A Novel System for Automatic Detection and Classification of Animal

Slavomir Matuska, Robert Hudec, Miroslav Benco, Patrik Kamencay, Martina Zachariasova
Department of Telecommunication and Multimedia
University of Zilina, Univerzitna 8215/1 01026
Zilina, Slovakia
slavomir.matuska, robert.hudec, miroslav.benco, martina.zachariasova, patrik.kamencay@fel.uniza.sk

Abstract— **In this paper, a novel system for automatic detection and classification of animal is presented. System called ASFAR (Automatic System For Animal Recognition) is based on distributed so-called 'watching device' in designated area and main computing unit (MCU) acting as server and system manager. Watching devices are situated in wild nature and their task is to detect animal and then send data to MCU to evaluation. The main task of whole system is to determine migration corridors of wild animals in designated area. To create object representation, visual descriptors were chosen and Support Vector Machine (SVM) was used to classify descriptors.**

Keywords—Support Vector Machine, automatic system, object recognition, visual descriptor, ASFAR

I. Introduction

In the recent years, many organizations and governments spent considerable resources on environmental protection of various animal species that are endangered in their natural environment, particularly by the building of new infrastructures. The growing rate of construction new road infrastructures leads to intervention in natural migration corridors of wild animals. This is the main reason to develop an integrated automatic system with elements of artificial intelligence to monitor the movement of animals which will provide data of wildlife migration in designated area. This system should replace currently standard methods (direct observation, field tracks, droppings and others) that can not cover a continuous period of time and even then, this is very time consuming. The information about migration potential of wild animals can be used for effective design of migration corridors in planning and contraction of new road infrastructures. In [1], animal recognition system in the Mojave Desert was used in the effort to protect endangered and threatened species in wild.

The outline of the paper is as follows. In the second section, proposal of a novel automatic system for animal detection and classification is presented in detail. In the third section, all necessary software equipment in main computing unit is given. Finally, experimental results are shown and discussed in fourth section. In the end, conclusion and future work is presented.

II. Automatic System Proposal

Based on collected data from designated area, the main task of a proposal automatic system is to create motion vectors of animals and desired migration corridors. Creation of desired automatic system for animal recognition (ASFAR) is a difficult and extensive problem, because of computational demands, size of transferred data and the environmental conditions. ASFAR will be putting in wild nature, often without access to the electricity network and cable internet connection. System needs to work 24 hour a days and as long as possible and therefore, there is a need to minimize the power consumption of the system. It is necessary to deploy a detection algorithm and algorithms for effective object description to reduce data transfer over communication module.

System solution of intelligence video system for determining the migration potential of wild animals is shown in Fig. 1.

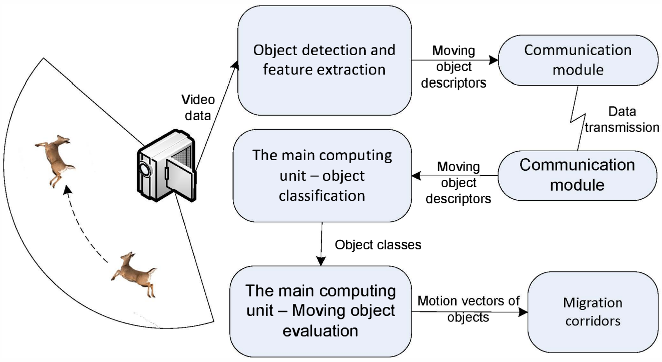

Fig. 1. ASFAR system solution

In the system solution, there are more standalone modules, called "watching device" and one device, called main computation unit (MCU).

A. *Watching device*

The main task of watching device is to detect animals in wild nature, effective describe moving animals and then send the descriptions to MCU for evaluation. Watching device schematic is shown in Fig.2. The main parts of the watching device are video camera, computation unit, control unit, communication unit, power supply unit and accessories like light and temperature sensors, infra red illumination and heating unit.

1) Video camera unit: In the term of effective evaluation of migration corridors in designated area, there is need to place a watching device into exact location to cover whole designated area. One watching device can cover different areas, depending on viewing angle of video camera unit.

