

APPLICATION OF UNMANNED AERIAL VEHICLES IN MONITORING OF COMMUNICATION ROUTES ON COUNTRY AREAS

Damian Wierzbicki

Military University of Technology, Poland
damian.wierzbicki@wat.edu.pl

Abstract. Along with intensive development of automotive transport there is a real need of assurance of monitoring for the purpose of the improvement of safety of traffic and to increase the flow of traffic, particularly in the case of large routes. Applying unmanned aerial vehicles (UAVs) equipped with compact digital cameras is an outstanding solution. Currently UAVs are widely used in many areas, particularly in transport. Principal uses encompass the supervision of safety, monitoring of the strain of movement, monitoring of construction of roads and of connected research with the course of communication nets. Selected aspects of applying UAV equipped with compact digital camera for the monitoring traffic flow routes are described in the article. Imagery data were obtained with the fixed-wing Trimble UX-5 platform, which was equipped with the compact digital camera. Analysis of the usefulness of images for the monitoring of the traffic intensity was performed on the selected traffic routes. Detection of vehicles on the video was based on the method of Haar Feature-based Cascade Classifier from the OpenCV collection of a library. The training of the classifier consisted in preparing several dozen different samples containing cars and trucks. After the operation of training the classifier analyzing the chosen sequences of images was possible. The effectiveness of the detection of vehicles at the images deprived of the smear equals 64 %. It is a satisfying result. Considering the fact that images were not made in the nadir they were only obliques. In addition, wooded and urban areas were registered on the images besides the road and vehicles. For that reason the classifier in many cases recognized the crowns of trees as individual vehicles. Research methods applied at the work can be used and tested for the detection of vehicles on roads, like motorways or dual carriageways.

Keywords: transport, UAV, photogrammetry, video, computer vision.

Introduction

Along with intensive development of automotive transport in Poland there is a real need of assurance of monitoring for the purpose of the improvement of safety of traffic and to increase the flow of traffic, particularly in the case of large routes [1]. Applying unmanned aerial vehicles (UAV) equipped with compact digital camera is an outstanding solution. Currently unmanned aerial vehicles are widely used in many areas, particularly in transport. Principal uses encompass the supervision of safety, monitoring of the strain of movement, monitoring of construction of roads and of connected research with the course of communication nets [2]. Unmanned aerial vehicles are usually equipped with single-frequency GPS receivers. In this case, the typical accuracy of determining the coordinates of the platform can reach up to several meters [3]. The determined coordinates of the UAV by the GNSS receiver are referred to the geocentric system XYZ or the geodetic system BLh or the local ENU system.

In addition, compact digital cameras are mounted on their boards. Such platforms are an important complement to modern imaging systems for the needs of intelligent monitoring of road communication networks. With the increasing traffic and traffic flow, the use of unmanned platforms is becoming more and more popular. The classic acquisition of data on vehicle traffic is based mainly on cameras permanently mounted in the road lane. Such a solution is limited only to the local area, and thus is expensive and labor-intensive to perform monitoring of larger areas [4]. The advantages of using unmanned aerial vehicles to monitor road infrastructure are:

1. low monitoring costs on even several kilometers of road sections;
2. providing high spatial and temporal resolution;
3. ability to collect image data in different spectral ranges.

The above advantages are complemented by the ability to record images using continuous video sequences. Monitoring of road infrastructures provides the users with up-to-date information on the volume of traffic at a given moment.

Compared to individual images, video sequences can provide more information about objects (cars) and their changes over time, so it is important to develop vehicle tracking algorithms. In the course of the conducted research [2], it was observed that the matching and segmentation of an object

is one of the main issues, when it comes to tracking objects in video sequences. In other studies, an object tracking algorithm based on subtraction of the image background and image adjustment using a critical point filter (multiresolution) has been proposed. Other [5.6] researchers proposed in their work to improve the standard sampling strategy for object detection and tracking based on detection and tracking filters in video. However, this method had some barriers. Newer articles have shown that you can track objects automatically using the Mpeg2 vector to mark active objects in a clip. Then, the separated vector motion features are used to create the model. Experiments have shown that the algorithm can achieve good trajectory tracking effects and at the same time be immune to interference [6].

The aim of the research presented in the paper is to evaluate the effectiveness of the use of Haar Feature-based for the detection of cars based on image sequences acquired from a low altitude. A ready set of samples on the basis of 526 car images was used for the analysis. The proposed solution is based on an algorithm available in the OpenCV library.

