Archives of
Transport System
Telematics

Volume 6

Issue 4

November 2013

Classification of animals to determine the migration potential at the construction of new infrastructure

S. MATUSKA[a] , R. HUDEC[a], M. BENCO[a], M. ZACHARIASOVA[a]
[a] UNIVERSITY OF ZILINA, Department of Telecommunication and Multimedia, Zilina, Slovakia
EMAIL: slavomir.matuska@fel.uniza.sk

ABSTRACT

At the planning and construction of new infrastructures, the information about migration potential of animals in a target area is needed. This information will be used to design of migration corridors for wild animals. To determine the migration potential of animals based on distributed video camera system, new methods for object recognition and classification are developed. In general, an object recognition system consists of three steps, namely, the image feature extraction from the training database, training the classifier and evaluation of query image of object/animal. In this paper, an extraction of local key point by SIFT or SURF descriptors, bags of key points method in combination with SVM classifier and two hybrid key points detection methods are proposed in detail.

KEYWORDS: SIFT, SURF descriptors, SVM classifier, animal classification

1. Introduction

Tasks of image recognition are currently being addressed in many fields of human activity. The various organizations and government spent considerable resources on environmental protection of various animal species that are endangered in their natural environment, particularly by the building of new infrastructures. There is a need to develop an integrated 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. Therefore, in this paper, object recognition process focusing on animal species as a part of integrated system providing data of wildlife migration is proposed.

In computer vision, idea of object recognition process is based on creation a representation of particular classes that characterizing the appearance of objects creating mentioned class. Moreover, this principle can be applied to classification an unknown objects to known class. Success rate of object recognition depends especially on good object representation. Moreover, object representation depends on good object characterization. Object characterization can be achieved by visual descriptors, shape descriptors or texture representation [1-3].

The paper is organized as follows: in the second part, the detailed object recognition process is presented. The third part is related to key points detection and descriptions following the classification part. Finally, in the fifth part the experimental results are described and discussed.

2. Object recognition process

Object recognition process is shown in Fig. 1 and can be divided into two parts: training and testing part. Task of training part is to create a classification model from the training data. Training data contain a collection of images of each class. The extraction of primary images features are extracted at their low-level by different methods. Most common used methods are

SIFT (Scale Invariant Feature Transform), SURF (Speeded Up Robust Features), OpponentSURF, OpponentSIFT etc [1, 2]. These methods will be described in detail in section 3. Moreover, a low-level features extracted from images are used to creation a classification model.

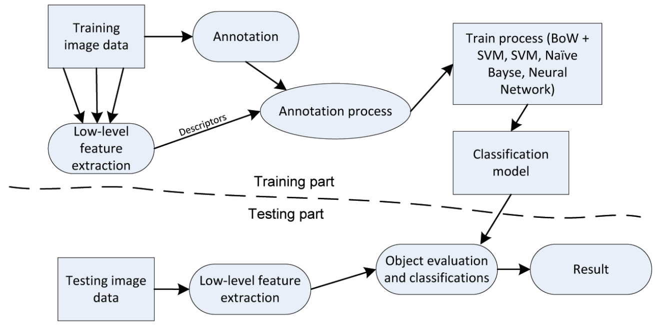

Fig. 1. Principle of object recognition process

To the input of the testing part the enter images and their still picture objects designated to the classification. Moreover, these objects have the same metadata description like data in training part. Based on these data, the classifier is able to regard the classification model successfully evaluate unknown objects to the appropriate class.

3. Visual descriptors

Visual descriptors are used to capture the local appearance of objects. They are calculated from the neighbor pixels. Visual descriptors need to be discriminative enough to distinguish a large number of object classes. Some of them are visually similar and they need to have also invariance to noise, changes of illumination and viewpoints [4]. Each visual descriptor consists of two parts: detector and descriptor.

3.1 Key points detectors

Task of detector is to find key points in the image. There are many methods to detect key points. In this part SIFT and SURF methods for key points detection will be described as well as two proposed hybrid methods SUSIFT (SUrf-SIFT) and SISURF (SIft-SURF).

SIFT: the difference of Gaussians operator is applied to an image at different scales to identify features of potential interest – key point. Then the precise position of key points is dedicated [4, 5].

SURF: detector is based on the determinant of the Hessian matrix. The discriminant value is used to classify the maximum and minimum of the function by second order derivative test [4, 7].