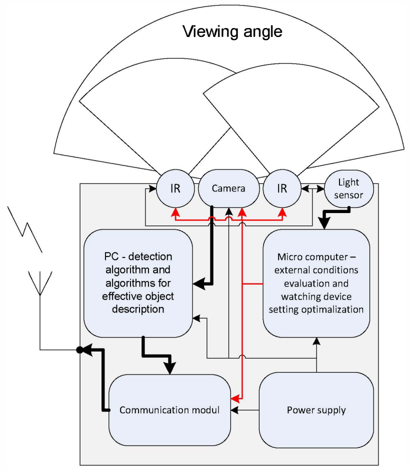

Fig. 2. Watching device schematic

In the system solution, there are three possible video camera unit variants:

- one camera with viewing angle 180° , Fig 3a.,
- three cameras with viewing angle at 120° , Fig 3b.,
- one rotation camera with 360° viewing angle, Fig 3c.

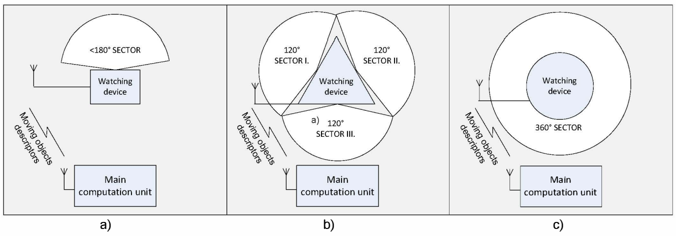

Fig. 3. Video camera unit variants

Sensing in infra red range, full HD resolution, 20 or more fps and built-in IR illumination are additional requirements for video camera unit.

2) Computation unit in watching device: Computation unit needs to have enough copmutation power to handle real-time moving object detection and segmentation from raw video data. The main task of this unit is to run all necessary computation that precedes sending data to MCU. These computations involve algorithms for object detection, segmentation and descriptions.

3) Control unit in watching device: The task of control unit is to monitor the external weather condition using build-in sensors and optimizing functionality, efficiency and precision of watching device. Using information from light sensor, control unit adjust the intensity of IR illumination and activates or deactivates day/night camera mode. The level of power supply is evaluated continuously and in the case of low power, a warning message is sent to MCU.

B. Main Computing Unit

Main computing unit is a server and management device for whole ASFAR. The tasks of MCU are:

- collect data from watching devices,
- using classification algorithms to evaluate unknown objects to known classes,
- determine moving vectors of animals and migration corridors,
- store all results in database,
- control and manage watching devices and others.

Schematic for proposal ASFAR system covering designated area is shown in Fig.4.

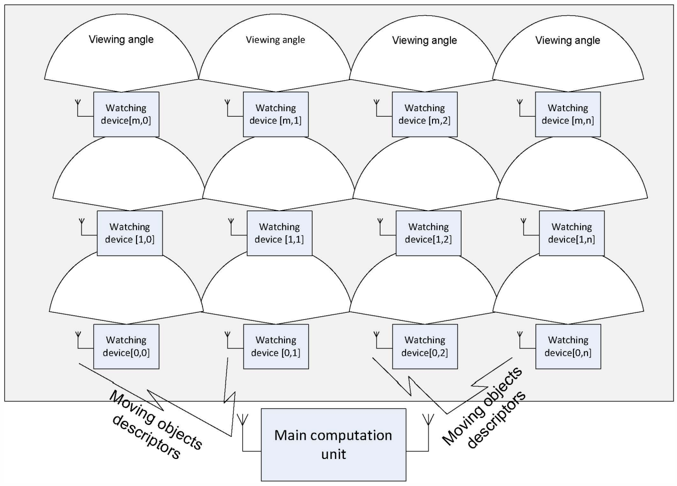

Fig. 4. ASFAR system proposal

III. SOFTWARE EQUIPMENT IN MCU

One of the main tasks of MCU is evaluate unknown object received from watching device to known classes. To perform this task, there is need to use methods for object recognition and classification. In ASFAR system, the method for object recognition and classification follow principle shown in Fig.5 is used. This method consists of two parts, training and testing part.

