of the dominant



(a) **(b)** **(c)** **(d)**

Figure 1: Input image (a), estimated dominant orientations (b), voting function (c), output (d).

orientation $\theta(\mathbf{p})$ of an image at pixel $\mathbf{p}(x,y)$. Our approach is based on a bank of 2D Gabor wavelet filters since they are known to be accurate [2].

The set of $k \times k$ Gabor kernels for an orientation θ, wavelength λ and odd or even phase, the filters are defined by

$$\widehat{g}_{odd}(x,y,\theta,\lambda) = \exp\left(-\frac{1}{8\sigma^2}(4a^2+b^2)\right)\sin\left(\frac{2\pi a}{\lambda}\right), \tag{1}$$

where $x = y = 0$ is the kernel center, $\sigma = \frac{k}{9}$, size of kernel k is determined by wavelength as $k = \frac{10\lambda}{\pi}$ and $\lambda = 2^{\log_2(I_w)-5}$ provides a good trade-off between computational complexity and a precision. To obtain even kernel, “sin” is simply substituted by “cos”. Next, a and b are defined as

$$\begin{aligned} a &= x\cos(\theta) + y\sin(\theta), \\ b &= -x\sin(\theta) + y\cos(\theta). \end{aligned} \tag{2}$$

Then, $\widehat{g}$ ’s DC component is subtracted from Gabor kernel and kernel’s coefficients are normalized, so that $L^2 = 1$.

To get the best characteristics of a local texture jet, a complex response of the convolution of an image I with each of n evenly spaced Gabor filter orientations is computed. For n even and odd pairs of Gabor filters (e.g. $n = 36$), the dominant orientation at pixel $\mathbf{p}(x,y)$ is chosen as the filter orientation which elicits the maximum complex response at that location.

The set of possible vanishing points for each pixel $\theta_{max}(\mathbf{p})$ with dominant orientation θ_{max} are all pixels along the line defined by ($\mathbf{p}$, $\mathbf{p}(\theta_{max})$). Let the angle of the line joining an image pixel $\mathbf{p}$ and a vanishing point candidate $\mathbf{v}$ is $\alpha(\mathbf{p},\mathbf{v})$, then $\mathbf{p}$ votes for $\mathbf{v}$ if the difference between $\alpha(\mathbf{p},\mathbf{v})$ and $\theta_{max}(\mathbf{p})$ is within the dominant orientation estimator’s angular resolution (coefficient $\gamma = 2$ sets selectivity).

$$vote(\mathbf{p},\mathbf{v}) = \begin{cases} 1 & \text{if} |\alpha(\mathbf{p},\mathbf{v}) - \theta_{max}(\mathbf{p})| \leq \frac{\gamma\pi}{n}, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

Next, the definition of an objective function for each vanishing point candidate $\mathbf{v}$ is straightforward

$$votes(\mathbf{v}) = \sum_{\mathbf{p} \in R(\mathbf{v})} vote(\mathbf{p},\mathbf{v}), \tag{4}$$

where $R(\mathbf{v})$ is a voting region, which includes all image pixels below the horizontal line $\mathbf{l}$ determined by the current vanishing point candidate $\mathbf{v}$.

Instead of usage of output independently per each frame, we rather run a smoothing filter (CONDENSATION) throughout the whole sequence to reduce influence of noise and to avoid the jumpy characteristic of output. Particle filters (sequential Monte Carlo) are often used in computer vision since they overcome many limiting assumptions of Kalman filters.

4 ROAD DETECTION

Vanishing points provides information about direction, however we do not have any information about a road surface and free space ahead of the robot. Thus, another algorithm based on adaptive color segmentation is needed - Gaussian Mixture Models (GMM) [1]. We use a hierarchical agglomerative k-means clustering (HAC) to construct GMM models from training area defined by a vanishing point. Each cluster c is represented by its mean μ, covariance matrix Σ and a mass m. In addition to c training models, n_l learned models exist, which represent "history of the road" with exponential forgetting. At the beginning, all color models are null. Each training model is compared with learned models

$$(\mu_L - \mu_T)^{\mathrm{T}}(\Sigma_L + \Sigma_T)^{-1}(\mu_L - \mu_T) \leq d_{similar}, \tag{5}$$

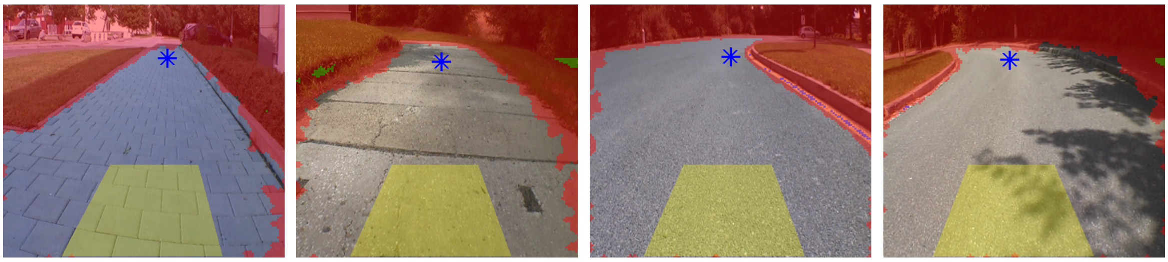
where μ is a mean vector, and Σ is a covariance matrix. If the training model overlaps any learned model, the learned model is updated according to formulas

$$\mu_{updated} = \frac{m_L\mu_L + m_T\mu_T}{m_L + m_T}, \quad \Sigma_{updated} = \frac{m_L\Sigma_L + m_T\Sigma_T}{m_L + m_T}, \quad m_{updated} = m_L + m_T, \tag{6}$$

where m is associated mass to the model. Otherwise, there are two possibilities. If all models are not full, then the new model is created. If all models are full, then the model with the lowest mass is discarded and a new one is created in its place. Next, all pixels of the image are assigned a "roadness" score, which is measured as a minimum of the Mahalanobis distance between each pixel and learned models to measure a degree of belonging to the road/non-road region of pixels outside the training area

$$D(\mathbf{p}, \mu_i) = \min_i((\mathbf{p} - \mu_i)^{\mathrm{T}})\Sigma_i^{-1}(\mathbf{p} - \mu_i)) \tag{7}$$

These values can be used as a input to some higher AI or to identify patches that create the road.



(a) **(b)** **(c)** **(d)**

Figure 2: Output of proposed algorithm.

5 CONCLUSION

We have presented a novel method for road detection in an outdoor environment. Our method neither need any additional sensor, nor previously learned models. Proposed algorithm works well on both, structured and unstructured roads with various types of surfaces and dynamically changing light conditions. In fact, whole algorithm is much more complicated. Due to the lack of space, only basic parts are described. Extended paper will be presented at IEEE ICRA 2011.

REFERENCES

[1] Dahlkamp, H.; Kaehler, A.; Stavens, D.; Thrun S.; Bradski G: Self-supervised Monocular Road Detection in Desert Terrain. In *Robotics: Science and Systems*, 2006.

[2] Rasmussen, C.: Grouping Dominant Orientations for Ill-Structured Road Following. In *IEEE International Conference on Computer Vision and Pattern Recognition*, 2004.