Vision based Horizon Detection for UAV Navigation

Stavros Timotheatos[1,2], Stylianos Piperakis[1,2], Antonis Argyros[1,2]
and Panos Trahanias[1,2]

[1] Institute of Computer Science, Foundation for Research and Technology
-Hellas (FORTH), Heraklion, Crete, Greece,
(stimotheat, spiperakis, argyros, trahania)@ics.forth.gr
[2] Department of Computer Science, University of Crete, Heraklion, Crete, Greece

Abstract. In this paper, we present a novel framework for horizon line (HL) detection that can be effectively used for Unmanned Air Vehicle (UAV) navigation. Our scheme is based on a Canny edge and a Hough detector along with an optimization step performed by a Particle Swarm Optimization (PSO) algorithm. The PSO's objective function is based on a variation of the Bag of Words (BOW) method to effectively consider multiple image descriptors and facilitate efficient computation times. More specifically, the image descriptors employed are $L*a*b$ color features, texture features, and SIFT features. We demonstrate the effectiveness and robustness of the proposed novel horizon line detector in multiple image sets captured under real world conditions. First, we experimentally compare the proposed scheme with the Hough HL detector and a deep learning HL estimator, a prominent example of line detection, and demonstrate a significant boost in accuracy. Furthermore, since from the horizon line the UAV roll and pitch angles can be derived, this scheme can be used for UAV navigation. To this end, to further validate our approach, we compare the horizon computed roll and pitch angles to the IMU ones obtained with a complementary filter.

Keywords: Horizon line detector, Unmanned Air Vehicle, Particle Swarm Optimization, Bag of Words, Pitch and Roll angles derivation.

1 INTRODUCTION

To operate an UAV in a realistic environment, it is crucial to monitor and control parameters and environment features associated with the system state. Information provided by vision-based sensors is of utmost importance for this task, as evidenced in a number of studies [1], [2]. More specifically, of particular interest is the detection of the so-called Horizon Line (HL), that in turn gives rise to two critical parameters regarding UAV's angular orientation, namely roll angle $\phi_{\mathbf{roll}}$ and pitch angle $\theta_{\mathbf{pitch}}$. The latter are directly involved in tasks such as autonomous navigation [3], [4], [5], [6].

HL detection algorithms usually rely on plausible assumptions, such as that HL forms a linear boundary, lies in the upper half of the image, sky and non-sky regions are equally probable, and the higher brightness and prominent texture of the upper part of the image (sky region) compared to the lower part (non-sky region). Methodologically, machine learning approaches and particularly deep learning [7], have led to state-of-the-art performance, at the cost of assigning significant computational resources during

the training phase. Moreover, such methods do not scale smoothly in cases of images that contain unlearned features and/or patterns. Workman et al. [8], employed Deep Convolutional Neural Networks (CNN) to compute HL. Alternative probabilistic approaches that rely on real-time feature inferencing offer robustness in the presence of irregular features, but significantly increase execution times. In [9], Oreifej et al. detect HL based on maximum a posteriori estimation applied to multiple candidate horizon lines. Moreover, they compute ϕ_{roll} and θ_{pitch} based on the detected HL.

Regardless of the employed methodology, most HL detection approaches fall into two main categories: (a) edge detection approaches and (b) segmentation into sky and non-sky regions. The authors in [10], were among the first to formulate a robust edge detection based scheme to identify HL. By employing the Hough transform, prominent lines are extracted and the *strongest* one is declared as HL. Lie et al. [2], compute an image edge map which is subsequently transformed to a multi-stage graph. Dynamic Programming is then utilized to detect HL as the shortest path in the graph.

A number of approaches in the second category focus on extracting the boundary between sky and non-sky image areas, based on various descriptors such as texture, color, SIFT, and Histogram of Gradients. Ahamd et al. [11] proposed an edge-less machine learning approach based on dynamic programming for classifying each image pixel based on the above descriptors.