SISURF: Hybrid SISURF method is the key points method detection using SURF detector assuming that in the key point neighbourhood at least one key point detected by SIFT detector is presented. SISURF key point is valid when (1) is true:

$$\min_i\left(\sqrt{\left(x_{KP\_SURF_i}-x_{KPS\_SIFT}\right)^2+\left(y_{KP\_SURF_i}-y_{KPS\_SIFT}\right)^2}\right) \leq \frac{\sum_{j=0}^{n}\min_j\left(\sqrt{\left(x_{KP\_SURF_j}-x_{KPS\_SIFT}\right)^2+\left(y_{KP\_SURF_j}-y_{KPS\_SIFT}\right)^2}\right)}{n} \quad (1)$$

where x_{KP_SURFi} and y_{KP_SURFi} are x and y coordinates of ith SURF key point, i = 0,1, ... n, where n is number of SURF key points, and x_{KPS_SIFT} and y_{KPS_SIFT} are coordinates of all SIFT key points.

SUSIFT: Hybrid SUSIFT method is the key points method detection using SIFT detector assuming that in the key point neighbourhood at least one key point detected by SURF detector is presented.

3.2 Key points descriptors

Task of key points descriptor is to describe key point by the n-dimensional feature vector. In this paper were used these descriptors: SIFT, SURF and Opponent colour descriptors.

SIFT: is the most widely used local visual descriptors. It has reasonable invariance to changes in illumination, rotation, scaling, and small changes in viewpoints. The SIFT descriptor of key point is obtained by first computing the gradient magnitudes and orientations of pixels in the neighborhood region of the key point, using the scale of the key point to select proper Gaussian kernel to blur the image. The orientation of histograms within the sub-regions around the key point are computed and combined into 128 dimensional SIFT feature vector. Produced vector is normalized to improve the invariance to changes of illumination. More detailed information about SIFT can be found in [4, 5, 6].

SURF: Results of SURF descriptor is feature vector of length 64 and is invariant to rotation, scale, brightness and after reduction to unit length or contrast. More detailed information about SURF can be found in [4, 7].

Opponent Color Descriptors: Opponent colour descriptors (OpponentSIFT and OpponentSURF) describe all of the channels in the opponent colour space using common descriptors. Opponent histogram is a combination of three 1D histograms based on the channels of the opponent colour space and this space is given by (2):

$$\begin{pmatrix} O_1 \\ O_2 \\ O_3 \end{pmatrix} = \begin{pmatrix} \frac{R-G}{\sqrt{2}} \\ \frac{R+G-2B}{\sqrt{6}} \\ \frac{R+G+B}{\sqrt{3}} \end{pmatrix} \quad (2)$$

where O1 and O2 contain red-green and yellow-blue opponent pairs and describe the color information in the image. The intensity information is represented by channel O3. R, G and B are channels of RGB colour space: red, green and blue. All three channels are described by SIFT or SURF and therefore they are called OpponentSIFT and OpponentSURF descriptors [8].

4. Animal classification

The collection of features or parameters characterizing the object by classifications methods to handle classification task are used. There are two phases of creation 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 regarding to classification model parameters successfully evaluate an unknown objects to the appropriate class [9, 10, 11]. In this work, for classification model combination bag of keypoints and Support Vector Machine (SVM) methods are used.

4.1 Bags of keypoints

Classification method called bags of keypoint is based on vector quantization of affine invariant visual descriptors of object in images. The main advantages of this method are their simplicity, computationally efficiency and invariance in affine transformation and change in illumination. The main steps of this method are:

- description of the object in images for a set of labeled training data collection,
- constructing a set of vocabularies using K-means algorithm,
- extracting bags of keypoints for these vocabularies,
- applying and training multi-class classifier using the bags of keypoints as features vectors [12].

4.2 Support vector machine

A SVM is a classification method belongs to the family of supervised learning methods that analyze data and recognize patterns. It is non-probabilistic binary linear classifier. SVM belongs to the group of model based classifiers. Training algorithm constructs the model that represents patterns as points in vector space. Task of SVM classifier is found an optimal hyperplane with maximum margin between data of two different classes. Development of the classification system includes separating data into training and testing sets. To separate data of different classes, SVM maps feature vectors into a higher dimensional space using a kernel function [13, 14]. In this work, radial basic function (RBF) kernel was used.