Materials and methods

As part of the tests carried out, the detection of vehicles was based on the method of Haar Feature-based Cascade Classifier from the OpenCV collection of a library. The suggested Haar like method is based on the detection of the so-called haar like features. These features include, for example, the occurrence of edges or lines. An additional advantage is the speed of calculating their value. Haar like features are defined by: a characteristic template - each feature is a mask consisting of two or three rectangles (Fig. 1), coordinates of the feature relative to the search window, feature size - width and height. Each feature is calculated as the difference in brightness values of pixels that make up the white and black rectangles as well as the sum of pixel values covered by the black rectangle. The components of this difference have values of weights that are inversely proportional to the size. In this case, the differences between the two areas are compensated [7].

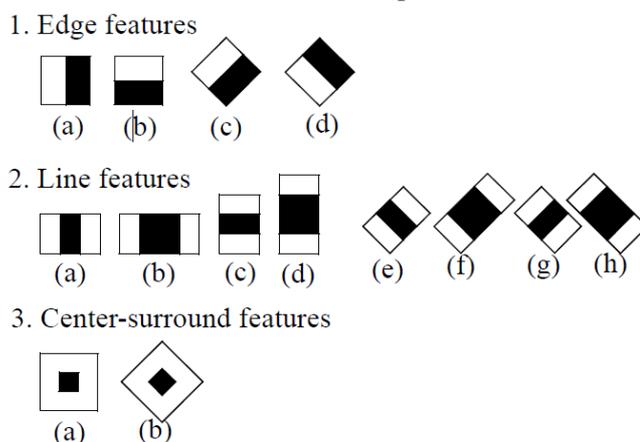


Fig. 1. Some examples of Haar-like features [8]

Rectangle features can be calculated very quickly using indirect image representations. The image in the place x, y contains the sum of pixels above and to the left x, y [7]:

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y'), \quad (1)$$

where $ii(x, y)$ – indirect image;
 $i(x, y)$ – the original one.

Calculations can be quickly made using the formulas:

$$s(x, y) = s(x, y - 1) + i(x, y), \quad (2)$$

$$ii(x, y) = ii(x - 1, y) + s(x, y), \quad (3)$$

where $s(x, y)$ – the cumulative row sum;
 $s(x - 1) = 0$, and $ii(-1 - y) = 0$.

The integral image can be calculated in one pass with the original image. Applying the classifier learning function requires providing a set of negative and positive samples. The main challenge is to find a very small number of useful features that will be able to complement each other to achieve the classifier's effectiveness. In order to support this task, a weak learning algorithm is used, the purpose of which is to correctly detect the single rectangle feature in order to optimally separate positive and negative samples. For each feature, the optimal threshold classification function is determined, such as the minimum number of samples that have been incorrectly classified. Weak classifier $h_j(x)$ therefore consists of the feature f_j , and the cut-off point of the feature θ_j and a parity p_j indicating the direction of inequality [9]:

$$h_j(x) = \begin{cases} 1 & p_j f_j(x) < p_j \theta_j \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

In this case x is a 20×20 pixel sub-window of an image. A classifier with such a structure enables fast detection of cars, because decisions are first made at the lowest levels, which leads to initial selection and leaving data meeting the assumed criteria.

Detection of vehicles based on images obtained from a low altitude is easier in relation to the detection of vehicles based on oblique images or characterized by low spatial resolution. However, it can be difficult to still correctly detect a car on a sequence of images. The main difficulties are as follows [10]:

1. Too low resolution organic images correct car detection;
2. Some cars may be obscured by the edges of trees;
3. The color of cars may be very bright or very dark - this effect has a negative effect on the correct detection of cars;
4. The presence of shadows can effectively disrupt the correct detection of cars;
5. The heterogeneous radiometric quality of images - caused by differences in lighting and other atmospheric factors can effectively reduce the possibility of correct car detection;
6. Expected car characteristics may be falsified by the occurrence of shadows.

Image data from the low level were obtained using the UX5 platform, which can be classified in the mini UAV category (Fig. 2).



Fig. 2. Trimble UX-5 platform before launch

The fixed-wing allows for a completely autonomous flight at a given altitude and with longitudinal and transverse coverage of acquired images. The system equipment includes:

1. Fixed-wing UX-5;
2. Controller enabling flight management;
3. UAV launcher.

The UX-5 platform allows you to perform an autonomous flight along a given route and to obtain high resolution digital images or video sequences. Pictures are acquired using an automatic infrared trigger. Flight safety is controlled automatically, but the operator intervention and the initiation of emergency safety procedures is possible at any time [11]. Table 1 presents the basic parameters of the platform.

