

# A Novel System for Non-Invasive Method of Animal Tracking and Classification in Designated Area Using Intelligent Camera System

Slavomir MATUSKA, Robert HUDEC, Patrik KAMENCAY, Miroslav BENCO, Martina RADILOVA

Dept. of Telecommunications and Multimedia, University of Zilina, Univerzitna 8215/1, 01026 Zilina, Slovakia

slavomir.matuska@fel.uniza.sk, {robert.hudec, patrik.kamencay, miroslav.benco, martina.radilova}@fel.uniza.sk

Manuscript received July 16, 2015

**Abstract.** *This paper proposed a novel system for non-invasive method of animal tracking and classification in a designated area. The system is based on intelligent devices with cameras, which are situated in a designated area, and a main computing unit (MCU) acting as a system master. Intelligent devices track animals and then send data to MCU for evaluation. The main purpose of this system is detection and classification of moving animals in a designated area and then creation of migration corridors of wild animals. In the intelligent devices, background subtraction method and CAMShift algorithm are used to detect and track animals in the scene. Then, visual descriptors are used to create representation of unknown objects. In order to achieve the best accuracy in classification, key frame extraction method is used to filtrate an object from detection module. Afterwards, Support Vector Machine is used to classify unknown moving animals.*

## Keywords

Animal recognition, object tracking, migration corridors, ASFAR system

## 1. Introduction

Object detection and classification has been very popular research area for many years. Many works were published focusing in object detection and classification in surveillance systems, traffic monitoring, human-machine interface, smart applications and different security solutions [1–4]. But there are only a few publications, which are focusing on detection and classification of wild animals [5], [6]. This was one of the motivations for creation an Automatic System For Animal Recognition (ASFAR system). Another reason was a growing rate of construction of new road infrastructures. This resolved in intervention in natural migration corridors of wild animals. The system is based on intelligent devices equipped with a camera, small computation unit like Raspberry Pi or Odroid-XU3, intelligent sensors and transmission module. These devices

are situated in the designated area, where it is necessary to detect animal movement and collect information. Collected data are used to create migration corridors for particular animals in the designated area. Information about migration corridors can be essential in many areas, particular in planning and constructing of new road infrastructures, in building green migration corridors or any environmental studies. ASFAR system can also use collected data to replace currently standard method for monitoring of the movement of wild animals like field track, direct observation, satellite tracking, Global Positioning System (GPS) tracking or droppings tracking.

For people it is easy to see, track and classify moving objects in real life or in video sequence based on early experiences. This object classification in computer vision is the task of recognizing a given object in the image or video sequence. Humans perform this task as extremely trivial and they can recognize objects even if the objects are rotated or scaled. However, this is one of the hardest challenges for computer vision systems today. In order to succeed in this task, it is necessary to apply algorithms and methods of computer vision. Firstly, the objects representation needs to be created. In ASFAR system, for object representation, descriptors of local visual features like SIFT, SURF and etc. were chosen. Secondly, the Support Vector Machine (SVM) in combination with Bags of visual key points were used to create classification model using local visual features in object classification. To detect and track animals in video sequences, combination of background subtraction method and CAMShift algorithm was used. In this part, a key frame extraction method was used in order to filtrate regions of interest and in this way improve accuracy in later object classification.

## 2. Related Work

A deep convolutional neural network for species recognition in wild nature on camera-trap data was published in [6]. Their dataset was captured with motion triggered camera trap and included 20 animal species. The moving objects were segmented from background using graph-cut

based algorithm. The overall species recognition accuracy was 38.315%. According to their results, they achieved superior performance in comparison with traditional bag of visual words model. Another work for animal classification based on animal head from still images was presented in [7]. To detect effectively shape and textures on animal head, they proposed a new set of gradient oriented features, Haar of Oriented Gradients. Experimental results, on a big dataset consisting of 14379 images from 12 different animal species, validate the superiority of their approach. The local visual features and SVM was used for animal classification in work presented in [5]. The main purpose of their work was to build a vision tools for field biologists to study the currently threatened animal species in Mojave Desert. According to their results, the proposed LBP-like operator outperformed classical SIFT descriptor and achieved an average accuracy of 77.89%.