In this work, we propose a novel HL detector that employs Particle Swarm Optimization (PSO) [12], to assess candidate horizon lines. The latter are initialized via Canny edge detection followed by the application of the Hough transform (HT) [13]. The main novelty of the current work lies in the optimization step where adequate image descriptors are utilized to evaluate alternative HL solutions. This is accomplished via the PSO's objective function which is formulated on top of a variant of the Bag of Words (BOW) encoding of the image content. The image descriptors used in the latter are color, texture, and SIFT features. Accordingly, rich image-based information is considered, whereas the PSO scheme facilitates efficient computation times.

The proposed HL detector is evaluated with multiple image sets captured under real world conditions. In the conducted experiments, we compare the proposed scheme with the Hough HL and CNN HL detectors and demonstrate a significant boost in accuracy. Furthermore, since from the HL the UAV roll and pitch angles can be derived, this scheme can be used for UAV navigation. To this end, to further validate our approach, we contrast the horizon computed roll and pitch angles to the IMU ones.

The paper is organized as follows. Section 2 establishes the proposed framework for HL detection. Experimental and comparative results are presented in Section 3 and the paper concludes with a brief discussion and future work suggestions in Section 4.

2 Novel Horizon Detection Framework

The proposed HL detection framework consists of two main modules: (a) computation of candidate HLs and (b) HL derivation. The first module comprises edge detection followed by HT, whereas the latter employs PSO to derive HL. A detailed technical presentation is in order.

2.1 Computation of Candidate HLs

A block diagram illustrating the processing steps in the current module is depicted in Figure 1. Initially, the input image is converted to grayscale and, subsequently, a Canny edge detector is applied. Edges are identified as local maxima of the intensity gradient, where in our case the corresponding gradient is calculated using the derivative of a Gaussian filter. The computed edge map image is used as input to a line detector based

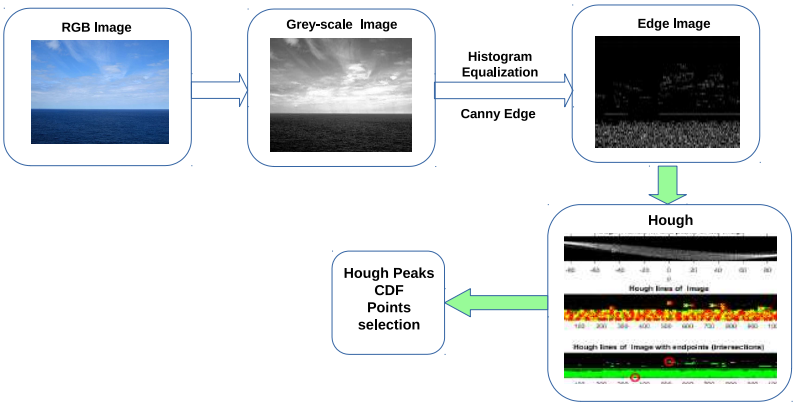

Fig. 1. Computation of candidate HLs.

on the HT, where the image space (x,y) is transformed to the (ρ,θ) parametric space. Consequently, a line in the image space corresponds to a point in Hough space and vice versa. Next, Hough accumulator cells are populated with one vote each time a non-background point in the image is detected. The bin with the highest number of votes represents a Hough peak, and a potential line in the input image.

Accordingly, we formulate the set of candidate lines, $\mathbb{S}_{HL}$, as prominent points in the Hough parametric space. More specifically, two subsets of points are constructed, $\mathbb{S}_{SP}$ and $\mathbb{S}_{RP}$, with their union representing $\mathbb{S}_{HL}$. $\mathbb{S}_{SP}$ denotes the most probable Hough peak points P_H, according to their cumulative distribution (CDF) and $\mathbb{S}_R$ consists of randomly selected points from the remaining Hough space $H - \{P_H\}$.

The rationale for formulating candidate HLs as the union of the above two subsets $\mathbb{S}_{SP}$ and $\mathbb{S}_R$ lies in the nature of the adopted optimization method, namely PSO. In particular, $\mathbb{S}_{SP}$ feeds as candidate HL solutions to the PSO lines with high image contrast and hence high probability of being the sought HL. Given that in many cases the actual HL might be a weak image line, $\mathbb{S}_R$ is used to provide randomly selected Hough points to PSO and thus facilitate successful convergence to the correct solution, so that the swarm can effectively explore the parameter space, without getting stuck to local minima.