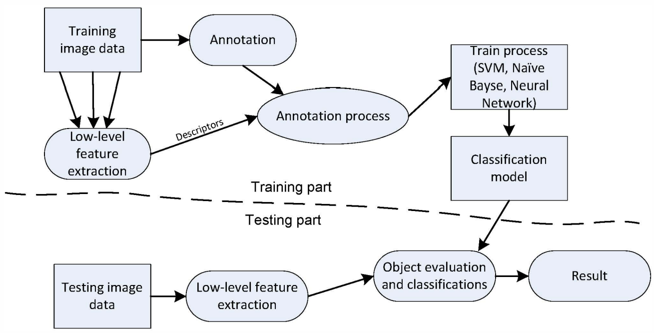

Fig. 5. Object recognition process

In training part, visual descriptors are extracted from training image dataset and they are used to create a classification model. In testing part, classification model is use to evaluate an unknown objects to the appropriate class [2],[3].

A. Visual descriptors

Visual descriptors are used to capture the local appearance of objects. In ASFAR, the visual descriptors were chosen to create an object description, because their compute efficiency and good performance. Each visual descriptor consists of key point detector and descriptor [4].

1) Key Points Detector: Task of detector is to find key points in the image. There are many methods to detect key points. In ASFAR, SIFT (Scale-invariant feature transform), SURF (Speeded Up Robust Features) were used. More information about these method can be found in [5],[6] for SIFT descriptor, in [7],[8] for SURF descriptor.

2) Key Points Descriptors: Task of key point descriptor is to describe key point by the n-dimensional feature vector. For this task, in ASFAR system three types of key points descriptors were chosen, namely SIFT, SURF and Opponent colour descriptors. More information about key point descriptors can be found in [5],[6] for SIFT, [7],[8] for SURF and [9] for Opponent colour descriptors.

B. Animal Classification

In ASFAR system, combination Bags of visual key points (BOW) and Support Vector Machine (SVM) methods were used to create a classification model. First, training data collections are used to set up the classification model parameters to distinguish different classes. Then, the classifier is able to evaluate an unknown object to the appropriate class.

1) Bags of Visual Keypoint: This method was presented in [10] and is based on quantization of affine invariant visual descriptors. The main advantage is their computationally efficiency, simplicity and invariance in affine transformation.

2) Support Vector Machine: This classification method is widely used in computer vision and machine learning and is related to the family of 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 [11],[12]. In ASFAR, radial basic function (RBF) kernel was used.

IV. Experimental Results

In the experiments, software equipment in MCU was tested. Training dataset contain of images separated into 5 classes: fox, deer, wolf, brown bear and wild boar. Each class contains about 60 images with appropriate animal. Testing database consists of 50 images, 10 images per class. The example of images from training database is shown in Fig. 6. Classification model was created following the principle of object recognition shown in Fig. 5. First, visual descriptors from training images were extracted. Then, visual vocabulary for these descriptors was created using bag of key points method, where K-means algorithm is used to separate descriptors into clusters.

Fig. 6. The images from training database

Two matching algorithm for assignment training descriptors to clusters were used. BruteForce (BF) matcher finds nearest cluster centre in vocabulary in term of Euclidean distance between extracted descriptor and clusters centre. FlannBased (FB) using fast nearest neighbor search in large datasets to find nearest cluster centre. Then, BOW image descriptor extractor compute an image descriptor using the set visual vocabulary. These data were used in SVM classifier to create a classification model.

In the experiment, a total 2 key point detectors SURF and SIFT were used. Afterwards, to describe a key point 4 descriptors SIFT, SURF, OpponentSIFT and OpponentSURF were used. In BOW, BruteForce and FlannBased matchers were used. Combination detector, descriptor and matcher were used to create one standalone run. For each run, number of descriptors participating in constructing visual vocabulary was changing from 8000 to 20000 per class increased of the 2000. Also, in data preparing for SVM classifier, number of descriptors was changing from 8000 to 20000 per class increased of the 2000. Together, there were 98 possible combinations in one run.