In [1], the simple system for object classification from video sequences based on the object shape was presented. They used 4 different classes, namely 4-leg animal, car, human and other objects and achieved overall success rate of 86.67%. In [4], a new approach for object detection and tracking in a multiple camera network was presented. This new approach is based on a new algorithm using mean shift segmentation and the depth information derived from stereo vision. The segmented objects are tracked by their novel Bayesian Kalman filter with simplified Gaussian mixture. They used non-training based object recognition algorithm to track and identify similar objects in nearby cameras. Another approach of motion detection was introduced in [8]. They proposed a new version of the original temporal averaging algorithm with adaptive updating speed of the background. Their goal was making the algorithm more robust in various scenarios. A real-time 3D video tracking system for monitoring primate groups was presented in [9]. They used 4 CMOS color cameras to monitor up to 4 animals in one cage. The presented system can follow a number of animals wearing only individual color markers. Their main challenges were the reliability of the position measurements and behavior classification.

### 3. Proposed System Solution

The proposed ASFAR system will be putting in the countryside, without access to the electricity network or cable internet connection. The system needs to work 24 hours a day and as long as possible on the battery. Therefore, there is a need of minimalized power consumption and optimized every process. ASFAR system solution for determining the migration potential of wild animals is shown in Fig. 1.

In the system solution, there are more standalone modules acting as slaves, called “watching devices” and one device acting as master called main computation unit (MCU). The system solution was firstly introduced in [10] as the partial results of our work. The visual descriptors like SIFT and SURF were tested on static images and

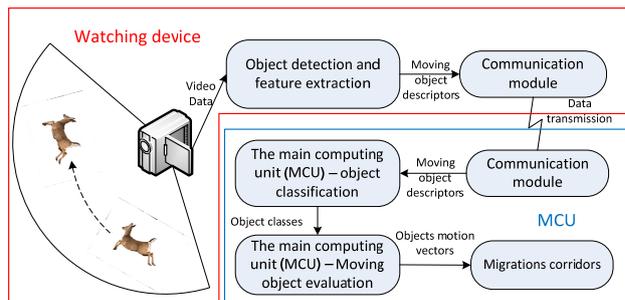


Fig. 1. ASFAR system solution [10].

experimental results showed that this approach is suitable for animal recognition. In comparison with the work presented in [10], our new paper is focusing on software solution of the ASFAR system and brings more experimental results on the real world data.

#### 3.1 Watching Device

The main task of the watching device is animal’s detection and effective description of moving animals in the wild nature. Then, the device sends the descriptors to the MCU for evaluation. The main parts of the watching device are video camera, computation unit, control unit, communication unit, power supply unit and accessories like light, temperature and motion detection sensors, infrared illumination and heating unit.

#### 3.2 Main Computation Unit

The main computation unit is the master, server and management device for the whole ASFAR system. The MCU tasks are:

- Data collection from watching devices.
- Unknown objects evaluation.
- Animals motion vectors and migration corridors determination.
- Result storing.
- Controlling and managing watching devices.

### 4. Proposed Software Solution

ASFAR software proposal is a difficult and extensive task, which can be divided into two main parts:

- Training part.
- Testing part.

The main task of the training part is to create classification model using training data. This model has to characterize animal classes and has to be able to evaluate unknown moving objects into one of the known class. In the testing part, classification model and training data are used to test the model, to determine the model accuracy and last to create the migration corridors for wild animals. The block diagram of the ASFAR software solution is shown in

