Situation Control of Unmanned Aerial Vehicles for Road Traffic Monitoring

Nikolai Vladimirovich Kim[1] & Mikhail Alekseevich Chervonenkis[2]

[1] Moscow Aviation Institute, National Research University, Russian Federation

[2] PAWLIN Technologies Ltd, Dubna, Russian Federation

Correspondence: Mikhail Alekseevich Chervonenkis, PAWLIN Technologies Ltd, Dubna, Moscow, Russian Federation.

Abstract

This paper aims to introduce an approach to the organization of road traffic monitoring by the means of unmanned aerial vehicles (UAVs), which is based on the automatic situation control of UAVs. The research includes analysis of existing methods of on-board automatic detection of emergency and abnormal traffic situations with UAV artificial vision systems (AVS), preliminary classification of these situations including the allocation of emergencies and disastrous situations. The paper presents the choice of UAV controls in compliance with the recognized situation. The traffic situation identification method introduced in the paper is based on Bayes and Neyman-Pearson criterion. Furthermore, the research involves the analysis of the existing approaches to the detection of moving and stationary vehicles by the means of UAV AVS. The paper proposes vehicles detection method based on the image segmentation, along with the use of machine learning methods, particularly the artificial neural network method known as Deep Learning. The research provides solutions for vehicle tracking and velocity detection problems in order to describe traffic situations. The proposed approach contributes to the efficiency of UAV in road traffic monitoring by means of the management and detection processes automation.

Keywords: automatic situation control, unmanned aerial vehicles (UAV), vehicle detection, road traffic monitoring, Bayesian framework, Neyman-Pearson Criterion, deep learning, image segmentation, artificial vision systems (AVS)

1. Introduction

The paper represents the way of organizing road traffic control by an operator on the basis of receiving and analysing incoming video information from various observing systems and, in particular, those set up on unmanned aerial vehicles (UAVs).

The goal of this monitoring is to increase traffic capacity and safety of the controlled road sections by timely:

• Detection of the movement obstacles in the pre-jam and jam stage, as well as other abnormal situations, including emergencies, and other traffic hindrances;

• Measures aimed at preserving peoples' health and wellbeing, the state protection objects, etc.;

• Elimination of abnormal situations consequences;

• Prevention of road accidents and so forth.

Abnormal situations detection and classification are required for an operator to determine the means of consequences elimination for these particular situations. The implementation of these processes implies high information workload for an operator, which can lead to the delays in the required decision-making and cause operator errors.

The research proposes an automation approach to the following processes:

1. Abnormal situations detection;

2. The preliminary classification of abnormal situations;

3. Collecting video data corresponding to a predetermined class, which is necessary for an operator to make

decisions.

The observation system, which is responsible for implementing the functions above on board of an UAV, belongs to the artificial vision systems (AVS).

The traffic monitoring organization scenario under consideration comprises four steps:

1. In normal mode, UAV AVS performs video monitoring of a specific road section (patrolling), detecting the vehicles, and estimating their movement characteristics;

2. In case of an abnormal situation, UAV AVS should detect it and notify the operator;

3. After locating the abnormal situation, UAV AVS should perform its preliminary classification;

4. In accordance with the situation class, UAV AVS collects and transmits the information, which is necessary for an operator to make decisions.

It is obvious that depending on the characteristics and severity of the predicted consequences of abnormal situations, an operator should plan and take a range of appropriate measures, e.g., call emergency services (or road safety service). For this purpose, it is necessary to provide the operator with video information, allowing him to analyse various facts related to this special situation with precision. In some cases, such information cannot be obtained by the UAV while patrolling.

Thus, according to the suggested monitoring scenario, it is necessary to implement situation control of UAV at step 4, which will allow calculating various trajectories depending on the specific road traffic situations (classes of abnormal situations).

2. Methodology

2.1 Vehicle Recognition

Detection (recognition) of vehicles is crucial for traffic situations recognition. It requires solving the problem of vehicles location and speed determining based on photo or video capture using UAVs. There are two basic approaches to vehicle detection: either detecting moving vehicles either on video or on selected frames irrespective of their speed. The methods of moving vehicles detection are more accurate (Gavrilov & Lej, 2013). For some abnormal situations it is enough to detect only moving vehicles, but to recognize a wide range of abnormal situations (such as congestion (traffic jams), consequences of road accidents, lack of active traffic etc.) it is also important to detect motionless or slowly moving objects.

Vehicle detection methods based on boundaries detection (such as Canny filter (Canny, 1986) are not always successful, since their use can lead to false activations and combining vehicle adumbrations with shadows, road cracks, road-side, road marking, other vehicles etc. Adumbration can be unstable and certain aspects of a vehicle are impossible to detect. The methods of segmentation and/or marking features on image areas are more appropriate. Below is the vehicle detection scheme for this case:

1) Image scaling and determination of the region of interest (ROI), usually one or more rectangles including the road and the roadside. In this case, navigation data are used.

2) ROI image segmentation.

3) Primary segments filtration (in particular, rejection of very large or very small segments) and combining several adjacent segments, as well as receiving a set of regions.

4) Region filtration.