Average score of animal classification for the best 5 runs in combination SIFT or SURF detector and SUFT descriptor is shown in Fig.7. Horizontal axis defines specification for particular run, in the order: detector, matcher, number of descriptors in clustering process, number of descriptors in learning process in figure from 7 to 10.

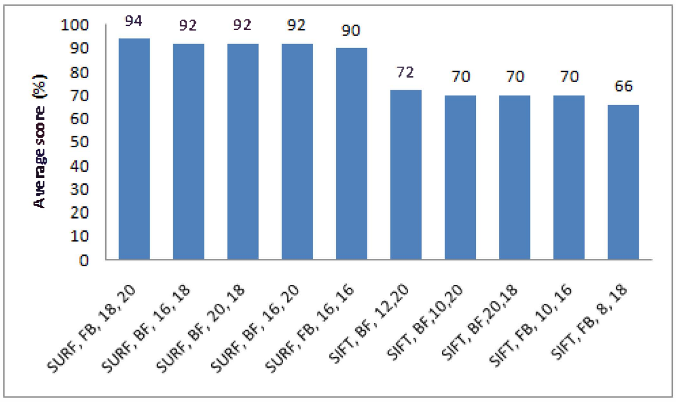

Fig. 7. Average score for SIFT descriptor

Confusion matrix for best run in combination SURF detector, SIFT descriptor, FlannBased matcher, 18000 descriptors in clustering process and 20000 descriptors in learning process with classification score 94 % is shown in Table 1.

TABLE I. CONFUSION MATRIX FOR BEST RUN WITH SIFT DESCRIPTOR

True classes →	Wild boar	Brown bear	Wolf	Fox	Deer
Wild boar	**10**	0	0	1	0
Brown bear	0	**10**	0	0	0
Wolf	0	0	**9**	0	0
Fox	0	0	1	**8**	0
Deer	0	0	0	1	**10**

Average score of animal classification for the best 5 runs in combination SIFT or SURF detector and SURF descriptor is shown in Fig.8.

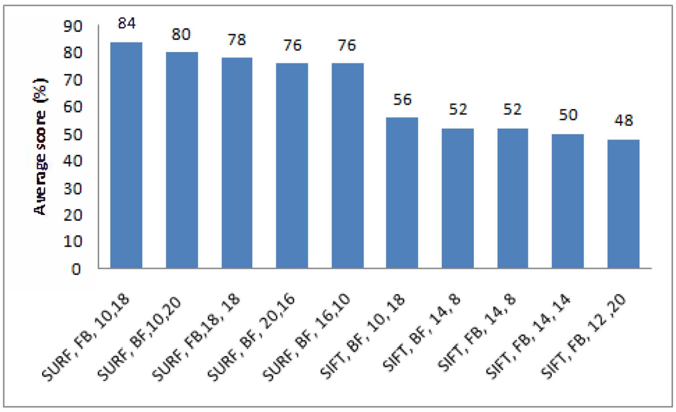

Fig. 8. Average score for SURF descriptor

Confusion matrix for best run in combination SURF detector, SURF descriptor, FlannBased matcher, 10000 descriptors in clustering process and 18000 descriptors in learning process with average classification score 84 % is shown in Table 2.

TABLE II. CONFUSION MATRIX FOR BEST RUN WITH SURF DESCRIPTOR

True classes →	Wild boar	Brown bear	Wolf	Fox	Deer
Wild boar	**5**	0	0	1	0
Brown bear	3	**10**	0	0	1
Wolf	1	0	**10**	0	0
Fox	1	0	0	**9**	0
Deer	0	0	0	0	**9**

Average score of animal classification for the best 5 runs in combination SIFT or SURF detector and OpponentSIFT descriptor is shown in Fig.9.