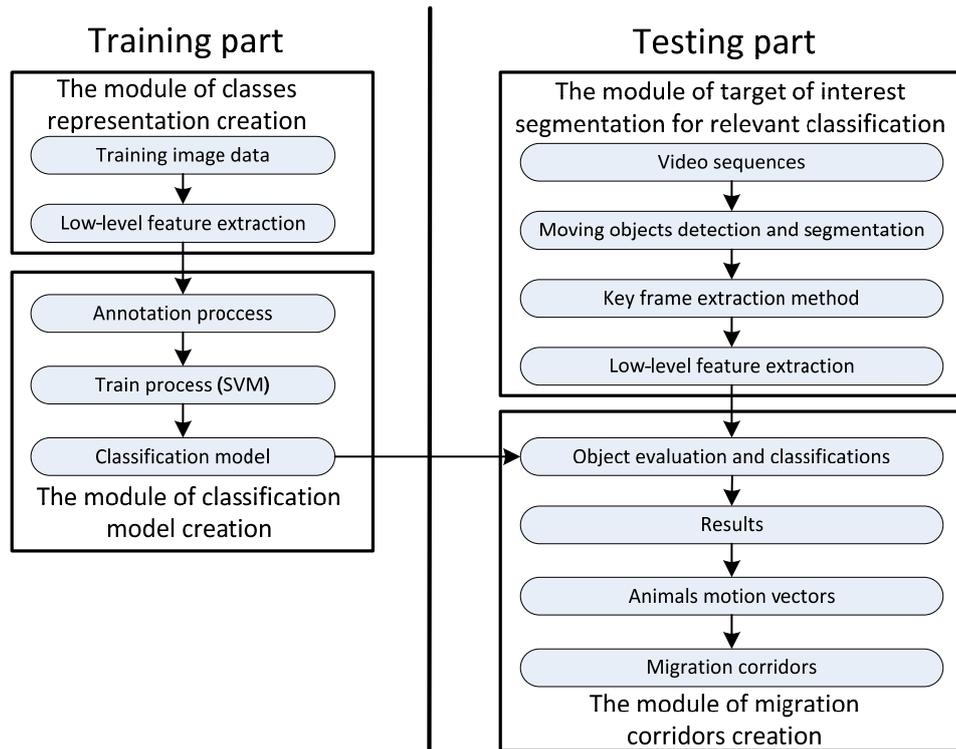


Fig. 2. Block diagram of ASFAR software solution.

Fig. 2, and can be divided into 4 parts:

- The module of classes' representation creation.
- The module of classification model creation.
- The module of target of interest segmentation for relevant classification.
- The module of migration corridors' creation.

#### 4.1 The Module of Classes Representation Creation

In the ASFAR system, the first and the fundamental step is the creation of representation for single classes. For this task, descriptors of local visual features were chosen, namely SIFT, SURF, OpponentSIFT and OpponentSURF [11–13]. These are well-known descriptors, which are widely used in object recognition systems. Every descriptor consists of two parts, key point's detection and key point's description. Two hybrid key point's detectors called SISURF (SiftSURF) and SUSIFT (SURfSIFT) were proposed. The first reason of the proposed detectors was to increase accuracy in object recognition, which was presented in [14] as the partial results of our work. The experimental results showed that SISURF hybrid descriptor outperformed others descriptors in the object recognition. The second reason was to reduce computation time of local descriptors according to achieve real-time animal detection and classification. The hybrid descriptors reduce the number of valid key points and in this way computation time of local descriptors is decreasing. The tests of the speed of the visual descriptors were done, but have not been publish yet.

SISURF hybrid key points detector combines SURF and SIFT detectors and its main idea is detection of key points using main SURF detector and then use SIFT control key point detector to filtrate main key points. The minimal Euclidean distance for every single main key point and the nearest control key point is calculated. From these values, average value of the minimal distances  $min_s$  is calculated using:

$$min_s = \frac{\sum_{j=0}^n min_j \left( \sqrt{(X_{SURF_j} - X_{SIFT})^2 + (Y_{SURF_j} - Y_{SIFT})^2} \right)}{n} \quad (1)$$

where  $X_{SURF_j}$  and  $Y_{SURF_j}$  are coordinates of the  $i^{th}$  main key point,  $i = 0, 1, \dots, n$ ,  $n$  is the number of main key points.  $X_{SIFT}$  and  $Y_{SIFT}$  are coordinates of all control key points. Circled control area is created around every main key point with radius equal to  $min_s$ . The main SURF key point is valid only under condition, when at least 1 control key point is located in the circled control area. The process of creation SISURF key point's detector is shown in Fig. 3.

SUSIFT hybrid key points detector also combines SIFT and SURF detectors. The process of creation SISURT key points is the same as SISURF, but the main key point detector is SIFT and SURF detector is used as the control detector.

#### 4.2 The Module of Classification Model Creation

In the second step, the classification model based on the machine learning using training data was created. This

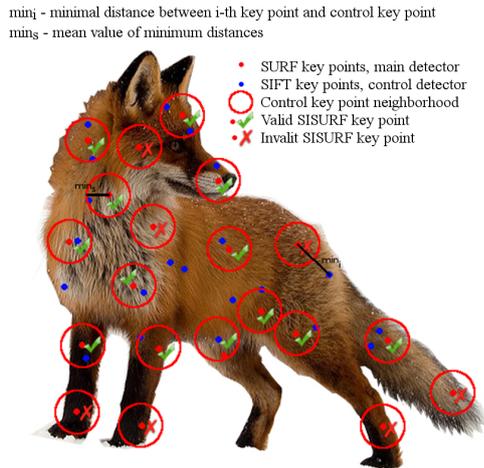


Fig. 3. The process of creation SISURF key point's detector.