5) Region descriptors creation.

6) Recognition of vehicles based on region descriptors.

7) Recognition of vehicles based on the velocity fields.

The segmentation results in having the processed images divided into a number of connected non-overlapping areas - segments:

$$X = \bigcup_i S_i, \tag{1}$$

where X – multitude of segmented image pixels, S_i – segment, i – integer index, and $S_i \cap S_j = \emptyset$, as $i \neq j$.

Segmentation is performed in such a way that a separate segment is homogenous enough, but different and adjacent ones vary in colour or brightness characteristics. The following methods of segmentation were investigated: Felzenszwalb-Huttenlocher Segmentation (Felzenszwalb & Huttenlocher, 2004), Quick Shift (Vedaldi & Soatto, 2008). Simple Linear Iterative Clustering (SLIC) (Achanta et al., 2012), Maximally Stable

Extremal Regions (MSER) (Matas et al., 2004), Model based clustering – MBC (Zhong & Ghosh, 2003). Currently, (Meuel et al., 2013) among those listed only FHS method provides the speed of response required for the on-board processing. Moreover, FHS method has the best quality of segmentation and the ability to combine large homogeneous fields into one region (this, for example, allows to combine most of the road into one segment rather than in many similar segments). FHS method combines all small segments (smaller than the threshold value) with adjacent large segments. As most of the methods listed, FHS uses the colour information.

In (Meuel et al., 2013) and in our study it is obvious that when using the FHS the image of a vehicle is usually sufficiently scaled up by using one, two, or three segments. Therefore, for vehicle detection all individual segments, merger of pairs of adjacent segments and merge of segment triples in which at least one is adjacent to the other two, are considered. The resulting set of points (further referred to as regions) is filtered according to the limited (top and bottom) area and a number of other geometrical parameters - area, occupation, size and aspect ratio of the minimum enclosing rotated rectangle (there is an efficient algorithm for constructing such rectangle) (Pirzadeh, 1998).

At this stage, the problem of vehicle image recognition is reduced to the binary classification of regions. To solve this problem, a set of features for each segment is determined and is used as the basis for the classification. The recognition method is based on machine learning.

Several works propose the construction of features by analysing geometric and textural properties of an object. In particular, in (Tuermer et al., 2011) modelling of vehicle features based on Histogram of Oriented Gradients, (HoG) (Dalal & Triggs, 2005) and Haar-feature (Lienhart & Maydt, 2002) is proposed, using AdaBoost as the recognition method (Polikar, 2006). HoG and Haar-feature are sensitive to rotation and therefore require multiplying copies of these features at the turning angle with a certain step. Another approach is to find a set of features inoptionoption to rotation. Fourier transformation by HoG can serve as such set.

However, recently algorithms synthesizing features by unsupervised learning have been replacing the methods of image object classification using features detected by deterministic algorithms (in particular HoG and Haar-feature) (Deng & Yu, 2014). We also offer a different approach to the problem of segment classification, based on applying the method of fibre-wise learning of multilayer auto-associative neural networks - Deep Learning (Bishop, 1995). The advantage of this method is that it is resistant to noise as well as to a wide range of distortion at a sufficiently representative sample.

The core of the method is that the original feature vector is transformed in a neural network cascade into a small dimension vector, and then this vector in turn is transformed back to the original vector in another neural network cascade. A specific learning scheme is applied to each class. During recognition, the class for which this transformation has been more accurate is selected.

Transformation of a region into a fixed set of features is required for the method mentioned above as well as for the most of other classification methods. To apply it, a segment has to be surrounded by the minimum enclosing rotated rectangle. An affine transformation transforms the given rectangle into a fixed one. The rectangle with an aspect ratio typical for a vehicle (f.eg. 2 : 5, the dimensions of the rectangle 80 x 200 pixels) will be used as a fixed rectangle with the size $H \times W$ (H and W are the height and the width of the selected rectangle). Image intensity values in the resulting rectangle are considered as a feature vector with the dimensions $C \times H \times W$, where C is the number of image channels (usually 3). The database of the region descriptors, which are consistent and admittedly inconsistent with vehicles, is created basing on the manually marked video records. Thus, the problem is reduced to the binary classification of the large dimension vectors, and is successfully solved by the Deep Learning method, provided the sampling is large enough. The last step of the vehicle detection is connecting the regions each of which is recognized as a vehicle. It should be done if these regions have a significant area of intersection (in relation to the area of their connection), but the combined region must meet top size constraints.

2.2 Detection of Moving Vehicles

There are many works, on moving vehicles detection, for example (Gavrilov & Lej, 2013; Meuel et al., 2013; Zhang et al., 2012). Proposed ways of detection are generally as follows:

I. Detecting static objects and tracking their movements.

II. Detecting and filtering areas or points of movement.

III. Modelling a super pixel (fulfilling segmentation of points in space and time) (Meuel et al., 2013). (In this paper, we will consider this modelling as it is a subject of a separate study).

Way I is essentially an object tracking method implementation of which is considered in the next paragraph. For the Way II we researched the following methods of movement detection:

  - Correlation method (finding the maximum of the correlation function between the adjacent frames).

