Article

Attitude estimation using horizon detection in thermal images

Adrian Carrio ⓘ**, Hriday Bavle and Pascual Campoy**

Abstract
The lack of redundant attitude sensors represents a considerable yet common vulnerability in many low-cost unmanned aerial vehicles. In addition to the use of attitude sensors, exploiting the horizon as a visual reference for attitude control is part of human pilots' training. For this reason, and given the desirable properties of image sensors, quite a lot of research has been conducted proposing the use of vision sensors for horizon detection in order to obtain redundant attitude estimation onboard unmanned aerial vehicles. However, atmospheric and illumination conditions may hinder the operability of visible light image sensors, or even make their use impractical, such as during the night. Thermal infrared image sensors have a much wider range of operation conditions and their price has greatly decreased during the last years, becoming an alternative to visible spectrum sensors in certain operation scenarios. In this paper, two attitude estimation methods are proposed. The first method consists of a novel approach to estimate the line that best fits the horizon in a thermal image. The resulting line is then used to estimate the pitch and roll angles using an infinite horizon line model. The second method uses deep learning to predict attitude angles using raw pixel intensities from a thermal image. For this, a novel Convolutional Neural Network architecture has been trained using measurements from an inertial navigation system. Both methods presented are proven to be valid for redundant attitude estimation, providing RMS errors below $1.7°$ and running at up to 48 Hz, depending on the chosen method, the input image resolution and the available computational capabilities.

Keywords
Attitude estimation, horizon detection, thermal image, unmanned aerial vehicle, Convolutional Neural Network, deep learning

Received 19 May 2017; accepted 2 September 2018

Introduction and related works

Unmanned aerial vehicles (UAVs) are increasingly being used in both military and civilian applications such as search and rescue, surveillance, precision agriculture, etc. In order to perform precise navigation, it is necessary to accurately estimate the state of the UAV. In particular, estimating the attitude of an UAV is a critical aspect, as inaccuracies or sensor failures can lead to entire system failure and accidents. These requirements demand the use of redundant attitude sensors onboard UAVs. Image sensors are of particular interest onboard micro air vehicles due to their desirable properties, i.e. small-size, lightweight and low-power. The attitude provided by an inertial measurement unit (IMU) or from an inertial navigation system (INS) can be fused with the measurements derived from an image sensor to improve attitude estimations. However, due to different atmospheric and illumination conditions, the operation of visible light image sensors can be affected and cause estimation errors. Thermal infrared image sensors provide a wider range of operational illumination conditions, including operability during the night. The size and cost of thermal image sensors has experienced a great reduction in the last years, allowing it to be used onboard UAVs. Several applications involving thermal

Computer Vision & Aerial Robotics Group, Universidad Politécnica de Madrid (UPM-CSIC), Madrid, Spain

Corresponding author:
Adrian Carrio, Computer Vision & Aerial Robotics Group, Technical University of Madrid, Madrid, 28003, Spain.
Email: adrian.carrio@upm.es

cameras onboard UAVs have already been proposed, such as person detection,[1] water stress assessment in agriculture[2] and aircraft detection.[3] Thermal cameras can therefore be exploited as reliable and accurate imaging sensors, capable of providing attitude measurements, even in the presence of challenging illumination and atmospheric conditions, thus providing sensing redundancy and more reliability for UAV systems.

Quite some research has been carried out on attitude estimation using imaging sensors. Many efforts have been focused on horizon line detection alone, due to its relevance for visual geo-localization, port security, etc. Shabayek et al.[4] present a comprehensive review of the available algorithms for vision-based horizon line detection and attitude estimation for UAVs. They propose a horizon segmentation approach using polarization information of the surrounding. Fefilatyev et al.[5] present machine learning techniques for horizon segmentation, comparing the performance of SVM, J48 and Naive-based classifiers but attitude is not estimated. Horizon detection in infrared images is presented by Bazin et al.[6] In contrast to our approach, catadioptric images are used. Due to the unavailability of catadioptric mirrors, the authors manually synthesize the catadioptric images. Ahmad et al.[7] present a detailed comparison between two machine learning-based approaches for horizon line detection (edge and edgeless-based). Additionally, they propose a fusion strategy, combining both approaches for better results, but they fail to provide attitude estimations using the detected horizon. Other learning-based approaches are present in Hung et al.,[8] Ahmad et al.,[9] Gurghian et al.,[10] Porzi et al.[11] and Ahmad et al.[12]