model needs to characterize simple classes and needs to evaluate a valid class for unknown objects. In the ASFAR system, the combination of bag of visual key points (BOW) method and SVM were used. BOW was presented in [15] and is based on quantization of affine invariant visual descriptors. The main advantage is its computation efficiency, simplicity and invariance in affine transformation. A BOW vocabulary is created using k-means algorithm and local visual descriptors. These descriptors are assigned to the nearest cluster center using BruteForce or FlannBased matcher. Then, the final visual key point descriptor is created as a normalized histogram of one of the center of vocabulary. Final descriptors serve SVM as training data. Support Vector Machine is widely used in object classification and is related to the family of the supervised learning methods. The main task of SVM classifier is to find an optimal hyper plane with the maximum margin between data of two different classes. Radial Basic Function (RBF) was used as the SVM kernel function.

### 4.3 The Module of Target of Interest Segmentation for Relevant Classification

The main task of this module is detection and tracking of moving objects and marking relevant targets of interest for object classification. The combination of background subtraction method and CAMShift algorithm were used. First, moving objects are separated from the background using background subtraction method. After a successful determination and detection of the moving objects, CAMShift algorithm is applied to find optimal object size, position and orientation. Then, in order to achieve the best accuracy in object classification, key frame extraction method (KEM) is used to filtrate regions of interest. The proposed method consists of two parts:

- Early filtration.
- Late filtration.

Early filtration is applied on the moving region of interest (ROI) in real-time and its task is to do first filtration of valid regions of interest. ROI needs to fulfill two condi-

tions in order to participate in objects recognition:

- Size condition - ROI size must be more than a threshold value.
- Movement condition - ROI must consist at least of 40% moving pixels.

The threshold value for Size condition was determined empirically from testing video sequences and its value was set to 10000 pixels. This condition should filtrate ROIs where the unknown object is not clearly visible in ROI, for example when an object is entering or leaving the scene or is partially overlapped by other objects. The second threshold value was also determined empirically. In testing video sequences average percentage of moving pixels in regions of interest was about 50% and therefore the threshold was set to 40%.

Late filtration is applied on a set of the regions of interest belonging to the tracking object, after this object has left the scene. Late filtration consists of two conditions:

- Edge condition - Number of edges in ROI.
- Length condition - Number of frames with tracking object in video sequences.

Canny detector is used to detect edges in each ROI. Edges are represented by white points in output binary image from Canny detector. Number of white points is counted and average value of white points in images is calculated using:

$$edge_{\text{mean}} = \frac{\sum_{i=0}^n edgeCount_i}{n} \quad (2)$$

where  $n$  is the number of images which belong to a single tracking object and  $edgeCount_i$  is the number of white points in an individual ROI. If the number of edges in ROI is lower than the average value, the particular ROI is used in object classification. Idea of this condition is to remove these ROIs, where the unknown object is overlapping by some other objects from background, such as trees, stones etc. The objects which are not at least 40 frames at video sequence are removed from valid detected objects according to the second condition. The false detected objects in video sequences are also eliminated using this condition.

### 4.4 The Module of Migration Corridors Creation

In the last part of ASFAR system, moving animals are classified to appropriate classes and all information about their movement in the particular area is stored in database. Then, migration corridors in the particular area are created using these data.

## 5. Experimental Results

The whole ASFAR system was programmed and developed in C++ programming language using OpenCV



Fig. 4. The images from the training database.

Animal species	Video length [min]	Total frame count	Frames with animal
Wild boar	08:55	13375	7753
Brown bear	07:23	11075	6255
Wolf	04:52	7300	5011
Fox	01:21	2025	2025
Deer	13:48	20700	19093

Tab. 1. The information about video sequences.

libraries. The system was tested on static video sequences with resolution  $1920 \times 1080$ . In video sequences, moving objects were represented by animals in their natural conditions or in the zoo. These videos were created as a part of the international project E!6752 – DETECTGAME. In video sequences, there were 5 different animal species, namely wild boar, deer, fox, brown bear and wolf. The example images from the training database are shown in Fig. 4.

Information about video length, total number of frames in video sequences and number of frames with animals in the scene per animal species are given in Tab. 1.

In the experiments, we follow the block diagram of the ASFAR system solution. The system was tested in more standalone runs, where one standalone run consists of a combination of key point detector, key point descriptor and descriptor matcher in BOW. Four detectors, SIFT, SURF, SISURF and SUSIFT, four descriptors SIFT, SURF, OpponentSIFT and OpponentSURF were tested and for feature matching, BruteForce (BF) or FlannBased (FB) matchers were used.