  - Lucas-Kanade method of optical flow determining (Lucas & Kanade, 1981) based on the images pyramid and on a fixed grid.

  - Farnebäck method (developed by Gunnar Farnebäck (Farnebäck, 2001), based on quadratic (relative to coordinates and time) models of image intensity representation.

Lucas-Kanade method provides a possibility of velocity field modelling for each point on the image pyramid, but is less quick /in operation and less accurate than other methods. Lucas-Kanade method is efficient when applied to a large grid and is suitable for determining the total frames shift.

Farnebäck method performs well; it allows producing a list of moving objects and tracking them. Farnebäck method operates fast enough (e.g. circa120 FPS, single-core mode, processor Intel i7 3300 MHz, 1080x760) Correlation method performs very slowly when applied to a fine grid and does not ensure sufficient accuracy when applied to a wide grid.

Farnebäck method is adequate for moving vehicles detection.

All methods of traffic search are susceptible to false responses, which are less frequent on roads and more frequent on side-lines or in the bushes. Therefore, the results of moving objects detection should be filtered or combined with the results other methods.

Farnebäck method output is a set of points that are identified as moving. Particular regions can be defined for each point (not all of these regions correspond to real vehicles). The sets of such regions are classified based on learning in the same way as the regions of individual images are classified (see the previous section).

2.3 Tracking the Recognized Vehicles and Assessing Their Speed

To describe a traffic situation, it is important not only to recognize vehicles, but also to assess their speed, and, in some cases, their motion paths. Likewise, the highest accuracy of speed assessment is achieved by applying the algorithms of moving objects tracking.

The initial data for the tracking system are represented by the sets of descriptions of objects recognized in each frame (substantively the image regions of corresponding vehicles). A specific vehicle recognized in the current frame is described by a features vector $\vec{Z} = (z_1, z_2, ..., z_n, ..., z_L)$, where L is the total number of features. This feature set includes information about the object location, its geometric characteristics, and primary evaluation of its speed (based on the optical flows). Using vector $\vec{Z}$ makes it possible to make an overall assessment of the 2-D coordinates of the object $\vec{X}(\vec{Z})$ (for example, the centre of mass, or the centre of the enclosing rotated rectangle, both of which in practice may coincide with the two coordinates of $\vec{Z}$). The result of the frame processing is a set of descriptors TC - $U = (\vec{Z}_1, \vec{Z}_2, ..., \vec{Z}, ..., \vec{Z}_K)$, where K is the total number of vehicles detected /in the frame. For the frame number t a set of objects descriptors U_t is formed.

For recognized objects tracking, two monitoring levels are proposed:

1) The level of generating hypotheses of specific vehicles' movements;

2) The level of tracking specific vehicles based on Kalman filters (Welch & Bishop, 2001; Kleeman; Kelly, 1994).

Hypotheses are generated based on the f sequential frames analysis (wherein f is a small number, such as 3 and 4; f value is determined by the time during which objects acceleration can be neglected in the analysis of their movement). In these frames all the pairs of objects are selected; one of the objects is located in the current frame and the other in another frame, but at a limited distance (calculated on account of the information about limiting speed of the vehicle, image scale and the frame rate). The hypothesis is that the detected couple of objects belongs to the same vehicle (and with some precision the movement is straight and steady – at the last f - frames).

After that, the set of hypotheses produced is compared with the hypotheses adopted earlier to detect congruence.

If a hypothesis is not coherent with the previously adopted one, it is added to the list of hypotheses (*HL*) to which a Kalman filter is applied. To confirm or reject the hypotheses in the HL, the search for the best

confirmation of the Kalman filter forecast is used as well. If a hypothesis is not confirmed within f_l frames, it is deleted. The output of the algorithm at each step is a set $S = \left((\vec{Z}_1, \vec{V}_1), (\vec{Z}_2, \vec{V}_2), \dots, \dots, (\vec{Z}_R, \vec{V}_R) \right)$, where R is the number of Kalman filters, and $\vec{V}_i$ is the speed calculated by means of Kalman filter:

$$\vec{V}_i = \frac{\vec{F}_{Ki} - \vec{X}_i}{\Delta} , \tag{2}$$

where Δ is the time step beetween the frames, $i = \overline{1, R}$ is the object index, Ki - Kalman filter for i object, $\vec{X}_i$ – current position of i object, /and/ $\vec{F}_{Ki}$ – Kalman filter forecast for the next step of i object.

3. Results

3.1 Classes of Traffic Situations

Traffic situation classes' description should be based on the sets of vehicles descriptors recognized in the current and previous frames.

In the simplest cases, traffic situations are described by a set of static quantities (statistics), calculated because of a vehicle scalar speeds set. To convert the vector velocity into scalar form, it is possible to project the velocity on the road centreline (the scalar product of the velocity vectors and the directrix vector of the road centreline). In particular, the velocity distribution diagrams, the number of the vehicles, average (see section Detection of abnormal situations) and maximum speed, mean square deviations etc. can be selected as static quantities.