Several research works have focused on attitude estimation based on horizon line detection. Cornall et al.[13] propose a method for aircraft attitude estimation using the detected horizon line. Dusha et al.[14] propose the use of computer vision techniques for horizon detection along with optical flow in order to compute the pitch and roll of the aircraft along with the three body rates. They use an extended Kalman filter (EKF) to propagate and track the horizon line candidates in successive image frames; however, the authors fail to provide accuracy metrics for validating their approach. Ettinger[15] implemented a horizon line detection algorithm using differences in the color distribution between the ground and sky regions. The computed horizon line is then used for attitude estimation and control of a fixed wing UAV. However, the estimated attitude is not compared to any ground truth data for validating the obtained results. Bao et al.[16] also propose one of the earliest computer vision-based methods for horizon line detection for micro aerial vehicles (MAVs) and attitude estimation. Nonetheless, the

performance of this approach is hard to evaluate from the manuscript, as no detections are depicted and no computed attitude data are provided. Hwangbo and Kanade[17] implement an attitude estimation method using line segments extracted from images to correct the accumulated error of onboard gyroscopes. The algorithm is only proven to be effective in urban areas, assuming a single vertical vanishing point or multiple horizontal vanishing points. Thurrowgood et al.[18] present a technique for estimating, controlling and stabilizing the 2-DOF attitude (roll and pitch angles) of an UAV using a dual fisheye camera placed back-to-back incorporating color and intensity information. Images need to be stitched together in order to obtain a spherical view of the environment. In the presence of insufficient features, failures while stitching the images can lead to errors in the horizon detection, which is performed using computer vision techniques. The attitude is then estimated using the position, shape and orientation of the detected horizon profile. Moore et al.[19] use the same setup as Thurrowgood et al.[18] and apply machine learning techniques to detect the horizon in several sequences of images. Moore et al.[20] apply the previous machine learning techniques for horizon detection in order to obtain the roll and pitch angles along with the heading of the UAV. The aforementioned methods require a dual camera setup and are critically conditioned by illumination. Mondragón et al.[21] use catadioptric cameras for horizon line detection as well as attitude and heading estimation of UAVs. The reported work requires higher resolution images as compared to our proposed approach. Dusha et al.[22] use a basic front-end image processing based on edge detection techniques to detect horizon line candidates and optical flow to reject false detections. The attitude is then computed from the detected horizon line, and body rates are computed using optical flow. An EKF with a constant velocity model is used to filter attitude estimations. In spite of the promising results, this technique fails to accurately detect the horizon line without an additional optical flow computation. Tehrani et al.[23] propose to fuse attitude estimations from a UV-filtered panoramic vision sensor with angular rate estimations from optic flow. This method requires a sophisticated camera system including a combination of a CCD image sensor, an ultraviolet filter and panoramic mirror lens. Grelsson and Felsberg[24] and Grelsson et al.[25] propose a new method for attitude estimation. Nonetheless, no performance evaluation, i.e. comparison with the ground truth data, is reported. Dumble and Gibbens[26] present a computer vision-based horizon line detection and fitting method for estimating roll and pitch angles. The authors use an m-estimator to fit a line to the extracted profile by using iterated

re-weighted least squares minimization. The attitude is then determined using an infinite horizon line model.

Several other researches have also been carried using digital elevation models, but these methods require a pre-generated terrain model in order to detect the horizon and estimate the attitude and typically fail to provide real-time performance. Naval et al.[27] present a method for estimating the position and orientation of a camera using a single image of a mountain, by aligning the terrain skyline with a synthetic skyline generated from a digital elevation map. Behringer[28] use terrain silhouettes which provide unique features to be matched with the data from a digital elevation map. The best matched data are used to provide the camera orientation. Woo et al.[29] propose a vision-based system for UAV localization in mountain areas. The UAV uses a digital elevation map, infrared image sequences and an altimeter to estimate the position and orientation of the UAV. Gupta and Brennan[30] implement a novel method for estimating the roll, pitch and yaw angles by fusing image and inertial sensors. The horizon lines obtained from the captured video are compared to the horizon lines from a digital elevation map, obtaining absolute roll, pitch and yaw orientation. A Kalman filter is used to fuse the IMU measurements with the scene matching estimates. Grelsson et al.[31] present a recent work on attitude estimation for airborne vehicles, with high-resolution images (30 MP) and fisheye lens, but it is quite far from offering real time performance. Dumble and Gibbens[32] propose a position and attitude estimation technique by aligning the measured horizon line profile to an off-line pre-generated terrain-aided reference. The results are obtained using simulated flights sequences, which take almost 4 s to generate and an additional

53 ms for complete optimization. Some of these approaches and their performance are summarized in Table 1.