For each run, number of descriptors participating in constructing visual vocabulary was changing from 8000 to 20000 per class, with the step 2000. Also, in data preparing for SVM classifier, number of descriptors was changing from 8000 to 20000 per class, with the step 2000. For each run, the following evaluation was done:

- Overall accuracy.
- Accuracy per class with the effect of key frame extraction method.
- Recall, precision and F1 measure.
- Confusion matrix.

### 5.1 Precision, Recall and F1 Measure

Precision is defined as a proportion between the num-

ber of valid frames specified by the module and the number of all frames classified as valid frames. Precision is defined by:

$$Precision = \frac{t_p}{t_p + f_p} \tag{3}$$

where  $t_p$  defines the number of frames correctly classified as a valid class and  $f_p$  defines the number of frames incorrectly classified as a valid class. Recall is defined as a proportion between the number of valid frames specified by the module and the number of all valid frames. Recall is defined by:

$$Recall = \frac{t_p}{t_p + f_n} \tag{4}$$

where  $t_p$  defines the number of frames correctly classified as a valid class and  $f_n$  defines the number of frames which were not classified as valid, but they should be. F1 measure combines precision and recall in harmonic mean and F1 measure is defined by:

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{5}$$

### 5.2 Results

The overall results for the best combinations of the classification models for animal recognition from video sequences are shown in Tab. 2. One combination consists of detector, descriptor, matcher, number of descriptors used in the vocabulary building process and the number of descriptors used in SVM machine learning. The best accuracy 86.07% was achieved in combination of SURF detector, OpponentSIFT descriptor, BruteForce matcher, 14000 descriptors in the vocabulary building process and 16000 descriptors in machine learning. The best accuracy with the proposed hybrid key points detectors was 81.98% and it was achieved in combination of SISURF detector, OpponentSURF descriptor and FlannBased matcher, 14000 and 12000 used descriptors.

Classification model	Overall accuracy [%]
SURF, OpponentSIFT, BF, 14000, 16000	86.07
SURF, OpponentSURF, FB, 12000, 14000	85.94
SIFT, OpponentSIFT, BF, 14000, 16000	84.95
SURF, OpponentSURF, FB, 14000, 12000	84.09
SISURF, OpponentSURF, FB, 14000, 12000	81.99
SURF, SIFT, BF, 16000, 18000	81.38
SURF, OpponentSIFT, BF, 10000, 12000	80.99
SURF, SURF, BF, 14000, 14000	80.06
SURF, SURF, FB, 12000, 12000	79.74
SUSIFT, OpponentSIFT, FB, 12000, 14000	79.23
SISURF, OpponentSIFT, BF, 8000, 10000	77.89

Tab. 2. Overall accuracy for the combination of the best classification models.

The final classification score and influence of KEM method in object recognition is shown in Fig. 5. The shortcuts in the bottom part of the image mean which parts of KEM method are used in ROI filtration in order to improve