Figure 1 shows examples of vehicle velocity distribution histograms on a section of a road in various standard situations. The examples were obtained by means of traffic processes modelling.
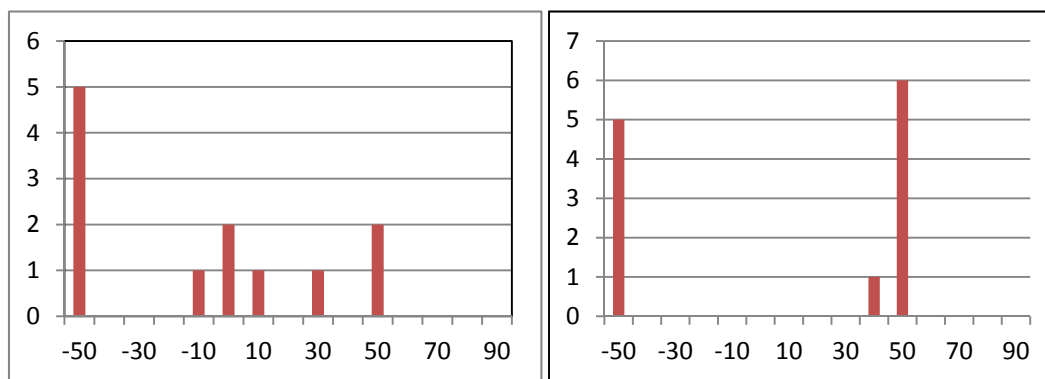
Figure 1. The first histogram shows an example of traffic with some vehicles moving in both directions at a normal speed, some at a low speed, and some being stationary. The second histogram shows a standard situation with two-way traffic (the number of vehicles is displayed on the vertical axis, the velocity in km/h – on the horizontal axis)

More complex optionoptions of descriptions shall include the location of vehicles in relation to one another and to the lines of traffic, the direction of the velocity vector in relation to the road centreline.

In general, the process of classification or recognition (or, in some cases, detection) within the scope of tasks solved by SVS, consists of specifying the class of the observed traffic situation on the basis of video information, and of relating the discovered set of features to the area characteristic to the corresponding class.

Let us denote the set of situation classes by X = {x_1 〚, x〛 _ (2), ... x_m, ..., x_M}, where M is the total number of situation classes.

Vectors Y = (y_1, y_2, ..., y_n, ..., y_N), characterizing different classes of situations in the feature space are referred to as implementation vectors, N is the total number of features used to describe the situation. The complete set of features used is called the feature dictionary.

By convention we assume that all traffic situations can be divided into five classes (M = 5) that comprise the source classes alphabet.

Herewith, class x_1 corresponds to a standard traffic situation where the traffic capacity of a specific road segment in a given season and at a given time of day or night lies within the specified tolerance limits.

We assume that classes x_2, x_3, x_4, x_5 refer to abnormal situations that push current values of the road traffic capacity outside the specified (for a normal situation) tolerance limits.

Class x_2 is an abnormal situation not resulting in direct financial losses, but disturbing the traffic capacity and pushing it outside the tolerance limits specified for a standard situation as well.

Abnormal situations that belong to the classes x_3, x_4 are emergencies, followed by financial losses including but not limited to vehicle damage of varying severity.

Situation x_5 is catastrophic and results in human losses.

Moreover, we assume that special situations x_2, x_3, x_4, x_5 occur due vehicles collisions and the higher the class of the situation, the more severe are the consequences.

For detection and classification of traffic situations we will use statistical technology of recognition.

In accordance with the above, when a particular situation (regardless of its class) occurs, it must be detected (step 2).

3.2 Abnormal Situations Detection

Let us denote a standard situation occurrence probability by $P(x_1)$ and an abnormal situation occurrence probability by $P(x_m)$ with index $m = \overline{2, M}$.

Note that at this point abnormal situations are not classified.

Object recognition task can be considered a special case of recognition, if the decision taken contains only two outcomes: whether an abnormal situation is detected or not. This approach simplifies the solution that allows applying various statistical detection criteria without calculating posterior probabilities by Bayes formula (Hazewinkel et al., 2001b).

To determine the decisive boundaries it is suitable to use likelihood factor (ratio) (Neyman and Pearson, 1933).

$$\lambda = p(Y|x_m)/p(Y|x_1) \tag{3}$$

and threshold (critical) likelihood factor

$$\lambda_0 = P(x_1)(R_{m1} - R_{11})/P(x_m)(R_{1m} - R_{mm}), \tag{4}$$

where $\boldsymbol{P(x_n)}$ – a priori probability of the situation $\boldsymbol{x_n}$; $\boldsymbol{P(Y|x_n)}$ – conditional probability density of the feature $\boldsymbol{Y}$, upon condition that the source of information is the situation $\boldsymbol{x_n}$; $\boldsymbol{n = \overline{1, M}}$ – situation classes indices; $\boldsymbol{R_{mn}}$- losses caused by identification errors (situation $\boldsymbol{x_n}$ identified as $\boldsymbol{x_m}$).

The decision as to whether a situation is standard or an abnormal is made based on correlation $\boldsymbol{\lambda}$ between $\boldsymbol{\lambda_0}$.