2.2 HL Derivation

A schematic representation of the second module is shown in Figure 2a which actually comprises PSO. The latter is a stochastic evolutionary algorithm in generations, that

features particle populations and a set of rules for evolution. The population is a set of points in the parameter space of the objective function to be optimized. The particles evolve in generations that lead in a solution that optimizes a defined objective function. Therefore, PSO particles, which correspond to a set of proposed solutions, move around a specific solution-space, searching and evolving to converge to the solution that optimizes the objective function.

Accordingly, the HL returned by PSO is optimized with respect to the image-based criteria that are expressed by the objective function. All particles are initialized from $\mathbb{S}_{HL}$ and the PSO variant developed by Oikonomidis et al. [14] is utilized, which is based on the canonical PSO. The main assets of the PSO are: (a) a low number of objective function evaluations are used for convergence, (b) the objective function derivatives are not required, and (c) depends on only a few parameters.

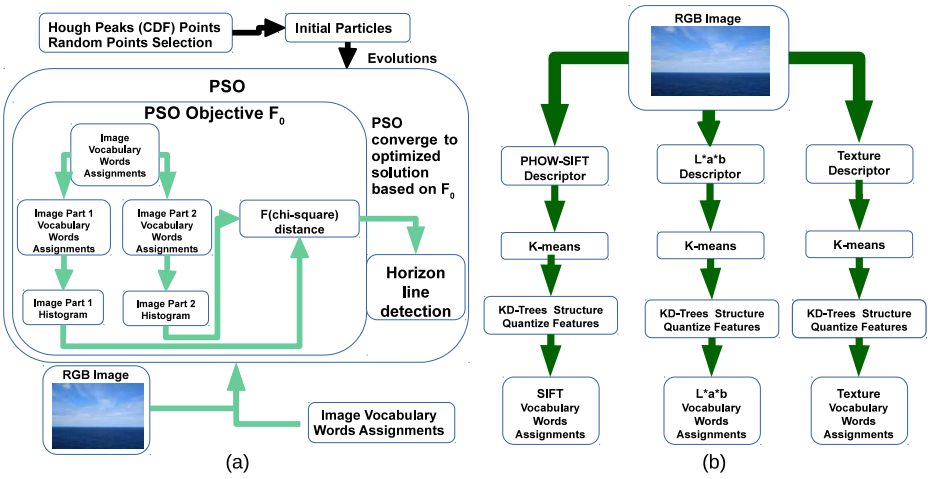

Fig. 2. (a) HL derivation; (b) BOW formulation.

Given that each PSO particle represents an image line, it segments the image into two regions. For each of the two regions, three feature-histograms are formed based on SIFT, color, and texture features. To facilitate computational efficiency, feature-histograms are not computed for each possible image region but are readily derived from corresponding histograms precomputed for the entire image. The latter are computed only once and give rise to histograms for specific image regions by means of the Bag of Words (BOW) scheme, outlined in the next subsection for the sake of clarity of presentation. The above computed histograms for each image region are used to formulate the objective function F_O of the PSO. For that, the corresponding histograms across the two regions are compared in terms of the χ-square distance (χ^2). F_O is calculated as the product of the three χ^2 distances:

$$F_O = \prod_{i=1}^{3}(1-\chi_n^2) \tag{1}$$

where the three χ_n^2 distances refer to each descriptor used (SIFT, color, and texture). F_O is calculated on each PSO iteration for every particle and represents a score. Essentially,

each particle (ρ, θ) corresponds to a line that separates the image into two regions. χ_n^2 measures the distance between the BOW representations of the two regions. Thus, the $1-\chi_n^2$ measures the similarity of these regions. Given the above, the objective function F_O measures the similarity of the two regions with respect to all three descriptors. Overall, the PSO-based minimization of the F_0 function results in an optimized Hough point $(\rho^\star, \theta^\star)$ that separates the image into two maximally dissimilar regions.