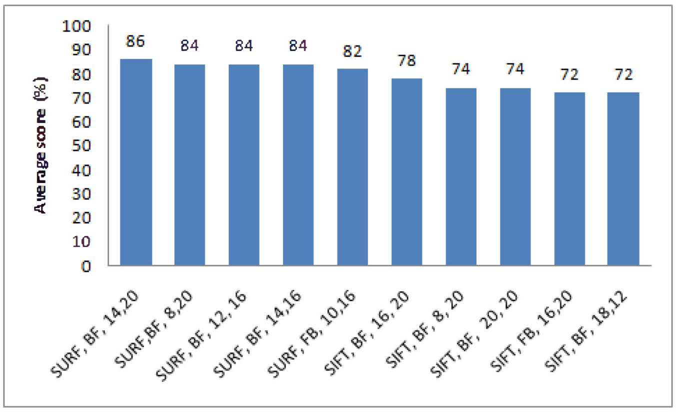

Fig. 9. Average score for OppoenentSIFT descriptor

Confusion matrix for best run in combination SURF detector, OpponentSIFT descriptor, BruteForce matcher, 14000 descriptors in clustering process and 20000 descriptors in learning process with average classification score 86 % is shown in Table 3.

TABLE III. CONFUSION MATRIX FOR BEST RUN WITH OPPONENTSIFT DESCRIPTOR

True classes →	Wild boar	Brown bear	Wolf	Fox	Deer
Wild boar	**5**	0	0	1	0
Brown bear	0	**10**	0	0	0
Wolf	0	0	**10**	0	0
Fox	4	0	0	**8**	1
Deer	1	0	0	1	**9**

Average score of animal classification for the best 5 runs in combination SIFT or SURF detector and SURF descriptor is shown in Fig.8.

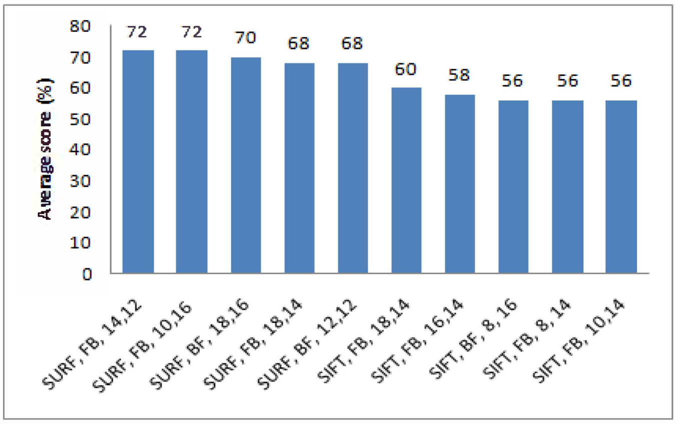

Fig. 10. Average score for OpponentSURF descriptor

Confusion matrix for best run in combination SURF detector, OpponetSURF descriptor, FlannBased matcher, 14000 descriptors in clustering process and 12000 descriptors in learning process with average classification score 72 % is shown in Table 4.

TABLE IV. CONFUSION MATRIX FOR BEST RUN WITH OPPONENTSURF DESCRIPTOR

True classes →	Wild boar	Brown bear	Wolf	Fox	Deer
Wild boar	**6**	0	2	1	2
Brown bear	3	**10**	1	0	0
Wolf	1	0	**7**	0	0
Fox	0	0	0	**8**	1
Deer	0	0	0	1	**7**

V. CONCLUSION AND FUTURE WORK

In this paper, a novel automatic system for animal recognition (ASFAR) was presented. System collects data from its agents – watching devices, located in wild nature. Watching device detect animals and then send data to MCU. MCU evaluate these data and create migration corridors for animals in designated area. In MCU, to classify unknown object BOW method with SVM classifier were used. This classification method was tested in our experiments. From the realized experiment is evident that highest classification score 94 % was achieved by algorithm in combination SURF detector, SIFT descriptor, FlannBased matcher, 18000 descriptors in clustering process and 20000 descriptors in learning process. Classification success rate 92 % was achieved in three more runs in combination SURF detector, SIFT descriptor, BruteForce matcher and variable descriptors in clustering and learning process. Maximum success rate only 72 % was achieved by combination SIFT detector, SIFT descriptor, arbitrary matcher and arbitrary number of descriptors in clustering and learning process. Moreover, in combination with other key point descriptors (SURF, OpponentSIFT, OpponentSURF), SURF detector outperformed SIFT detector in all case. The best success score with SIFT key point detector 78 % was achieved in combination with OpponentSIFT descriptor, BruteForce matcher, 16000 descriptors in clustering process and 20000 descriptors in learning. Promising classification success rate 86 % and 84 % was achieved by run in combination SURF detector, OpponentSIFT descriptor, BruteForce matcher, 14000 descriptors in clustering process and 20000 descriptors in learning process respectively SURF detector, SURF descriptor FlannBased matcher, 10000 descriptors in clustering process and 18000 descriptors in learning process.