Assuming that the losses in case of a right decision are

$$R_{11} = R_{mm} = 0,$$

In addition, in given (1), (2) the condition (*Bayes criterion*) is obtained:

if

$$\lambda = p(Y|x_m)/p(Y|x_1) > \lambda_0 = P(x_1)(R_{m1})/P(x_m)(R_{1m}), \tag{5}$$

then the situation observed is abnormal $X = x_m$,

else the situation observed is standard $X = x_1$.

For the detection process, it is important to consider the errors of object recognition when an abnormal situation is not identified (α – type one error) or mistakenly identified (β – type two error).

If there is a possibility of one of the abnormal situation identification errors leading to unacceptable consequences, Neyman-Pearson criterion must be used to limit the tolerable values of these errors.

According to the criterion, the detection algorithm conditions are the following:

$$\min \alpha \text{ provided } \beta \leq \beta_0 \text{ or } \min \beta \text{ provided } \alpha \leq \alpha_0, \tag{6}$$

where α_0, β_0 - given maximum limits for type one and type two errors respectively.

Likelihood factor λ is compared with a threshold value λ_0 calculated from /the/ formulae (Neyman and Pearson, 1933):

$$\alpha_0 = \int_{\lambda_0}^{\infty} p(\lambda|x_m)d\lambda \quad \text{or} \quad \beta_0 = \int_{-\infty}^{\lambda_0} p(\lambda|x_m)d\lambda \ .$$

Let us assume that we take average speed of vehicles measured per second as a feature while detecting abnormal situations:

$$Y = \frac{1}{K}\sum_{k=1}^{K} V_k,$$

where K – the number of vehicles, V_k – the speed of vehicle k.

To simplify the calculations, we also assume that feature distribution for each - n = 1, m-th situation follows the normal law (Bryc, 1995)

$$p(y|x_n) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{[y-M(y|x_n)]^2}{2\sigma^2}}. \tag{7}$$

Let us consider the examples of dangerous situations detection based on Bayes and Neumann-Pearson criteria.

For a specific controlled road section (discarding the feature measurement - km/h), it is assumed that

$$E(y|x_1) = 60, \ \sigma(y|x_1) = 15,$$
$$E(y|x_m) = 20, \ \sigma(y|x_m) = 5,$$

here and hereafter $E(|x_m)$ – mathematical expectation of y, and $\sigma(y|x_m)$ – conditional root-mean-square (RMS) deviation of y, $m = \overline{1,M}$.

Option 1. Bayes criterion

Priori probabilities:

$$P(x_1) = 0.99; \ P(x_m) = 0.01.$$

Tolerable loss: $R_{m1} = 0.9; \ R_{1m} = 0.1.$

Option 2. Bayes criterion

$$P(x_1) = 0.99; \ P(x_m) = 0.01; R_{m1} = 0.99; \ R_{1m} = 0.01 \ .$$

Option 3. Neumann-Pirson criterion (Neyman and Pearson, 1933)

Tolerable error (failure to identify an abnormal situation (AS))

$$\beta_0 = 0.001.$$

Let us assume that the following values of the feature (average speed) were taken:

20, 30, 40, 50.

It is required to determine what speed corresponds to an abnormal situation (AS), and what to a standard situation (SS).

The computational results on the ground of (3), (4) are represented in Table 1.

Table 1. Computational results

Speed	20	30	40	50
ption 1	AS	AS	AS	SS
Option 2	AS	AS	SS	SS
Option 3	AS	AS	AS	SS

Thus, the results of the decisions made are largely dependent on the prescribed initial data. For example, on reducing the assigned losses caused by the failure to identify AS (Option 1 in relation to Option 2) at the average speed of 40 km/h, it is decided that the situation is standard (Option 2), but in Option 1 the situation is identified as an abnormal.

Similar results can be obtained by /using/ various algorithms. Option 3 shows that Neumann-Pearson criterion can provide the same solutions as Bayes criterion but without loss evaluation R_{mn}.

After detecting an abnormal situation it is necessary to classify the situation at the next step of the scenario (step 3).

4. Discussion and Conclusions

4.1 Abnormal Situations Classification

Let us assume that a collision of 2 vehicles occurred in vision of UAV AVS, i.e. the AS development was observed in dynamics. The collision condition is:

$$D_{qr} = 0,$$

where D_{qr} – the distance between the vehicles with indices q, r.

In such case, the closing speed (dimension - m/s) before the collision can be taken as a feature characterizing AS classes

$$Y = \dot{D}_{qr}.$$

Connection between the AS class and closing speed is random in character, so statistical methods will be applied to AS classification (recognition) as well as to their detection.

In general, recognition or classification (including those of AS) in statistical methods are based on the calculation of the posterior probability by Bayes formula (Hazewinkel et al., 2001b)

$$P(x_m|Y) = \frac{P(x_m)P(Y|x_m)}{\sum_{k=2}^{M} P(x_k)P(Y|x_k)}, \tag{8}$$

where $P(x_m), P(x_k)$ – priori probabilities of situations x_m, x_k; $P(Y|x_m)$ – conditional probability density of the feature vector Y, upon condition that the source of information is the situation x_m; $m, k = 2,3,4, M$ – situation classes indices; $M = 5$.