5. Experimental results

Training database consists of 5 classes: wild boar, brown bear, wolf, fox and deer. The examples of images from training database are shown in Fig 2. 10 images per class were randomly chosen from training database and were used as test database.

Tested method follows principle scheme of object recognition process shown in Fig.1. First, the low-level features from training images were extracted. In the next step, the extracted descriptors together with annotation record in order to create a representation of particular class were used.

These data enters the process of constructing vocabulary using k-means clustering algorithm. In [8] was proved, that number of cluster equal 1000 present a good trade-off between accuracy and speed. Then, bag of key points for vocabulary were extracted. To extract bag of key points, algorithms for matching training descriptors with cluster centre in vocabulary were used. For each feature data extracted from test image by selected descriptor, BruteForce matcher finds a cluster centre in vocabulary. To the designation of feature vector and cluster centre distance, the Euclidean distance was used. Similar approach how to find out the minimum distance of feature vector and cluster centre is called FlannBased matcher. Thus, extracted bag of keypoints for SVM classifier serve to creation a classification model for particular classes were used.

Fig. 2. The images from training database

In the experiment, a total 4 key point detectors, namely, SURF, SIFT, SUSIFT and SISURF were used. Moreover, to describing a key point by four descriptors: SIFT, SURF, OpponentSIFT or OpponentSURF and two matchers: Brute Force or Flann Based were used too. All combinations of detectors, descriptors and matchers were combined into standalone runs and they were programmed in C++ language with support of OpenCV (Open source Computer Vision) library. In the clustering process, 15.000, 20.000, and maximum descriptors per class were chosen to construct the vocabulary. Moreover, for training classifier, 15.000, 20.000, and maximum extracted bags of keypoints were used.

Average score of animal classification for combination SIFT descriptor, SIFT, SURF, SISURF, SUSIFT detectors, two matchers and variable number of descriptors used in clustering process is shown in Fig. 3. In Fig.4 is shown average score of animal classification for combination SURF descriptor, SIFT, SURF, SISURF, SUSIFT detectors, two matchers and variable number of descriptors used in clustering process.

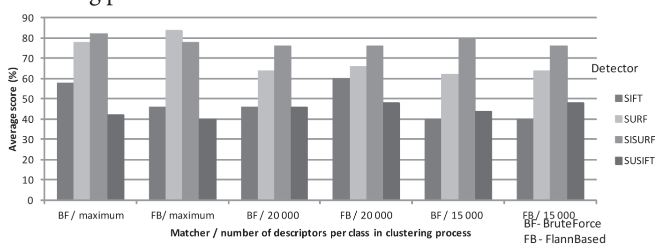

Fig. 3. Average classification score for SIFT descriptor

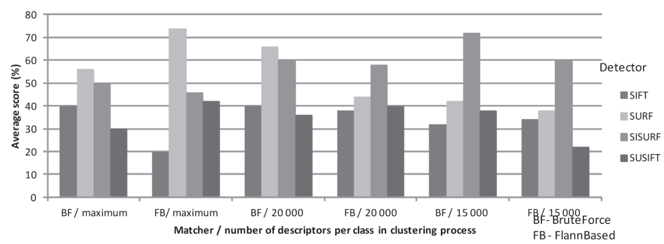

Fig. 4. Average classification score for SURF descriptor

Average score of animal classification for combination OpponentSIFT descriptor, SIFT, SURF, SISURF, SUSIFT detectors, two matchers and variable number of descriptors used in clustering process is shown in Fig. 5.

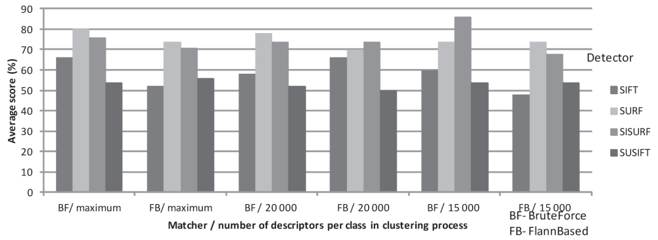

Fig. 5. Average classification score for OpponentSIFT descriptor

In Fig.6 it is shown the average score of animal classification for combination OpponentSURF descriptor, SIFT, SURF, SISURF, SUSIFT detectors, two matchers and variable number of descriptors used in clustering process.