The system is operable for wind speed not exceeding 18 m/s and in weather conditions not worse than a light rain. The images can be captured from an altitude ranging from 75 to 750m with ground resolution from 2.4 to 24 cm. The image data were obtained by Sony NEX-5R digital camera, which is one of the most commonly selected sensors mounted on board unmanned aerial cars [12; 13].

Table 1

Technical parameters of Trimble UX-5

Type	Body with the wings
Weight	2.5 kg
Wing span	1 m
Wing surface	34 dm ²
Dimensions	100 x 65 x 10 cm
Motor	Electric motor – 700W
Battery	14.8 V, 6000 mAh

Video sequences are taken with a super bright Voigtlander lens with 15 mm focal length and the maximum aperture equal F4.5. The video sequences were recorded in the AVCHD format with a resolution of 1920 x 1080 pixels and H.264 compression and at a rate of 25 fps.



Fig. 3. Trimble UX-5 platform before launch

The selected test area was around the town of Owsianka located 60 kilometers north-east of Warsaw (Fig. 3). The image data from the low altitude was obtained taking into account the height of the coating and the properties of the lens. The test video material was obtained from a 100 m altitude with a lens angle of 110°. A complete collection of samples [14; 15] developed on the basis of 526 car images was used for analyzes. The images were without scales. The size of the rectangular filter was 20x20 pixels. The simplified algorithm for car detection is presented below:

Algorithm 1

Car tracking with Haar Feature-based Cascade Classifier

```

import
cascade source = 'cars.xml'
video source = 'video1.avi'
while True:
frame_n = video.read()
if (type(frame_n) == type(None));
break
    for frame_n
        Converse each frame_n from RGB to GRAY
for
do detect cars use MultiScale Cascade detect
do
detect potential region of interest
use filter detected regions
if it is a new region, add to the collection
do count = count+1
end

```

The classifier training consisted of preparing several hundred different samples containing the selected object – a car. The object on each sample should be made up of approximately the same number of pixels. After training the classifier, it was possible to analyze the selected image sequences.

Results and discussion

The analysis involved 200 selected video frames from video material lasting 2:04 minutes, with processing 25 frames per second for 1920×1080 videos on average. The selection criteria took into account that there should be at least one car on each image frame. The maximum number of vehicles that could be registered in a single image was seven. Next, five representative sets of data containing 40 video frames were created. Table 1 shows the results of vehicle (car) detection effectiveness on individual sequences of images.

Table 2

Effectiveness of car detection on image sequences

Dataset	Set_1	Set_2	Set_3	Set_4	Set_5
Detected cars, %	64.9	65.2	62.7	63.8	63.4

For the first set of images, the average number of detected cars is below 3, and the median is equal to 3. The minimum number of detected cars is 1, and the maximum is equal to 5 cars. In turn, the number of incorrectly classified cars on the video frames of set No. 1 is from 0 to 4. The average value of erroneously classified cars (missed cars) on the video frames of set No. 1 is less than 2, while the median is equal to 1. Based on the result of the image set No. 1, it can be stated that the effectiveness presented in the work of the research method is 64.9 %. Therefore, the total value of misclassified cars on the video frames of set No. 1 is 35.1 %.

For the image sequence No. 2, the average number of detected cars is above 3, and the median is equal to 3. The minimum number of detected cars is 1, and the maximum is equal to 5 cars. In turn, the number of incorrectly classified cars on the video frames of set No. 2 is from 0 to 4. The average value of missed cars on the frames of video set 2 is less than 2, while the median is equal to 2. Based on the result of the sets of pictures No. 2 it can be stated that the effectiveness presented in the work of the research method is 65.2 %. Therefore, the total value of misclassified cars on video frames of set No. 2 is 34.8 %.

For the image sequence No. 3, the average number of detected cars is below 3, and the median is equal to 3. The minimum number of detected cars is 1, and the maximum is equal to 5 cars. In turn, the number of incorrectly classified cars on the video frames of set No. 3 is from 0 to 4. The average value of missed cars on the video frames of set No. 3 is less than 2, while the median is equal to 1. Based on the results of the image sets No. 3, it can be concluded that the effectiveness presented in the work of the research method is 62.7 %. Therefore, the total value of misclassified cars on the video frames of set No. 3 is 37.3 %.