In particular, the simplest means of classification is by using the ideal observer criterion. We also accept as a condition that all the losses caused by identification failures are equal. The method is based on hypothesis selection corresponding to the maximum posteriori probability $P(x_m|Y)$.

Let us consider the AS classification procedure as applied to the traffic monitoring task.

We assume that $P(Y|x_m)$ is known, Y – an univariate vector ($N = 1, Y = (y_1), suppose\ y = y_1$)

Assume that feature distribution for class x_2 follows the exponential law and conditional probability density is calculated by the formula (Hazewinkel et al., 2001a)

$$p(Y|x_2) = p(y|x_2) = \begin{cases} \lambda e^{-\lambda y}, y \geq 0, \\ 0, y < 0. \end{cases}$$

Feature distribution for classes x_3, x_4, x_5 corresponds to the formula (5) (normal law).

Let us assume the following initial data:

parameter $\lambda = 4$ (expert analysis for the class x_2);

for classes x_3, x_4, x_5:

$$E(y|x_3) = 6,\ \sigma(y|x_3) = 1.5,$$
$$E(y|x_4) = 12,\ \sigma(y|x_4) = 2.3,$$
$$E(y|x_5) = 25,\ \sigma(y|x_5) = 3.5.$$

Probability densities plots are represented in Figure 2 (vertical axis - densities, horizontal axis - closing speed in m / s).
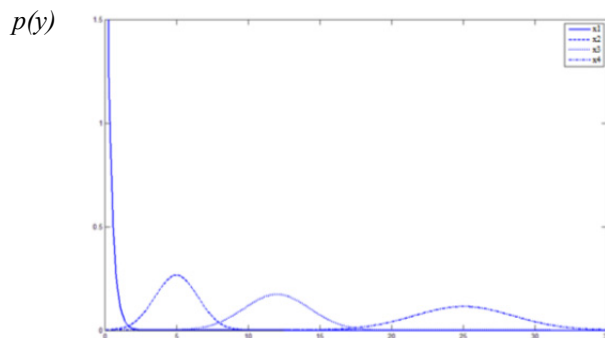
Figure 2. Vehicles closing speeds probability densities for different situation classes

Figure 2 shows /indicates/ 4 probability densities descriptive of different AS classes. Herewith, the probability density plots on the y-axis intervals (0-2.5); (5-10); (15-20) are mutually intersected.

These intersections are the most difficult for AS to be classified. However, densities duality allows (in this example) applying detection algorithms (with two possible outcomes).

Detection is not necessary for the remaining areas of the graph, so the following rules are applied:

1. If $2.5 \leq y \leq 5$, then $X = x_3$;
2. If $10 \leq y \leq 15$, then $X = x_4$;
3. If $20 \leq y$, then $X = x_5$.

Consider two examples of abnormal situations classification (Figure 3 –AS on the straight-line section, Figure 4 –AS on the intersection).

a) Initial situation   b) Collision

Figure 3. AS development on the straight-line section

Figure 3 demonstrates two (non-adjacent) frames of a video sequence, registering the development of an AS.

In Figure 3, a) a motor car (MC) marked by square 1 was moving in the left lane. The van (V) marked by square 2 started left lane change/.

In the other frame (Figure 3 b) we can see the collision, which results in the motor car going to the roadside. The closing speed (based on the auto tracking results analysis) by the time of collision was equal

$$Y = 1.25$$

Let us assume as an example:

- Option 1

$$P(x_2) = 0.1; \ P(x_3) = 0.9; R_{m1} = R_{1m} = 1$$
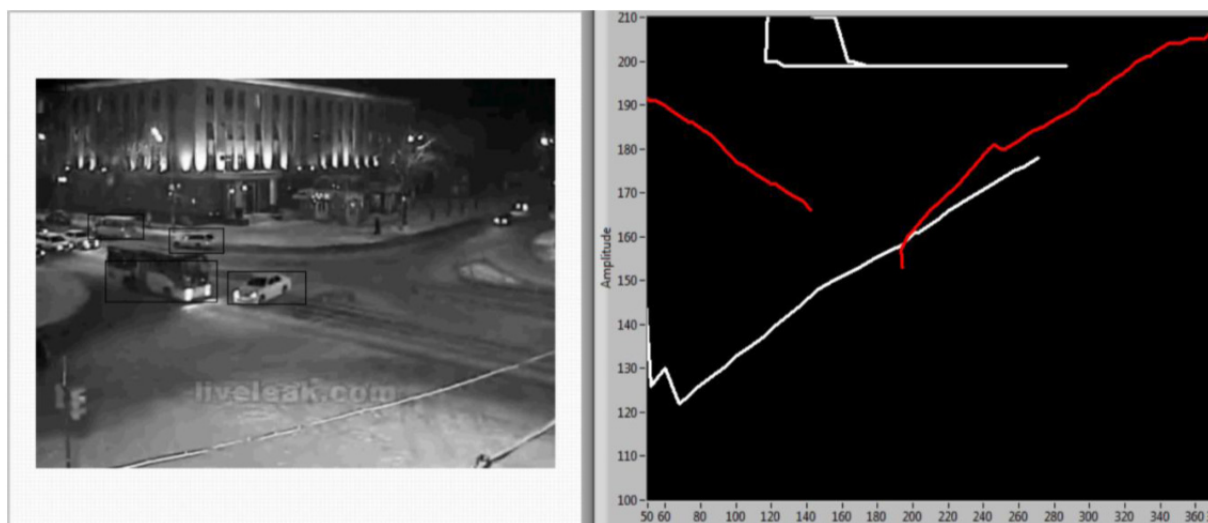
and in accordance with Bayes criterion,