In this paper, two methods for attitude estimation using horizon detection in thermal images are proposed. The first method uses a novel approach to estimate the line which best fits the horizon in a thermal image. The slope and intercept of this line are then used to estimate pitch and roll angles following the horizon model presented in Dumble and Gibbens.[26]

In the second method, a novel Convolutional Neural Network (ConvNet) architecture is trained to predict attitude angles directly from thermal images. Both methods are evaluated and compared with other methods reported in the literature. To the authors' best knowledge, this is the first time that thermal infrared imaging is used for attitude estimation using horizon detection.

The remainder of this paper is as follows. Firstly, the two proposed attitude estimation algorithms, infinite horizon and ConvNet-based, are presented. Secondly, the data acquisition procedure followed is explained. Then, the evaluation methodology is discussed. Next, the results obtained with the proposed methods are presented and compared with other relevant approaches in the state of the art. Finally, we present the conclusions and future works.

Attitude estimation algorithms

Infinite horizon method

The proposed infinite horizon method consists of two steps: estimation of the parameters of the line that best fits the horizon in an image and subsequent estimation

Table 1. Comparison of vision-based attitude estimation methods.

Reference	Image resolution (pix)	RMS pitch error (deg)	RMS roll error (deg)	Processing time (s)	Requires a terrain model
Dumble and Gibbens[26]	1024 × 768	0.01	0.28	0.024	No
Dumble and Gibbens[32]	1024 × 768	0.07	0.02	0.053	Yes
Dusha et al.[14,22]	1024 × 768	2.07	0.38	0.112	No
Ettinger[15]	245 × 192	5.12	2.08	0.050	No
Grelsson et al.[25]	5616 × 3744	0.25	0.25	13–25	Yes
	3248 × 4872				
Gupta and Brennan[30]	720 × 480	0.14	0.01	0.033	Yes
Hwangbo and Kanade[17]	320 × 240	0.61	0.49	0.200	No
Moore et al.[19,20]	80 × 40	1.49	1.49	0.002	No
Tehrani et al.[23]	200 × 200	1.83	1.62	0.008	No
Infinite horizon method	620 × 476	0.52	1.33	0.302	No
ConvNet-based method	128 × 128	1.53	1.12	0.021	No

Note: Averaged results on all the three sequences are provided for the infinite horizon method. For the ConvNet-based method, the training set is excluded and only the results on the test sets are averaged.
ConvNet: Convolutional Neural Network.

of the pitch and roll angles using the method proposed in Dumble and Gibbens,[26] which itself is an extension of the method proposed in Dusha et al.[14] for a pin-hole camera model.

The estimation of the parameters a and b of the line given by $v = au + b$ which best fits the horizon line is computed as follows. Firstly, radial and tangential distortion is removed from the images using the parameters extracted from the camera calibration. Then, a bilateral filtering operation is applied in order to remove noise while preserving edges, and the image is equalized in order to enhance contrast. The obtained image is then sampled into non-overlapping columns of constant width. For each of these subimages, an estimation of the row corresponding to the horizon will be obtained. For this, the average intensity value for each row in the subimage is computed. The first order derivative of the average intensity values along the vertical direction is obtained using a Sobel operator. These derivatives are accumulated along the vertical direction, from the top to the bottom of the image. Whenever the accumulated derivative exceeds a threshold value, that row is assigned as the one corresponding to the horizon. The threshold value was set to 0.5, which empirically provided best results.

Computing the horizon row in each of these subimages provides the (x, y) coordinates of a point in the image plane, where x corresponds to the first column of each sample and y is the row corresponding to the horizon in each subimage. A line is fitted to these points by minimizing $\sum_i \rho(r_i)$, where r_i is the distance between the i-th point and the line, and ρ_i is given by equation (1), where $C = 1.3998$. This value of C provides 95% asymptotic efficiency on the standard normal distribution. The points for each sub-image and the estimated line are shown in Figure 1(b) and (d).

$$\rho(r) = C^2 \left[\frac{r}{C} - \log\left(1 + \frac{r}{C}\right) \right] \quad (1)$$

Once a and b are known, the pitch and roll angles, here noted as θ and ϕ respectively, can be obtained using equations (2) and (3), where (f_x, f_y) are the focal lengths expressed in pixels and (c_x, c_y) are the coordinates of the principal point, all obtained from camera calibration.

$$\phi = -arctan\left(\frac{f_x}{f_y} a\right) \quad (2)$$

$$\theta = arctan\left(\frac{(b - c_y)cos\phi}{f_y} - \frac{c_x sin\phi}{f_x}\right) \quad (3)$$