For the image sequence No. 4, the average number of detected cars is below 3, and the median is equal to 3. The minimum number of detected cars is 1, and the maximum is equal to 5 cars. In turn, the number of incorrectly classified cars on the video frames of set No. 4 is from 0 to 4. The average value of missed cars on the video frames of set No. 4 is less than 2, while the median is equal to 2. Based on the result of the sets of images No. 4, it can be concluded that the effectiveness presented in the work of the research method is 63.8 %. Therefore, the total value of misclassified cars on the video frames of set No. 4 is 36.2 %.

For the picture sequence No. 5 the average number of detected cars is below 3, and the median is equal to 3. The minimum number of detected cars is 1, and the maximum is equal to 5 cars. In turn, the number of incorrectly classified cars on the video frames of set No. 5 is from 0 to 4. The average value of missed cars on video frames of set No. 5 is less than 2, while the median is equal to 1. Based on the result of the sets of images No. 5, it can be stated that the effectiveness presented in the work of the research method is 63.4 %. Therefore, the total value of misclassified cars on video frames of set No. 5 is 36.6 %.

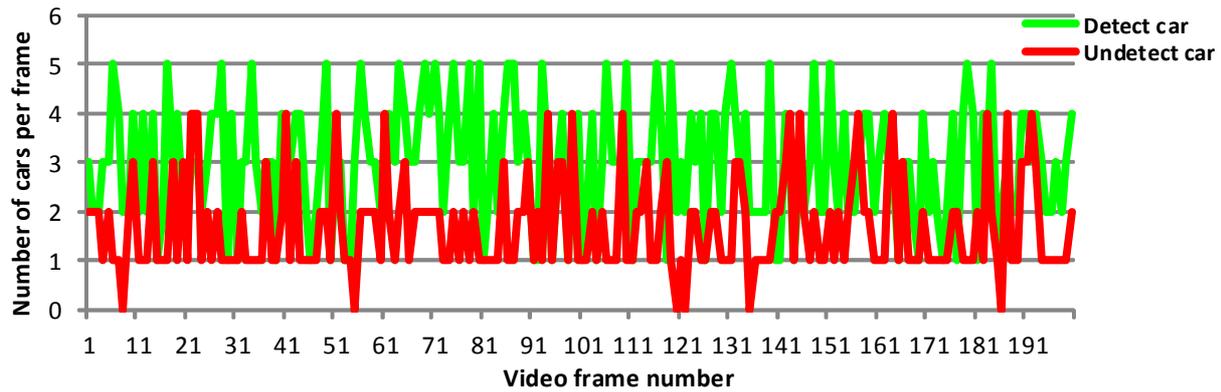


Fig. 4. Effectiveness of car detection on individual video frames

For the whole image sample, the average number of detected cars and their median is 3. The minimum number of detected cars is 1, and the maximum is equal to 5 cars. The average value of missed cars on video frames is less than 2, while the median is equal to 1. On the basis of the whole sample of data, it can be concluded that the effectiveness presented in the work of the research method is 64.0 %. The figure below (Fig. 5) shows an exemplary video frame, on which cars were detected.



Fig. 5. Result of cars detected using Haar Feature-based Cascade Classifier

In the case of video data, observation of possible car movement from a height of 100 m (Fig. 5) allowed to obtain a field size of 35 to 61 m on one frame of the video image, respectively. The height of observation should enable conducting qualitative and quantitative analyzes of cars currently on the road. The set height of observation should ensure the registration of traffic within the entire roadway. The resolution should also allow identification of individual car details in the image (Fig. 4). This is important for correct classification of the object during automatic detection of video sequences. On individual image frames made from 100 m altitude (Fig. 4), the type of car and its color can be distinguished. In the case of automatic tracking of an object on video sequences, it is required that the tracked car be visualized using at least tens of pixels.

Slightly more accurate results were obtained in [15], where the resolution of the applied samples was 752 x 512 pixels. For the detection of cars in low resolution aerial images [10], an efficiency of 80 % was achieved using the Bayesian Neural Networks. In another research done by Hen et al. [16] the Haar-like method and also additional edge detection methods (Sobel and Prewitt) were used to detect cars. Detecting rate ranged from 76.2 to 92.1 %. When comparing the results of studies of other authors, it can be concluded that the obtained results from the presented experiment are similar. The limitation of the proposed solution is the fact that in some cases the results of the analysis may be very unstable, especially when the acquired images were oblique (off-nadir).