$$\lambda_0 = 0.111$$

- Option 2

$$P(x_2) = 0,8; \ P(x_3) = 0.2; R_{m1} = R_{1m} = 1; \ \lambda_0 = 0.25.$$

Hence, it is decided that for Option 1 resulting characteristic value corresponds to the class x_3, and for Option 2 – to the class x_2.

Figure 4 a) demonstrates an AS where a motor car and a bus collide at the intersection, and b) depicts the vehicles motion paths.

a). The frame before the collision   b). Selected vehicles motion paths

Figure 4. The vehicles collision at the crossroads

The paths of the vehicles that passed the intersection before the AS are highlighted in white, whereas the collision participants are highlighted in red.

The calculated closing speed was equal to $Y = 12 \ м/с$.

In accordance with the probability densities plots (Figure 2), this characteristic value falls within the range corresponding to class x_3 and requires no further classification.

4.2 Flight Control Selection in Accordance with the Situations Class

Generally, the total scope of video information about any AS can be obtained if UAV AVS moves over the surface of the hemisphere covering the AS area. However, such surveillance of the region of interest is, in practice, impossible and redundant.

Since the location of the observed scene elements that can provide an operator with useful information is not known in advance, it is not possible to select optimal UAV paths.

Therefore, it can be assumed that the scope of information that an operator needs to make reasoned decisions correlates with the complexity of the AS class.

Let us select a surveillance strategy based on expert analysis for each AS class.

Denote UAV control activities by - S (m), where m – the number of the situation class.

$S(1 - 2)$ Assume that for classes x_1, x_2, an operator has enough information received by the UAV AVS in the patrolling mode.

$S(3)$. Class x_3 requires UAV flying over the AS area (UAV circling flight) at the patrolling altitude.

$S(4)$. Class x_4 requires UAV flying over the AS area at various altitudes (circling flight at the patrolling altitude and circling flight at minimum flight altitude (for the given conditions), e.g., 50 m.

$S(5)$. Class x_5 requires UAV helical ascent flight over the AS area. The starting segment of the helical path coincides with the patrolling path, and its end is determined by the minimum flight altitude.

Figure 5 shows an example of UAV flight path during the AS of class x_5 surveillance.

The AS location is indicated by a cross. Path patrolling is a rectangle located along the controlled road section, 300 m width. UAV flight altitude is 400 m. When approaching the place of the AS, UAV starts to drift down, shifting to a helical path (a constant banked turn).

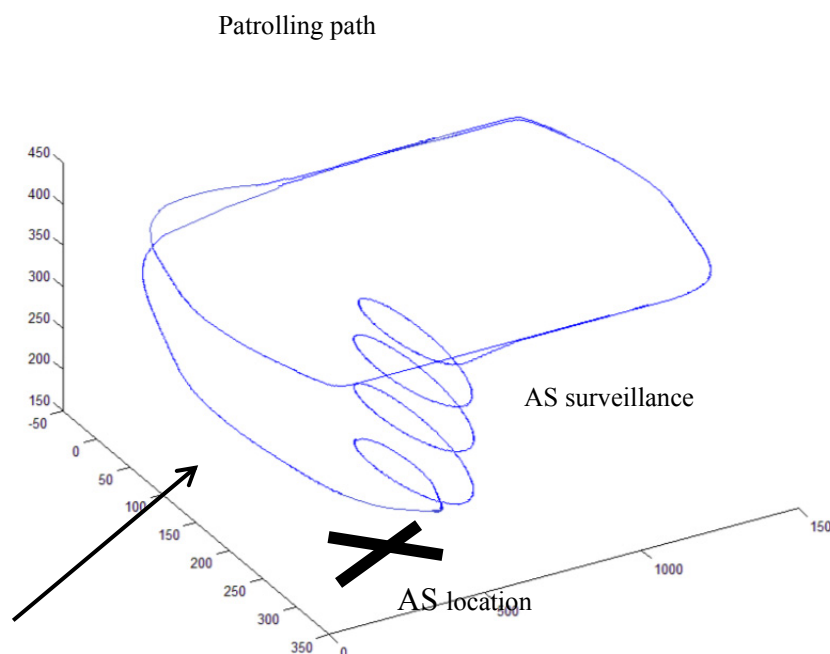Parameters in the example: helix circumference - 300 m, propeller pitch - 60 m.

Figure 5. UAV flight path in an AS monitoring mode

Because of building such path, it may be assumed that the scene (the abnormal situation location) is observed from the surface of the sphere sector. In this case, at the top of the helix UAV will be situated on the generatrix of zenith angle ($\alpha = 20,55$ gr.).