Since this method samples the horizon line at multiple points, it is expected to be somewhat robust to partial occlusions of the horizon. However, large occlusions, caused for example by large clouds or fog, could noticeably affect the performance of both of the
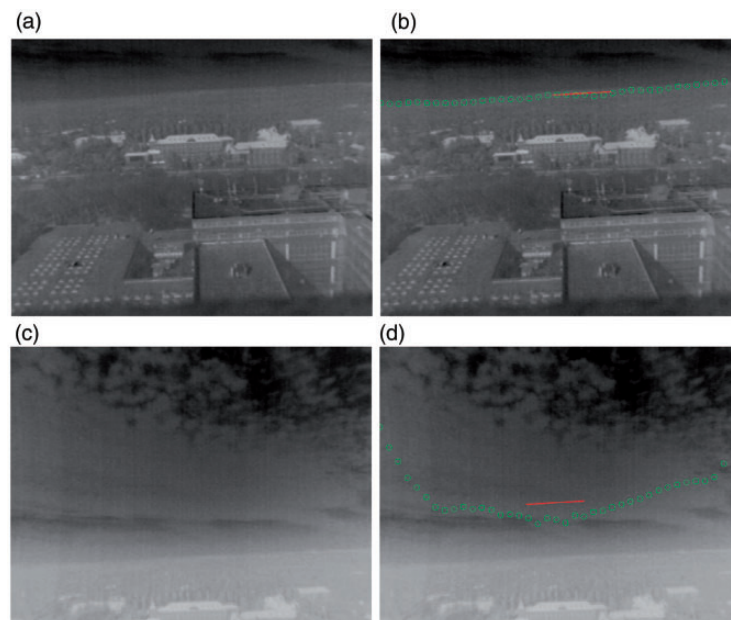

Figure 1. Input frames (left) and frames showing the detected horizon line (right) using the infinite horizon method. (a) and (b) Images showing correct estimation. (c) and (d) Images showing how the estimation can fail due to the autogain feature in the thermal camera, which here reduces the intensity gradient between sky and ground in order to get the most detail out of the scene.

methods presented in this paper. Sparse fog is the only scenario that could prevent correct operation of an infrared image sensor while a visible spectrum image sensor would still be able to perceive the horizon, provided that good illumination is available.

ConvNet-based method

As an alternative to the infinite horizon method, we propose to approach the attitude estimation problem as a regression learning problem. Given a thermal image capturing the horizon, a ConvNet model trained with images and their corresponding camera attitude angles can be used to predict the corresponding pitch and roll camera angles for unseen images. The trained model is not expected to perform better than the IMU/INS used to acquire the data. However, a model could be trained using a high-end IMU/INS, replicated and used in multiple UAVs at a lower cost. These vision-based measurements could also be fused for redundancy with those coming from a low-cost IMU to achieve much more accurate measurements than using the IMU only.

Different sets of parameters are learned for estimating pitch and roll, but the same model architecture is used for both angles. We propose to use different sets of parameters to totally decouple both estimations. However, it seems reasonable that some of the image features extracted to estimate pitch could also be used to estimate roll and vice versa. Keeping a model for pitch and another for roll also allows to run them in parallel, increasing the performance. In our implementation, images are downsampled to a resolution of 128 × 128 pixels and distortion is removed before they are fed to the model. The model has been designed empirically, trying to keep a low number of parameters in order to achieve speed and requiring a relatively low number of images for successful generalization. The proposed model has almost 12 M trainable parameters and its architecture is summarized in Table 2.

Data acquisition

Sensors

A Thermoteknix MicroCAM 384 analog thermal camera has been used in the experiments. This camera weights around 80 g and features a 7.5 mm lens. The camera provides analog images of 384 × 288 at 60 fps. An A/D video converter has been used, providing a 640 × 480 digital image with black padding, which is cropped to obtain a 620 × 476 image. These images are acquired at 20 fps using the Robot Operating System.

A high-end SBG Ellipse2-N-G4A2-B1 INS with integrated Global Navigation Satellite System

Table 2. ConvNet model architecture proposed.

Layer	Size	Number of parameters
Input	128 × 128	–
2D Convolution (5 × 5 kernel)	48 × 124 × 124	1248
Max pooling	48 × 62 × 62	–
Dropout	48 × 62 × 62	–
Flatten	184,152	–
Fully connected	64	11,808,832
Fully connected	1	65

Note: The model takes a 128 × 128 pixel image as an input and predicts the pitch or roll angle of the camera in degrees.

(GNSS) has been used for measuring attitude. The system is able to estimate roll and pitch angles with an RMS accuracy of 0.1°. The sensor is 46 × 45 × 24 mm and weights 47 g.