2.3 Bag of Words-BOW

BOW is a method for image content representation that has been used extensively for image recognition and classification. In our framework BOW is employed for obtaining feature-histograms for any image region from the corresponding histogram of the entire image and is additionally used for computational efficiency. As mentioned above, three such histograms are used in our framework based on appropriate local descriptors, namely SIFT features, $L*a*b$ color features, and texture features. The latter are extracted via the Leung-Malik [15] filter bank.

Subsequently, the vocabulary of visual words for each different image feature descriptor is constructed, a process often mentioned as BOW encoding. In our case, the visual vocabulary is computed from a number of words (K-means centers) by running the K-means algorithm on features. Finally, BOW quantizes features using the vocabulary for finding the nearest K-means centers, based on the Euclidean distance metric for each descriptor, thus forming the visual words. The described process for BOW formulation is presented in Figure 2b.

2.4 Derivation of Roll and Pitch Angles

Having computed HL, the UAV's roll and pitch angles are readily available from a sequence of images acquired from the UAV camera as in [9]. The roll angle ϕ_{roll} can be directly determined from the HL since it corresponds to the angle formed by the slope of the HL. Since it is orthogonal to the camera's rotation axis, the HL rotates the same as the roll angle of the UAV camera and is invariant to all other motions, such as UAV translation, pitch and yaw rotations. This is true because under these motions no image motion is induced for the points at infinity and the HL slope remains constant [1]. Accordingly, in this work we first compute the horizon ϕ_{roll} angle. Similarly, pitch angle θ_{pitch} computation is based on the assumption that HL is captured from the UAV camera in infinite distance. Hence, θ_{pitch} is calculated using the current and the initial HL positions as in [9], where f is the camera focal length:

$$\phi_{roll} = \tan^{-1}(-\frac{cos(\theta^\star)}{sin(\theta^\star)}),\ \ \theta_{pitch} = \tan^{-1}\left((\frac{\rho_i^\star}{sin\theta_i^\star} - \frac{\rho_0^\star}{sin\theta_0^\star})/f\right) \tag{2}$$

3 Results

In this section both qualitative and quantitative experimental results are presented in order to demonstrate the effectiveness of the proposed HL detection framework. A dataset of 300 RGB images has been employed for conducting experimentation with the proposed framework. Moreover, an implementation on an actual UAV facilitated further evaluation under real world conditions.

3.1 Qualitative Results

A dataset consisting of 300 RGB images has been compiled to facilitate in-depth experimentation. It contains images from three different sources: (a) a set of images from the recently published Horizon Lines in the Wild dataset [8] and (b) a set of images captured with a mobile camera mounted on a moving UAV, termed as VHS. The content and quality of the images in the dataset are analogous to images commonly acquired by UAV-cameras during flight. Accordingly, brightness, blur, contrast, resolution, and horizon orientation vary considerably across the images as it is the usual case in UAV captured data.

In the following, we present HL detection results on sample images from the above dataset. For comparison purposes the results obtained via the application of the proposed framework, termed $PSO-HL$, are contrasted to the ones obtained with the $Hough-HL$ [10]. Figure 3a presents indicative results for various environment scenes. In all cases, we used a value $\sigma = 0.1$ for the Gaussian filter in the Canny operator, a step size of $\rho = 0.1$ and $\theta = 0.1$ for the HT parameters, and the initial PSO particles and number of evolutions were both set to 50.

Figure 3b presents results on images acquired VHS set. The presented images depict HL detection in difficult landscape scenes. Mountains, sky, sea, and trees along with blurring effects, low brightness, and different orientations formulate challenging images for HL detectors. Still, in all cases the proposed $PSO-HL$ succeeded in accurately detecting HL and considerably outperformed $Hough-HL$.