Such flight path of an UAV allows collecting enough information about an AS for an operator to make further decisions.

A new approach to organizing road traffic monitoring by means of UAV artificial vision systems has been proposed.

The approach proposed is based on the observation processes automation that allows performing UAV control in compliance with prevailing road traffic situations.

The traffic situation identification methods are defined.

The application of the proposed approach will contribute to the road traffic monitoring efficiency by means of UAVs.

Future research will focus on the experimental study of the proposed methods.

Acknowledgment

References

Achanta, R. et al. (2012). *Implementation of the SLIC super pixel algorithm to work with OpenCV.* Retrieved from http://www.github.com/PSMM/SLIC-Superpixels

Bishop, C. (1995). Neural networks for pattern recognition. Oxford: Clarendon Press.

Bryc, W. (1995). *The normal distribution: Characterizations with applications*. New York: Springer-Verlag.

Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence,* 679-698.

Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* 1063-6919/05.

Deng, L., & Yu, D. (2014). Deep learning: Methods and applications. Now Publishers.

Farnebäck, G. (2001). Very high accuracy velocity estimation using orientation tensors, parametric motion, and simultaneous segmentation of the motion field. *Proceedings of the Eighth IEEE International Conference on Computer Vision* (Vol. I., pp. 171-177). Vancouver, Canada.

Felzenszwalb, P., & Huttenlocher, D. (2004). Efficient graph-based image segmentation. *International Journal of Computer Vision,* 167-181. http://dx.doi.org/10.1023/B:VISI.0000022288.19776.77

Gavrilov, A., & Lej, S. (2013). The study of methods and algorithms for detecting moving objects in the video stream. Moscow. Bauman Moscow State Technical University. Retrieved from www.sntbul.bmstu.ru/file/out/568767

Hazewinkel, M., ed. (2001a). Exponential distribution, In *Encyclopedia of mathematics*, Dordrecht: Springer.

Hazewinkel, M., ed. (2001b). Bayes formula. In *Encyclopedia of mathematics*. Dordrecht: Springer.

Hormoz Pirzadeh's homepage. Rotating calipers: Minimum area enclosing rectangle. (1998, December 1). Retrieved from http://cgm.cs.mcgill.ca/~orm/maer.html

Kelly, A. (1994). *A 3D state space formulation of a navigation Kalman filter for autonomous vehicles*. Pittsburgh, Pa.: Carnegie Mellon University, the Robotics Institute.

Kleeman, L. (1995). *Understanding and Applying Kalman Filtering*. Department of Electrical and Computer Systems Engineering, Monash University, Clayton.

Lienhart, R., & Maydt, J. (2002). An extended set of haar-like features for rapid object detection. *International Conference on Image Processing* – ICIP*, 1*, 900-903. http://dx.doi.org/10.1109/ICIP.2002.1038171

Lucas, B. D., & Kanade, T. (1981). An iterative image registration technique with an application to stereo vision. *Proceedings of Imaging Understanding Workshop,* 121-130.

Matas, J., Chum, O., Urban, M., & Pajdla, T. (2004). Robust wide-baseline stereo from maximally stable extremal regions. Image and Vision Computing, 761-767.

Meuel, H., Reso, M., Jachalsky, J., & Ostermann, J. (2013). Superpixel-based segmentation of moving objects for low bitrate ROI coding systems. Advanced Video and Signal Based Surveillance (AVSS). *Proceedings of the 10th IEEE International Conference,* 395, 400. http://dx.doi.org/10.1109/AVSS.2013.6636672

Neyman, J., & Pearson, E. (1933). On the problem of the most efficient tests of statistical hypotheses. Philosophical Transactions of the Royal Society A: *Mathematical, Physical and Engineering Sciences*, 289-337.

Polikar, R. (2006). Ensemble based systems in decision making [A tutorial article on ensemble systems including pseudocode, block diagrams and implementation issues for AdaBoost and other ensemble learning algorithms.]. *IEEE Circuits and Systems Magazine, 6*(3), 21–45. http://dx.doi.org/10.1109/MCAS.2006.1688199

Tuermer, S., Leitloff, J., Reinartz, P., & Stilla, U. (2011). Evaluation of selected features for car detection in aerial images. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial*

Information Sciences, 341-346.

Vedaldi, A., & Soatto, S. (2008). Quick shift and kernel methods for mode seeking. In Proc. ECCV.

Welch, G., & Bishop, G. (2001). An Introduction to the Kalman Filter, SIGGRAPH. Course 8. In *Computer Graphics, Annual Conference on Computer Graphics & Interactive Techniques*. ACM Press, Addison-Wesley, Los Angeles, CA, USA (August 12–17).

Zhang, J., Liu, L., Wang, B., Chen, X., Wang, Q., & Zheng, T. (2012). High speed automatic power line detection and tracking for a UAV-based inspection. *Proceedings of the International Conference on Industrial Control and Electronics Engineering (ICICEE),* 266-269.

Zhong, S., & Ghosh, J. (2003). A unified framework for model-based clustering. *Machine Learning,* 1001-1037.

Copyrights