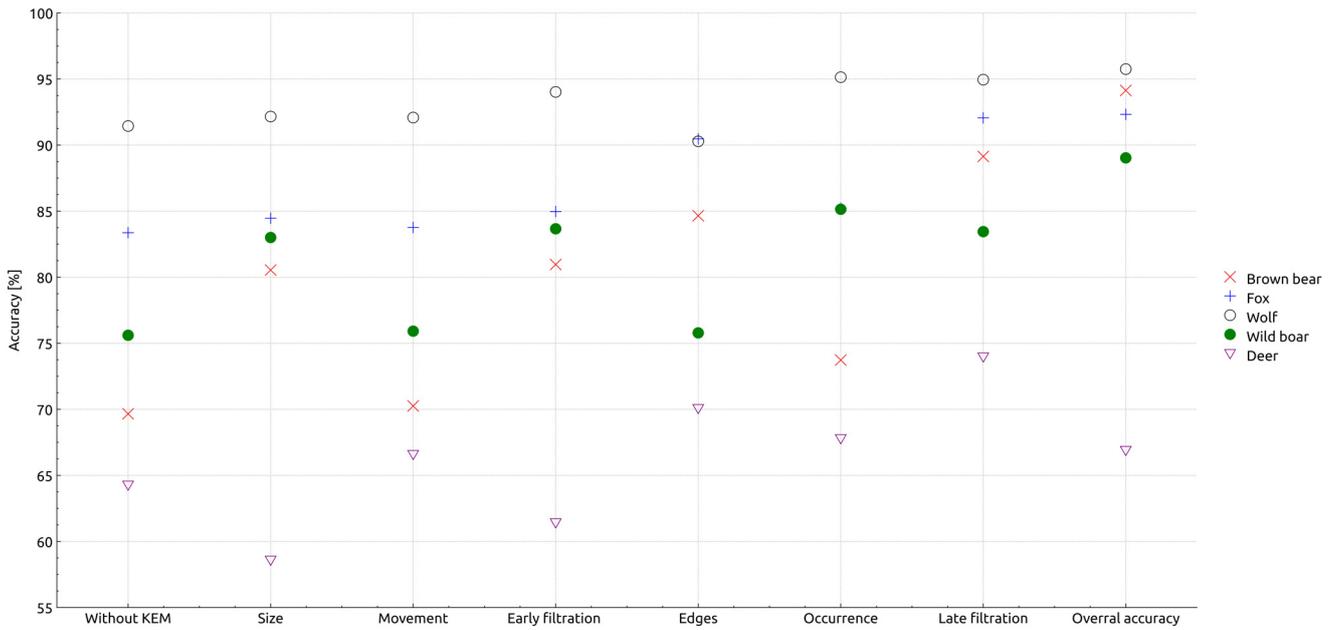


Fig. 5. Influence of KEM in animal classification from video sequences.

classification accuracy:

- Without KEM – no filtration is made and all ROIs are used in classification.
- Size – Only the ROIs which passed Size condition are used in classification.
- Movement – Only the ROIs which passed Movement condition are used in classification.
- Early filtration - Only the ROIs which passed early filtration are used in classification.
- Edges - Only the ROIs which passed Edges condition are used in classification.
- Occurrence - Only the ROIs which passed Length condition are used in classification.
- Late filtration - Only the ROIs which passed late filtration are used in classification.
- Overall accuracy - Only the ROIs which passed early and late filtration are used in classification.

From Fig. 5 the positive influence of KEM method in achieving the best accuracy in object recognition is evident. The best improvement using KEM method about 24.48% was achieved for class brown bear with the final accuracy of 94.13%. The lowest improvement only about 2.63% was achieved for class deer with final accuracy of 66.89%. The best final accuracy per class was achieved for class wolf at 95.75% with improvement about 4.3%. For class wild boar, final classification accuracy of 89.3% was achieved. It was improvement of about 13.43% in comparison without using KEM method. The high classification accuracy of 92.32%.with a slight improvement of about 8.95% was also achieved for class fox.

The best whole model accuracy was achieved at 86.07% in combination of SURF detector, OpponentSIFT descriptor, BruteForce matcher, 14000 descriptors in the

vocabulary building process and 16000 descriptors in machine learning. Using KEM method for this model, overall accuracy was improved of about 12.27%. Recall, precision and F1 measure for this model are shown in Tab. 3.

Animal species	Precision	Recall	F1 - measure
Wild boar	0.927	0.890	0.908
Brown bear	0.987	0.941	0.963
Wolf	0.701	0.957	0.809
Fox	0.668	0.923	0.775
Deer	0.996	0.668	0.800

Tab. 3. Precision, recall and F1 measure.

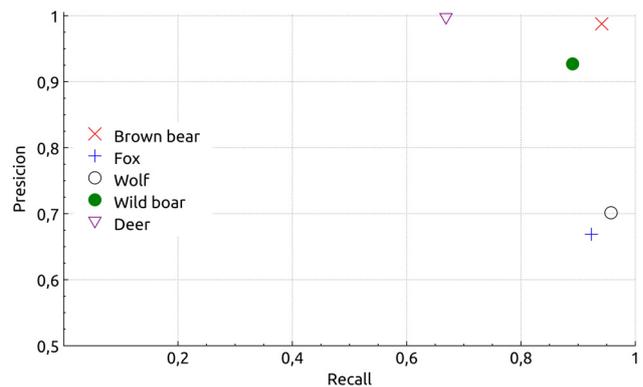


Fig. 6. Graph recall-precision.

True classes →	Wild boar	Brown bear	Wolf	Fox	Deer
Wild boar	966	39	12	3	22
Brown bear	4	1043	0	1	8
Wolf	92	22	721	37	156
Fox	23	4	19	517	210
Deer	0	0	1	2	800

Tab. 4. Confusion matrix for the best classification model.