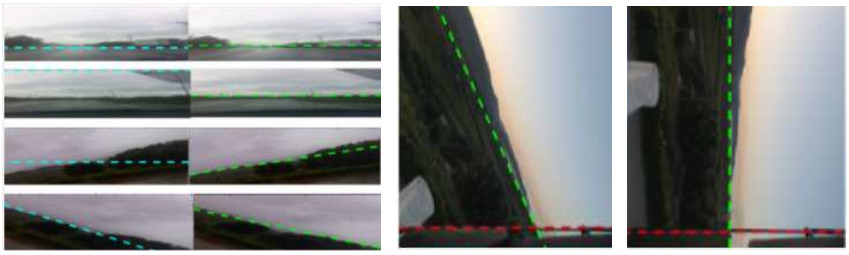

Fig. 3. HL detection in challenging images. (a) Moving camera, left column: $Hough-HL$ (cyan); right column: $PSO-HL$ (green). (b) Left column: Wide angle; right column: 90 deg orientation, $Hough-HL$ (red), and $PSO-HL$ (green).

3.2 Quantitative Results

Two sets of experiments were conducted to quantitatively assess the proposed framework. In the first case, the above mentioned dataset of 300 RGB images has been utilized. The detection results of $PSO-HL$, $Hough-HL$ [10], and Deep Learning ($CNN-HL$) [8], were acquired and visually verified by a trained human-operator. In all cases, the results were deemed as correct or erroneous depending on whether the obtained HL visually agreed with the one suggested by the operator. $PSO-HL$ significantly outperformed $Hough-HL$ and $CNN-HL$ with a detection accuracy of **77.66%** contrasted to the very low figure of **39%** and **62.66%**.

Finally, an experiment using an Asctec Firefly UAV in a realistic environment was performed to assess correct estimation of ϕ_{roll} and θ_{pitch}. Images were captured with the on-board RGB-camera at 20 FPS, 640×480 resolution, 74° field of view, and a focal length f of 3.67mm. HL detection and computation of ϕ_{roll} and θ_{pitch} are performed on a remote PC exchanging data with the Firefly over WIFI. Given that the mentioned experiments were conducted outdoors, no ground-truth information is available. To this end, we utilized a complementary filter to provide accurate ϕ_{roll} and θ_{pitch} estimates from the on-board IMU data. The latter estimates, are employed as valid substitutes for the missing ground-truth data.

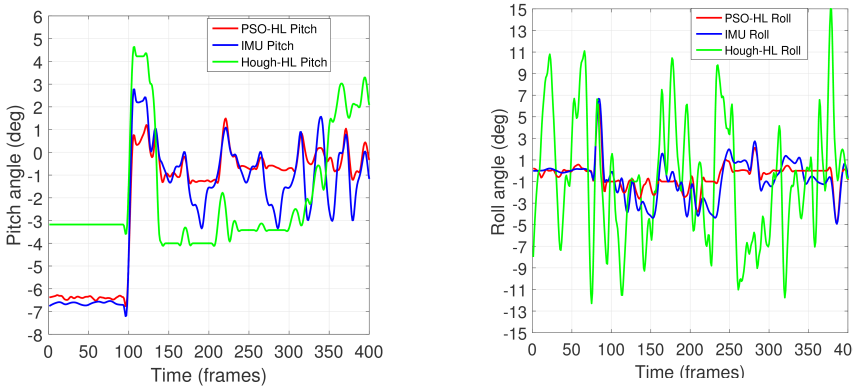

Fig. 4. UAV angles estimation. Left column: UAV's pitch angle; right column: UAV's roll angle, IMU data (blue line), $PSO-HL$ (red line), and $Hough-HL$ (green line).

Figure 4 illustrates the ϕ_{roll} and θ_{pitch} estimation results. In both cases, angles estimation results are presented for IMU data, $PSO-HL$, and $Hough-HL$. The ϕ_{roll} root mean square error for $PSO-HL$ over IMU data is **1.55** deg, compared to **8.02** deg for $Hough-HL$. The θ_{pitch} root mean square error for $PSO-HL$ over IMU data is **1.04** deg, compared to **2.90** deg for $Hough-HL$. Evidently, the estimation of both angles was carried out with improved accuracy by the proposed optimization methodology.