Expected errors

Errors inherent to sensing technologies, as well as inaccuracies when using mathematical models, emerge while operating in real world scenarios. In the following paragraphs, we summarize the most relevant errors intervening.

INS sensor errors

Even though attitude measurements from an IMU or INS sensor are valid for many applications, they still contain inaccuracies. The main sources of errors in an INS sensor are:

- *Sensor bias*: For a given physical input, the measurements provided by the IMU can be offset by a specific value, known as the bias.
- *Sensor noise*: When the IMU measures a constant signal, a random noise is always present in the measurement. Due to this random sensor noise error, the error in the final solution can grow unbounded.
- *Scale factor*: Scale factor is a relation between the input to the IMU and its corresponding output. For a given input, the IMU provides an output proportional to the input but scaled.
- *Poor filtering*: Bad performance in filtering can happen mainly when the system modelled is highly non-linear or due to a poor covariance estimation.

Image-derived errors

Analogously, errors are present in image-based attitude estimations due to different causes, which are enumerated next:

- *Camera model:* Errors coming from the assumption of a pinhole camera model and from the camera calibration process.
- *Horizon line extraction*: Errors coming from the assumption of lack of a terrain profile in the horizon (flat horizon) and from image noise.
- *Infinite horizon model*: Using a mathematical model such as the infinite horizon model (vanishing line of the ground plane) is an approximation which implies errors.

Video-inertial dataset

The INS and thermal camera were mounted on a platform restricting relative translation and rotation between the sensors. A video-inertial dataset was acquired using this platform in order to train and evaluate the proposed algorithms. These images were obtained by rotating the platform manually from a rooftop, situated about 40 m above the ground.

Three different video-inertial sequences with rotations in both pitch and roll axes were acquired. The first sequence was captured at dusk and is used for training the ConvNet model (927 frames). Two sequences are obtained for testing each of the methods proposed: one captured during the day (1025 frames) and one captured at night (1050 frames).

The scenes seen during the acquisition of the sequences used for training the ConvNet model and those seen for the evaluation of the algorithms are different as it can be observed in Figure 2. This was done in order to ensure the ConvNet model is generalizing properly and the method can be successfully applied on scenes different from those in the training set.

Evaluation methodology

Regarding the evaluation methodology, the INS data have been considered as ground-truth reference, comparing it with image-based estimations from both

Figure 2. Examples of RGB images (left) and TIR images (right) from the three video sequences captured. (a) and (b) Images captured during dusk, and the TIR images from this sequence were used for training the model. (c) and (d) Images captured with a lot more light during the day, pointing at a different location. (e) and (f) Images captured at night pointing at a different location from all previous sequences. Only infrared images were used in the experiments. RGB images are shown here to point out the sunlight differences in the scenes.

proposed methods. It is assumed that the aforementioned INS errors will be present.

We compare the attitude measurements for the pitch and roll angles using RMS error as expressed in equation (4), where x_i^{img} and x_i^{IMU} are synchronized measurements. x_i^{img} is the pitch/roll measurement obtained from frame i, and x_i^{IMU} is the pitch/roll measurement obtained from the INS. n is the total number of frames in the video sequence.

$$RMS = \sqrt{\frac{\sum_{i=1}^{n} \left(x_i^{img} - x_i^{IMU} \right)^2}{n}} \qquad (4)$$

When using the infinite horizon method, since the pitch/roll axes in the camera and INS sensor are not perfectly aligned, this misalignment will produce a

constant error between measurements from both sensors. These errors do not appear in the ConvNet-based method as the model is trained to follow the INS measurements. We propose to correct this misalignment in the infinite horizon method by extracting the median of the first 70 INS measurements (while the system is static) and using it as an offset value to correct the pitch and roll estimations.

ConvNet training

The ConvNet model proposed was trained using the Keras API with Theano as the tensor manipulation library. Each model was trained for a maximum of 2000 epochs with a mini-batch of size 32 with the Adam optimizer and using the mean square error as loss function. Early stopping was used to avoid overfitting, using 30% of the training images for validation. Training took about 15 min for both models (pitch and roll) using a Nvidia Geforce GTX 970 GPU. The training and validation errors are summarized in Table 3.

Table 3. Train and validation errors obtained during the training process of the ConvNet models.