Graphical representation for recall-precision is shown in Fig. 6. Confusion matrix for the best combination is shown in Tab. 4.

From the confusion matrix it is evident that 37 objects from fox class were incorrectly evaluated by the classifier as wolf. The total objects count for fox class was 560. On the other hand, 19 objects from wolf class were incorrectly evaluated as fox and also 12 objects as wild boar, total objects count for wolf class was 753. The total objects count for wild boar class was 1085, 92 objects were mistakenly classified as wolf and 23 as fox. The highest classification accuracy was achieved for brown bear class. Therefore, only 39 objects were classified as wild boar, 22 as wolf and 4 as fox from the total brown bear objects count 1108. Lots of objects from deer class were incorrectly evaluated. Moreover, 210 were classified as fox, 156 as wolf and 22 as wild boar from the total objects count for deer class 1196. This was the reason, why the lowest value of recall for deer class was achieved.

## 6. Conclusion

In this paper, the novel system for non-invasive method of animal tracking and classification was presented. The system is based on intelligent watching devices with camera, computation and transmission unit. Devices will be put in countryside, where it is necessary to determine migration corridors of wild animals. The presented system was tested on static video sequences and shows promising performances in comparison with state-of-art techniques, which were presented in Sec. 2. The best accuracy was achieved at 86.07%. In overall testing, hybrid descriptors failed to outperform classical descriptors in object classification accuracy. Because of the demand of real-time object classification, the combination with the best accuracy is not satisfactory. Therefore, the combination with hybrid detector can be the compromise between the speed and still acceptable accuracy at 81.98%. Now, the system is not capable of fully automatic process of migration corridor creation in large areas, because the system collects data only from one camera. With more watching devices in a designated area, the system will be able to collect data from all devices and in this way migration corridors for wild animals in larger areas can be created in a very efficient way.

## Acknowledgments

The work presented in the paper was supported by the University Science Park of the University of Zilina (ITMS: 26220220184) supported by the Research & Development Operational Program funded by the European Regional Development Fund and EUREKA project no. E! 6752 – DETECTGAME: R&D for Integrated Artificial Intelligent System for Detecting the Wildlife Migration.

## References

- [1] ZHI, K., CHEE, F.N., SIAK, W. K. Shape-based recognition and classification for common objects. In *Application in Video Scene Analysis Computer Engineering and Technology (IC CET)*. Chengdu (China), 2010, p. V3-13–V3-16. DOI: 10.1109/ICCEt.2010.5485747
- [2] ISLAM, M. K., JAHAN, F., MIN, J-H., et al. Object classification based on visual and extended features for video surveillance application. In *Proceedings of 2011 8th Asian Control Conference (ASCC)*. Kaohsiung (Taiwan), 2011, p. 1398–1401. Print ISBN: 9781612844879
- [3] BORJI, A., ITTI, L. Human vs. computer in scene and object recognition. In *Computer Vision and Pattern Recognition (CVPR)*. Columbus (OH, USA), 2014, p. 113–120. DOI: 10.1109/CVPR.2014.22
- [4] ZHANG, S., WANG, CH., CHAN, S-CH., et al. New object detection, tracking, and recognition approaches for video surveillance over camera network. *IEEE Sensors Journal*, 2015, vol. 15, no. 5, p. 2679–2691. DOI: 10.1109/JSEN.2014.2382174
- [5] WILBER, M.J., SCHEIRER, W.J., LEINER, P., et al. Animal recognition in the Mojave Desert: Vision tools for field biologists. In *Applications of Computer Vision (WACV)*. Tampa (FL, USA), 2013, p. 216–213. DOI: 10.1109/WACV.2013.6475020
- [6] GUOBIN, CH., HAN, T.X., ZHIHAI, H., et al. Deep convolutional neural network based species recognition for wild animal monitoring. In *Image Processing (ICIP)*. Paris (France), 2014, p. 858–862. DOI: 10.1109/ICIP.2014.7025172
- [7] WEIWEI, Z., JIAN, S., XIAOOU, T. From tiger to panda: Animal head detection. *IEEE Transactions on Image Processing*, 2010, vol. 20, no. 6, p. 1696–1708. DOI: 10.1109/TIP.2010.2099126
- [8] NIKOLOV, B., KOSTOV, N. Motion detection using adaptive temporal averaging method. *Radioengineering*. 2014, vol. 23, no. 2, p. 652–658. ISSN: 1805-9600
- [9] BALLESTA, S., REYMOND, G., POZZOBON, M., et al. A real-time 3D video tracking system for monitoring primate groups. *Journal of Neuroscience Methods*, 30 August 2014, vol. 234, p. 147–152. DOI: 10.1016/j.jneumeth.2014.05.022
- [10] MATUSKA S., HUDEC R., KAMENCAY P., et al. A novel system for automatic detection and classification of animal. In *ELEKTRO 2014*. Rajecke Teplice (Slovakia), 19-20 May 2014, p. 76–80. DOI: 10.1109/ELEKTRO.2014.6847875
- [11] BAY, H., ESS, A., TUYTELAARS, T., et al. SURF: Speeded Up Robust Features. *Computer Vision and Image Understanding (CVIU)*, 2008, vol. 110, no. 3, p. 346–359. ISSN: 1077-3142, DOI: 10.1016/j.cviu.2007.09.014
- [12] LOWE, D.G. Distinctive image features from scale-invariant key points. *International Journal of Computer Vision*. 2004, vol. 60, no. 2, p. 91–110. DOI: 10.1023/B:VISI.0000029664.99615.94
- [13] VAN DE SANDE, K.E.A., GEVERS, T., SNOEK, C.G.M. Evaluating color descriptors for object and scene recognition. *Pattern Analysis and Machine Intelligence*. 2009, vol. 32, no. 9, p. 1582–1596. DOI: 10.1109/TPAMI.2009.154
- [14] MATUSKA, S., HUDEC, R., KAMENCAY, P., et al. Classification of wild animals based on SVM and local descriptors. In *AASRI Procedia, 2014 AASRI Conference on Circuit and Signal Processing (CSP 2014)*. ISSN: 2212-6716. DOI: 10.1016/j.aasri.2014.09.006
- [15] CSURKA, G., DANCE, CH. R., FAN, L., et al. Visual categorization with bags of keypoints. In *ECCV International Workshop on Statistical Learning in Computer Vision*. Prague, 2004.