4 CONCLUSIONS

In this work a novel HL detection framework was presented and experimentally assessed. The framework is based on the Edge-Hough line detector, with a further optimization by a PSO algorithm. The latter utilizes an objective function that effectively fuses three different features, namely color, texture, and SIFT. BOW is employed to readily calculate the objective function for candidate HLs and provide a compact way for representing image content and fusing the comparison of different cues. In addition, the detected HL was utilized for computing two important UAV navigation quantities, namely roll and pitch angles.

Extensive experimentation with images acquired from a flying UAV revealed the accuracy and robustness of the proposed approach compared to the $Hough-HL$ and $CNN-HL$ detectors. Finally, the derived roll and pitch angles were contrasted to the

ones provides by the UAV's on-board IMU validating our approach. In future work, we aim at further improving HL detection via the inclusion of image vanishing points, continuously track the horizon, and additionally utilize it in real UAV navigation scenarios.

Acknowledgments

This work has been supported by the project "Advanced Research Activities in Biomedical and Agro alimentary Technologies" (MIS 5002469) which is implemented under the "Action for the Strategic Development on the Research and Technological Sector", funded by the Operational Programme "Competitiveness, Entrepreneurship and Innovation" (NSRF 2014-2020) and co-financed by Greece and the European Union (European Regional Development Fund).

References

1. D.Dusha, W.W.Boles and R.Walker, Fixed-Wing Attitude Estimation Using Computer Vision Based Horizon Detection, Australian Intl. Aerospace Congress, 2007.
2. W.N.Lie, T.C.I.Lin, T.C.Lin and K.S.Hung, A Robust Dynamic Programming Algorithm to Extract Skyline in Images for Navigation, Pattern Recogn. Lett., 2005.
3. R.J.D.Moore, S.Thurrowgood, D.Bland, D.Soccol and M.V.Srinivasan, A fast and adaptive method for estimating UAV attitude from the visual horizon, Intl. Conf. on Intelligent Robots and Systems, 2011.
4. J.A.Christian, Optical Attitude Determination from Horizon Orientation Using Image Segmentation, Journal of Guidance, Control and Dynamics, 2013.
5. S.Piperakis, E.Orfanoudakis and M.G.Lagoudakis, Predictive control for dynamic locomotion of real humanoid robots, Intl. Conf. on Intelligent Robots and Systems, 2014.
6. S.Piperakis and P.Trahanias, Non-linear ZMP based state estimation for humanoid robot locomotion, Intl. Conf. on Humanoid Robots, 2016.
7. S.Timotheatos, G.Tsagkatakis, P.Tsakalides and P. Trahanias, Feature Extraction and Learning for RSSI based Indoor Device Localization, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, 2017.
8. S.Workman, M.Zhai and N.Jacobs,Horizon Lines in the Wild, British Machine Vision Conf. 2016.
9. O.Oreifej and N.Lobo and M.Shah, Horizon constraint for unambiguous UAV navigation in planar scenes, Intl. Conf. on Robotics and Automation, 2011.
10. B.Zafarifar, H.Weda and P.H.N.de With, Horizon detection based on sky-color and edge features, Proc.SPIE, 2008.
11. T.Ahmad, G.Bebis, M.Nicolescu, A.Nefian and T.Fong, Fusion of edge-less and edge-based approaches for horizon line detection, Intl. Conf. on Information, Intelligence, Systems and Applications, 2015.
12. J.Kennedy and R.Eberhart, Particle swarm optimization, Intl. Conf. on Neural Networks, 1995.
13. Hough and V.C.Paul, Method and Means for Recognizing Complex Patterns, 1962.
14. I.Oikonomidis, N.Kyriazis, and A.A.Argyros, Efficient model-based 3d tracking of hand articulations using kinect, British Machine Vision Conference, 2011.
15. T.Leung and J.Malik, Representing and recognizing the visual appearance of materials using three-dimensional textons, Intl. Journal of Computer Vision, 2007.