In the future, the best runs will be used in MCU to evaluate unknown object. In ASFAR, there is need to develop good and effective method for animal detection and segmentation. Then, it is necessary to chose only the best images with animal in the term of the most relevant classification.

ACKNOWLEDGMENT

The work presented in the paper has been supported by the Slovak Science project Grant Agency, Project No. 1/0705/13 "Image elements classification for semantic image description" and EUREKA project no. E! 6752 – DETECTGAME: R&D for Integrated Artificial Intelligent System for Detecting the Wildlife Migration.

REFERENCES

[1] M.J. Wilber, W.J. Scheirer, P. Leitner, B. Heflin, J. Zott, D. Reinke, D.K. Delaney, T. E. Boult, "Animal Recognition in the Mojave Desert: Vision Tools for Field Biologists", Applications of Computer Vision (WACV), 2013 IEEE Workshop on, 15-17 Jan. 2013, Page(s): 206-213, ISSN: 1550-5790.

[2] M. Sonka, V. Hlavac ,R. Boyle, "Image Processing, Analysis and Machine Vision", Thomson Learning, part of the Thompson Corporation, ISBN: 10: 0-495-24438-4, 2008 J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73.

[3] C. T. Cao, "Object Recognition", Publisher: InTech, ISBN 978-953-307-222-7 2011.

[4] B. Ganesharajah, S. Mahesan, U. A. J. Pinidiyaarachchi, "Robust Invariant Descriptors for Visual Object Recognition", Industrial and Information Systems (ICIIS), 2011 6th IEEE International Conference on, 16-19 Aug. 2011, Page(s): 158 – 163, ISBN 978-1-4577-0032-3.

[5] D. G. Lowe.: Distinctive Image Features from Scale-Invariant Keypoints, The International Journal of Computer Vision, 2004, Pages 91-110, Print ISSN 0920-5691.

[6] M. Nixon, A. Aguando, "Feature extraction & image processing", second edition, ISBN 978-0-1237-2538-7, 2008.

[7] K.N. Ngan, H. Li, "Video Segmentation and Its Applications", Springer Science+Business Media 2011, ISBN 978-1441994813.

[8] T. Bay, T. Tuytelaars, V.G. Luc, "SURF, Speeded Up Robust Features", ETH Zurich, Preprint submitted to Elsevier, 10 September 2008.

[9] E. A. Koen, T. Gevers, G.M. Cees, "Color Descriptors for object category recognition", Pattern Analysis and Machine Intelligence, IEEE Transactions on (Volume:32 , Issue: 9), Sept. 2010, pages 1582 – 1596, ISSN: 0162-8828.

[10] E. A. Koen, T. Gevers, G.M. Cees, "Color Descriptors for object category recognition", Pattern Analysis and Machine Intelligence, IEEE Transactions on (Volume:32 , Issue: 9), Sept. 2010, pages 1582 – 1596, ISSN: 0162-8828.

[11] M. A. M. Shukran, Y.Y. Chung, W. Ch. Yeh, N. Wahid, "Image Classification Technique using Modified Particle Swarm Optimization", Modern Applied Science, Vol. 5, No. 5; October 2011.

[12] D. J. Sebald, J. A. Bucklew, "Support Vector Machine Techniques for Nonlinear Equlization", Signal Processing, IEEE Transactions on (Volume:48 , Issue: 11), Nov 2000, Page(s): 3217 – 3226, ISSN: 1053-587X.