## About the Authors ...

**Slavomir MATUSKA** was born in Banska Bystrica. He received his M.Sc. and Ph.D. degrees at the Dept. of Telecommunications and Multimedia, University of Zilina, Slovakia, in 2011, 2014, respectively. He has been a researcher at University of Zilina since September 2014. Areas of his research include digital image processing, semantic analysis of multimedia content, low-level pattern description and object classification.

**Robert HUDEC** was born in Revuca, Slovakia in 1974. He received his M.S. and Ph.D. degrees from the Technical University in Kosice, Faculty of Electrical Engineering and Informatics, Slovakia, in 1998 and 2003, respectively. Nowadays, he is an Associated Professor at the Dept. of Telecommunications and Multimedia, University of Zilina. Areas of his research include digital image filtration, segmentation, combined low-level description, object classification, video concealment and 3D face recognition. From 2005 his research interests cover also systems, services, terminals and intelligent textile sensors for e-/m-health applications. He received Werner von Siemens Excellent Award and Jozef Murgas Award for his research on image filtration in 2003 and 2007, respectively, and Award of „Vice premier and minister of school for science and research“ in the category: celebrity of research and develop-

ment till 35 years old for his research on e-/m-health systems and terminals in 2006.

**Miroslav BENCO** was born in Vranov nad Toplou, Slovakia in 1981. He received his M.Sc. degree in 2005 at the Dept. of Control and Information System, and Ph.D. degree in 2009 at the Dept. of Telecommunications and Multimedia, University of Zilina. Since January 2009 he has been a researcher at the Dept. of Telecommunications, University of Zilina. His research interest includes digital image processing, semantic analysis of multimedia content, image and video classification and retrieval for MPEG-7 standard.

**Patrik KAMENCAY** received his M.Sc. and Ph.D. degrees in Telecommunications from the University of Zilina, Slovakia, in 2009 and 2012, respectively. His Ph.D. research work was oriented to reconstruction of 3D images from stereo pictures. Since Oct. 2012 he has been a researcher at the Dept. of Telecommunications, Univ. of Zilina. His research interest includes holography for 3D display and construction of 3D images of the original objects.

**Martina RADILOVA** received her M.Sc. degree in Telecommunications from the University of Zilina, in 2010. Currently, she is a Ph.D. student at the Dept. of Telecommunications and Multimedia. Her doctoral work is in area of semantic analysis of web document. Her research interests include semantic description of images.