Conclusions

The paper presents the results of the research on the use of Haar Feature-based Cascade Classifier for the detection of cars based on image sequences acquired from a low altitude. A ready set of samples developed on the basis of 526 car images was used for the analysis. The proposed solution is based on an algorithm available in the OpenCV library for the Python environment. On the basis of

five test data sets it was possible to perform a correctness analysis of car detection. For test data sets, the detection efficiency was in the range of 62.7 % to 65.2 %. The obtained results of detection could be better, if an extended set of test data were used that served to train the classifier. The effectiveness of car detection based on a ready classifier is 64 %. The limitation of the proposed solution based on image sequences is the fact that in some cases the results of the analysis may be very unstable, especially when the acquired images were oblique (off-nadir). In this case, it was justified to use a classifier that is resistant to change of scale. However, it is not immune to the occurrence of rotation. Another limitation is the blurring of images as well as acquiring them in changing weather and lighting conditions. Future research will focus on improving the efficiency of the algorithm and developing its own set of samples to increase the effectiveness of car detection.

In conclusion, the proposed method of car detection can be successfully used to improve the road safety and increase the traffic flow in rural areas. The presented research results also show that UAVs equipped with compact digital cameras find an increasingly wider application in car monitoring and land transport.

Acknowledgements

Many thanks to Prof. Michal Kedzierski for his great help in acquiring data.

References

- [1] Fajczak-Kowalska Krenski C. Linear infrastructure of Polish motorways and national roads, its condition and influence on the safety of road traffic and road haulage „Logistyka”, 4/2014, pp. 265-271.
- [2] Zhou H., Kong H., Wei L., Creighton D., Nahavandi S. Efficient road detection and tracking for unmanned aerial vehicle. *IEEE transactions on intelligent transportation systems*, 16(1), 2015, pp. 297-309.
- [3] Kedzierski M., Fryskowska A., Wierzbicki D., Nerc P. Chosen aspects of the production of the basic map using UAV imagery. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 41, 873, 2016.
- [4] Zhou H., Creighton D., Wei L., Gao D. Y., Nahavandi S. Video driven traffic modelling. In “Advanced Intelligent Mechatronics (AIM), 2013 IEEE/ASME International Conference”. IEEE, 2013, pp. 506-511.
- [5] Colla S., Manesis C. Intergraph video and images exploitation capabilities. In “First International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2013)” (Vol. 8795, p. 87950A). International Society for Optics and Photonics, 2013.
- [6] Wang H. R., Zheng X. Q., Sun Y. B., Jia Z. R., Wang H. Z., Wang, H. Researching on the process of remote sensing video imagery. “In PIAGENG 2013: Intelligent Information, Control, and Communication Technology for Agricultural Engineering” (Vol. 8762, p. 87621D). International Society for Optics and Photonics, 2013.
- [7] Viola P., Jones M. Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition, “CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on (Vol. 1, pp. I-I)”*. IEEE, 2001.
- [8] Lienhart R., Maydt J. An extended set of haar-like features for rapid object detection. “In *Image Processing. 2002. Proceedings. 2002 International Conference on (Vol. 1, pp. I-I)”*. IEEE, 2002.
- [9] Menezes P., Barreto J. C., Dias J. Face tracking based on haar-like features and eigenfaces. *IFAC Proceedings Volumes*, 37(8), 2004, pp.304-309.
- [10] Zhao T., Nevatia R. Car detection in low resolution aerial images. *Image and Vision Computing*, 21(8), 2003, pp. 693-703.
- [11] Wierzbicki D., Krasuski K. Estimation of rotation angles based on GPS data from a UX5 Platform, *Measurement Automation Monitoring*, vol. 61, no. 11, 2015, pp. 516-520.
- [12] Wierzbicki D., Kedzierski, M., Fryskowska, A. Assessment of the influence of UAV image quality on the orthophoto production. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40(1), 2015.

- [13] Wierzbicki D. , Fryskowska A. Radiometric quality assessment of video sequences acquired from UAV photogrammetric systems “Environmental Engineering” 10th International Conference Vilnius Gediminas Technical University Lithuania, 27-28 April 2017.
- [14] Sobral A. Vehicle Detection HaarCascades, 2014 [online][06.02.2018], Available at: https://github.com/andrewsobral/vehicle_detection_haarcascades
- [15] Oliveira M., Santos V. Automatic detection of cars in real roads using haar-like features. Department of Mechanical Engineering, University of Aveiro, 3810, 2008.
- [16] Han S., Han Y., Hahn H. Vehicle detection method using Haar-like feature on real time system. World Academy of Science, Engineering and Technology, 59, 2009, pp. 455-459.