	RMS train error (deg)	RMS validation error (deg)
Pitch	0.30	0.60
Roll	0.19	0.69

Results

The average processing time for the infinite horizon method corrected is 302 ms (~3 Hz) and 21 ms (~47 Hz) for the ConvNet method, both tested in an Intel
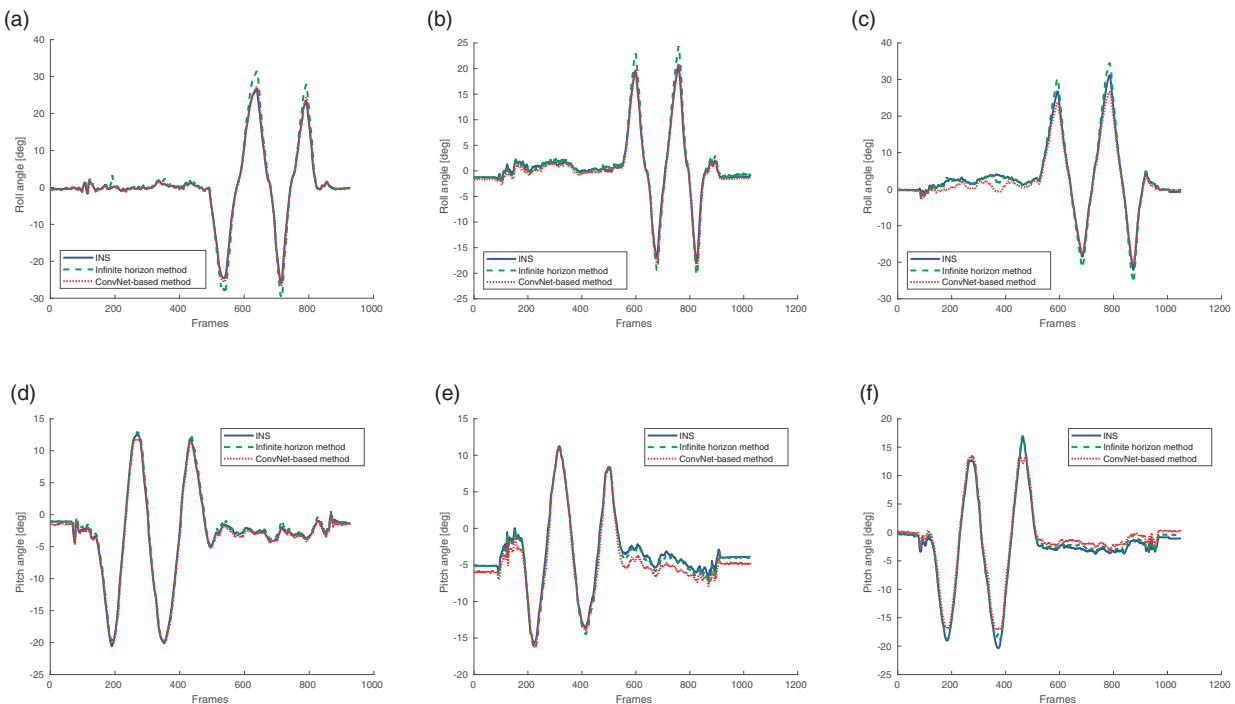

Figure 3. Attitude estimations for roll and pitch angles are respectively shown in the top and bottom rows. (a) and (d) The sequence captured at dusk and used for training the ConvNet. (b) and (e) The sequence captured during the day. (c) and (f) The sequence captured during the night.

Table 4. RMS pitch and roll errors on the three sequences (dusk/day/night) obtained with the infinite horizon method before and after the median-based correction, and using the ConvNet-based method.

	RMS pitch error (deg)	RMS roll error (deg)
Infinite horizon method raw	0.66/0.84/1.43	2.86/2.47/3.14
Infinite horizon method corrected	0.54/0.51/0.52	1.64/1.13/1.23
ConvNet-based method	0.23/1.40/1.65	0.50/0.69/1.55

ConvNet: Convolutional Neural Network.

Core i7 CPU running at 2.4 Ghz. No code optimizations were implemented.

Pitch and roll angles measured by the INS and estimated by both image-based methods are shown in Figure 3. In general, image-based estimations follow the INS measurements, with some exceptions. With respect to the roll angle, the infinite horizon method seems to over-estimate roll angles greater than $\pm 20°$. This could be explained by the intensity balance corrections performed by the autogain feature of the thermal camera. Another interesting effect is the bias observed in the pitch angle estimation in the test sequences by the ConvNet.

RMS errors are shown in Table 4. We can see how errors are higher in roll than in pitch. This can be explained by the fact that pitch angles reached during the tests were smaller. It can also be seen how the infinite horizon method slightly outperforms the ConvNet-based method in terms of accuracy, with the exception of the training set which should not be considered to evaluate the ConvNet method. However, the input resolution of the images in both methods is very different, as the infinite horizon method takes a 620×476 pixel image as input while the ConvNet-method uses a 128×128 pixel image, with 18 times less pixels. This explains the differences in both speed and accuracy.