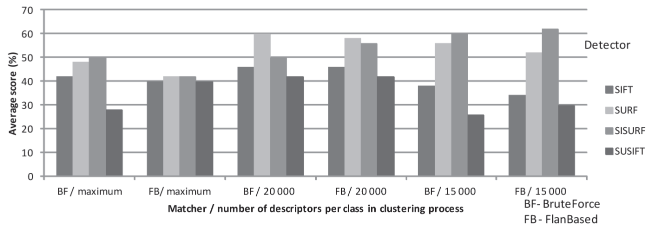

Fig. 6. Average classification score for OpponentSURF descriptor

Results per each class for the best run with average score 86 % are shown in Table.1.

Table 1. Confusion matrix for the best run, combination SISURF detector, OpponentSIFT descriptor, BruteForce matcher and 15.000 descriptors per class in clustering process

True classes →	Wild	Brown	Wolf	Fox	Deer
Wild boar	7	0	0	1	0
Brown bear	2	10	0	0	0
Wolf	0	0	9	1	0
Fox	0	0	1	7	0
Deer	1	0	0	1	10

5. Conclusion

In this paper, two hybrid key points detectors were presented and with other detectors and descriptors and combination bag of key points and SVM classifier were tested. From the realized experiment it is evident that the highest classification success rate of 86 % was achieved by algorithm in the combination of SISURF detector, OpponentSIFT descriptor, BruteForce matcher and 15.000 descriptors per class in clustering process. Moreover, success rate higher than 80 % was achieved by four more runs. Proposed hybrid key points detector SISURF achieved promising results, comparable with other key point detectors. Moreover, in same runs SISURF outperformed other standard detectors. On the other hand, SUSIFT detector achieved poor results with success rate of classification around at 50% only.

Acknowledgments

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.

Bibliography

[1] SONKA M., HLAVAC V., BOYLE R.: Image Processing, Analysis and Machine Vision, Thomson Learning, part of the Thompson Corporation, ISBN: 10: 0-495-24438-4, 2008

[2] CAO. T.C.: Object Recognition. ISBN 978-953-307-222-7 2011

[3] HOSSAIN K., PAREKH R.: Extending GLCM to include Color Information for Texture Recognition, International Conference on Modeling, Optimization and Computing, AIP Conference Proceedings, Volume: 1298, Pages: 583-588, 2010

[4] NGAN K.N., LI H.: Video Segmentation and Its Applications. Springer Science+Business Media 2011, ISBN 978-1441994813.

[5] NIXON M., AGUANDO A.: Feature extraction & image processing, second edition, ISBN 978-0-1237-2538-7, 2008

[6] UIJLINGS, J. R. R., SMEULDERS, A. W. M., SCHA, R. J. H.: The Visual Extent of an Object, International Journal of Computer Vision, Volume: 96, Pages: 46-63, Jan 2012

[7] BAY T., TUYTELAARS T., LUC V.G.: SURF, Speeded Up Robust Features, ETH Zurich, Preprint submitted to Elsevier, 10 September 2008

[8] KOEN E. A., GEVERS T., CEES 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

[9] ZHANG J., MARSZALEK M.: Local Features and Kernels for Classification of Texture and Object Categories: A Comprehensive Study, Journal of Computer Vision, 2006 Springer Science + Business Media, LLC. Manufactured in the United States

[10] NABER M., HILGER M., EINHAUSER W.: Animal detection and identification in natural scenes: Image statistics and emotional valence, Journal of Vision, Volume: 12, 2012

[11] GROSSBERG S., MARKOWITZ J., CAO, YQ.: On the road to invariant recognition: Explaining tradeoff and morph properties of cells in inferotemporal cortex using multiple-scale task-sensitive attentive learning, Neural Networks, Volume: 24, Pages: 1036-1049, 2011

[12] CSURKA G., DANCE CH. R., FAN L., WILLAMOWSKI J., BRAY C.: Visual Categorization with Bags of Keypoints, Xerox Research Centre Europe, France 2001.

[13] SHUKRAN M. A. M., CHUNG Y.Y, YEH W. CH., WAHID N.: "Image Classification Technique using Modified Particle Swarm Optimization," Modern Applied Science, Vol. 5, No. 5; October 2011

[14] SUPICHAI T.: "Image Classification," Available at: http://www.sc.chula.ac.th/courseware/2309507/Lecture/remote18.htm, accessed 18.08.2013.