These results are compared to the current state of the art methods in Table 1. Only references explicitly reporting attitude errors, image resolution and processing time have been included. While the infinite horizon method outperforms only some of the proposed methods, the ConvNet-based method outperforms all other methods using similar input image sizes either in terms of accuracy[15,23] or speed.[17]

Furthermore, the vision-based measurements obtained with the proposed methods were not filtered for simplicity of the solution, in contrast to some of the methods surveyed. Filtering techniques could definitely improve the accuracy of the proposed methods. More importantly, our approaches are based on thermal images, therefore being capable of providing attitude estimations across a much broader set of illumination conditions than any of the other methods surveyed.

Conclusions and future works

In this paper, two attitude estimation algorithms using horizon detection in thermal images have been proposed. The first method is based on the estimation of the line that best fits the horizon, which is used to obtain the pitch and roll angles using an infinite horizon line model. The second method uses a ConvNet trained with INS measurements to perform regression using raw pixel intensities. This ConvNet could be trained using a high-end INS, replicated and used in multiple UAVs at a lower cost. Also, these vision-based measurements could also be fused for redundancy with those coming from a low-cost IMU to achieve much more accurate measurements than using an IMU only.

The proposed methods have been exhaustively compared to current existing approaches. For the methods published with similar input image sizes, our ConvNet method outperforms current state of the art, either in terms of accuracy or speed. Furthermore, our approach could potentially be used under a much broader set of illumination conditions, including during the night.

Among the two methods proposed, higher accuracy is obtained with the infinite horizon method. However, this method takes as input an image of greater resolution and requires longer processing time. Both methods are proven to be valid for redundant attitude estimation, providing RMS errors below $1.7°$ and running at up to 48 Hz, depending on the chosen method, the input image resolution and the available computational capabilities.

As for future works, we plan to filter the vision-based measurements in order to improve accuracy by exploiting temporal information. Additionally, the use of recurrent neural networks will be explored. We will also evaluate the impact of partial occlusions of the horizon and test onboard operation capabilities.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ORCID iD

Adrian Carrio ⓘ http://orcid.org/0000-0001-6711-7279

References

1. Rudol P and Doherty P. Human body detection and geolocalization for UAV search and rescue missions using color and thermal imagery. In: *2008 IEEE aerospace conference*, pp. 1–8. Big Sky, Montana: IEEE, 2008.

2. Zarco-Tejada PJ, González-Dugo V and Berni JA. Fluorescence, temperature and narrow-band indices acquired from a UAV platform for water stress detection using a micro-hyperspectral imager and a thermal camera. *Remote Sens Environ* 2012; 117: 322–337.

3. Carrio A, Lin Y, Saripalli S, et al. Obstacle detection system for small UAVs using ADS-B and thermal imaging. *J Intell Robot Syst* 2017; 88: 583–595.

4. Shabayek AER, Demonceaux C, Morel O, et al. Vision based UAV attitude estimation: progress and insights. *J Intell Robot Syst* 2012; 65: 295–308.

5. Fefilatyev S, Smarodzinava V, Hall LO, et al. Horizon detection using machine learning techniques. In: *5th International conference on machine learning and applications (ICMLA'06)*, pp. 17–21. Orlando, Florida: IEEE, 2006.

6. Bazin J-C, Kweon I, Demonceaux C, et al. Dynamic programming and skyline extraction in catadioptric infrared images. In: *IEEE international conference on robotics and automation (ICRA'09)*, pp. 409–416. Kobe, Japan: IEEE, 2009.

7. Ahmad T, Bebis G, Nicolescu M, et al. Fusion of edgeless and edge-based approaches for horizon line detection. In *2015 6th International conference on information, intelligence, systems and applications (IISA)*, pp. 1–6. Ionian University, Corfu, Greece: IEEE, 2015.

8. Hung Y-L, Su C-W, Chang Y-H, et al. Skyline localization for mountain images. In: *2013 IEEE international conference on multimedia and expo (ICME)*, pp. 1–6. San Jose, California, USA: IEEE, 2013.

9. Ahmad T, Bebis G, Nicolescu M, et al. An edge-less approach to horizon line detection. In: *2015 IEEE 14th international conference on machine learning and applications (ICMLA)*, pp. 1095–1102. Miami, Florida, USA: IEEE, 2015.

10. Gurghian A, Koduri T, Bailur SV, et al. DeepLanes: end-to-end lane position estimation using deep neural networks. In: *CVPR workshops*, pp. 38–45, 2016. https://ieeexplore.ieee.org/document/7789502

11. Porzi L, Rota S Bulò et al. A deeply-supervised deconvolutional network for horizon line detection. In: *Proceedings of the 2016 ACM on multimedia conference*, pp. 137–141. Amsterdam, The Netherlands: ACM, 2016.

12. Ahmad T, Campr P, Čadik M, et al. Comparison of semantic segmentation approaches for horizon/sky line detection. In: *2017 International joint conference on neural networks (IJCNN)*, pp. 4436–4443. Anchorage, AK, USA: IEEE, 2017.

13. Cornall T, Egan G and Price A. Aircraft attitude estimation from horizon video. *Electron Lett* 2006; 42: 744–745.

14. Dusha D, Boles WW and Walker R. Fixed-wing attitude estimation using computer vision based horizon detection. 2007, https://eprints.qut.edu.au/6852/

15. Ettinger SM. *Design and implementation of autonomous vision-guided micro air vehicles*. PhD Thesis, University of Florida, USA, 2001. Vail, CO, USA.

16. Bao G, Zhou Z, Xiong S, et al. Towards micro air vehicle flight autonomy research on the method of horizon extraction. In: *Proceedings of the 20th IEEE instrumentation and measurement technology conference (IMTC'03)*, vol. 2, pp. 1387–1390. IEEE, 2003.

17. Hwangbo M and Kanade T. Visual-inertial UAV attitude estimation using urban scene regularities. In: *2011 IEEE international conference on robotics and automation*, pp. 2451–2458, 2011. https://ieeexplore.ieee.org/document/5979542

18. Thurrowgood S, Moore RJ, Bland D, et al. UAV attitude control using the visual horizon. In: *Proceedings of the Australasian conference on robotics and automation*, San Francisco California: Brisbane, Australia, 2010.

19. Moore RJ, Thurrowgood S, Bland D, et al. A fast and adaptive method for estimating UAV attitude from the visual horizon. In: *2011 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp. 4935–4940. IEEE, 2011.

20. Moore RJ, Thurrowgood S, Soccol D, et al. A method for the visual estimation and control of 3-DOF attitude for UAVs. In: *Proceedings of thirteenth Australasian conference on robotics and automation (ACRA 2011)*, vol. 142, 2011. http://www.araa.asn.au/conferences/acra-2011/

21. Mondragón IF, Campoy P, Martinez C, et al. Omnidirectional vision applied to unmanned aerial vehicles (UAVs) attitude and heading estimation. *Rob Auton Syst* 2010; 58:809–819.

22. Dusha D, Mejias L and Walker R. Fixed-wing attitude estimation using temporal tracking of the horizon and optical flow. *J Field Robot* 2011; 28:355–372.

23. Tehrani MH, Garratt MA and Anavatti SG. Low-altitude horizon-based aircraft attitude estimation using UV-filtered panoramic images and optic flow. *IEEE Trans Aerosp Electron Syst* 2016; 52: 2362–2375.

24. Grelsson B and Felsberg M. Probabilistic Hough voting for attitude estimation from aerial fisheye images. In: *Scandinavian conference on image analysis*, pp. 478–488. Espoo, Finland: Springer, 2013.

25. Grelsson B, Felsberg M and Isaksson F. Efficient 7D aerial pose estimation. In: *2013 IEEE workshop on robot vision (WORV)*, pp. 88–95. Clearwater, FL, USA: IEEE, 2013.

26. Dumble SJ and Gibbens PW. Horizon profile detection for attitude determination. *J Intell Robot Syst* 2012; 68: 339–357.

27. Naval PC Jr, Mukunoki M, Minoh M, et al. Estimating camera position and orientation from geographical map and mountain image. In: *38th Research meeting of the pattern sensing group, society of instrument and control engineers*. Tokyo, Japan: Citeseer, 1997.

28. Behringer R. Improving the registration precision by visual horizon silhouette matching. In: *Proceedings of the first IEEE workshop on augmented reality*, 1998. http://iwar1998.vgtc.org/

29. Woo J, Son K, Li T, et al. Vision-based UAV navigation in mountain area. In: *MVA*, pp. 236–239, 2007. https://dblp.uni-trier.de/rec/bibtex/conf/mva/WooSLKK07

30. Gupta V and Brennan S. Terrain-based vehicle orientation estimation combining vision and inertial measurements. *J Field Robot* 2008; 25: 181–202.

31. Grelsson B, Felsberg M and Isaksson F. Highly accurate attitude estimation via horizon detection. *J Field Robot* 2015; 33: 967–993.

32. Dumble SJ and Gibbens PW. Efficient terrain-aided visual horizon based attitude estimation and localization. *J Intell Robot Syst* 2015; 78: 205–221.