2014 AASRI Conference on Circuits and Signal Processing (CSP 2014)

Classification of Wild Animals Based on SVM and Local Descriptors

Slavomir Matuska, Robert Hudec, Patrik Kamencay, Miroslav Benco,
Martina Zachariasova*

Department of Telecommunication and Multimedi, University of Zilina, Univerzitná 8215/, 101026 Zilina, Slovakia

Abstract

In this paper, a novel method for object recognition based on hybrid local descriptors is presented. This method utilizes a combination of a few approaches (SIFT - Scale-invariant feature transform, SURF - Speeded Up Robust Features) and consists of second parts. The applicability of the presented hybrid methods are demonstrated on a few images from dataset. Dataset classes represent big animals situated in Slovak country, namely wolf, fox, brown bear, deer and wild boar. The presented method may be also used in other areas of image classification and feature extraction. The experimental results show, that the combination of local descriptors has a positive effect for object recognition.

1. Introduction

Semantic information about images and videos can be very valuable information in many areas, namely instance search, known-item search, semantic indexing, image retrieval and many others. Especially, animal

---

* Corresponding author. Tel.: +421 - 41 - 513 2245
E-mail address: slavomir.matuska@fel.uniza.sk.

classification and recognition can be used in surveillance systems or recognition systems in cars, to prevent car collision with wild animal.

In many applications, there is a need for real time object recognition, especially in surveillance systems and recognitions systems in cars. Animal recognition and classification can be also used in expert systems for determining of wild animal's migration corridors. It terms of real-time object recognition, SVM classifier was used for its fast testing speed and sufficient accuracies. Success rate of object recognition depends also on good object representation and characterization. Object characterization can be achieved by visual descriptors, shape descriptors or texture representation. In this paper, visual descriptors were used to object representation.

The outline of the paper is as follows. In the second section, an overview of related work is given. In the third section, object recognition process is presented. In fourth section is related to key points detection and descriptions following the classification section. Finally, experimental results are discussed in the sixth section, and it is followed by conclusion in the seventh section.

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 methods are SIFT, SURF, OpponentSURF, OpponentSIFT etc. These methods will be detailed described in section 3. Moreover, a low-level features extracted from images are used to creation a classification model.
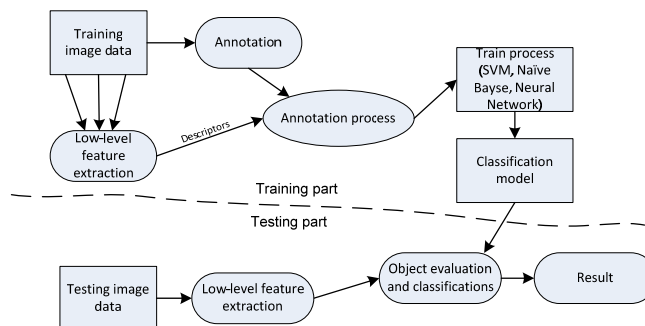

Fig.1. Object recognition process

To the input of the testing part enters an 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 regarding to classification model successfully evaluate an unknown objects to the appropriate class [1], [3].

3. Visual descriptors

Visual descriptors are used to capture the local appearance of objects. They are calculated from the neighbour 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 [2]. 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 will be described SIFT and SURF methods for key points detection and 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 [6].

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 [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. SISURF key point is valid when (1) is true:

$$
\min_i \left(\sqrt{\left(x_{KP\_SURF_i} - x_{KPS\_SIFT}\right)^2 + \left(y_{KP\_SURF_i} - y_{KPS\_SIFT}\right)^2} \right) <
$$

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

(1)

where x_{KP_SURFi} and y_{KP_SURFi} are x and y coordinates of i-th 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. Examples of valid SISURF key points are shown in the Fig 2.
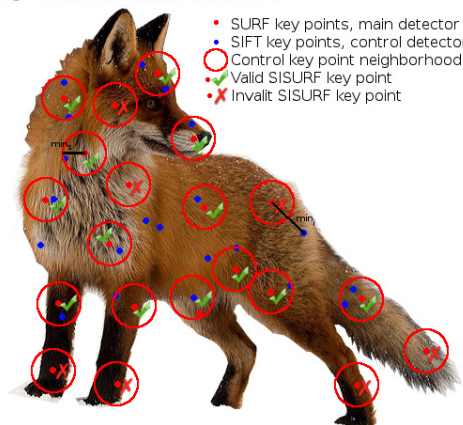

Fig.2. Example of SISURF detector

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.

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.

Information about SIFT descriptor can be found in [6]. SURF descriptor was described in detail in [7] and information about Opponent color descriptors can be found in [2], [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 [4], [5]. In this work, for classification model combination bag of keypoints and Support Vector Machine methods are used.

4.1. Bags of Keypoints (BOW)

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 [9].

4.2. Support Vector Machine

A SVM is classification method related to the family of supervised learning methods. There are two data types used in SVM classifier. To create a classification model, training data are used. To test and evaluate trained model accuracy, testing data are used. The main SVM classifier task is to separate training data in the higher dimensional space using a kernel function and find an optimal hyperplane with a maximum margin between data of two different classes [10]. 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 3. 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. 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.

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 all descriptors per class were chosen to construct the vocabulary. Moreover, for training classifier, 15.000, 20.000, and all extracted bags of keypoints were used.
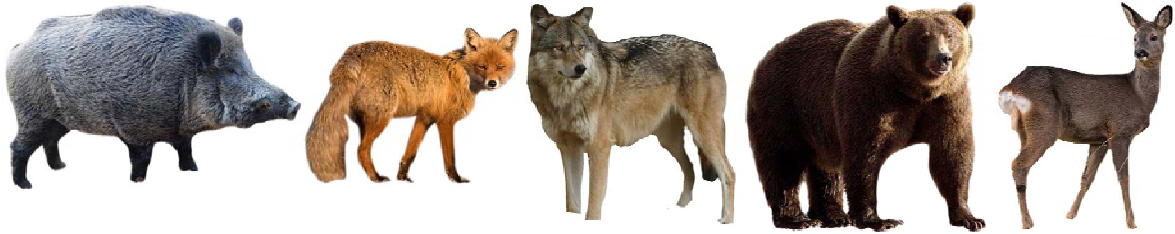

Fig.3. The images from training database

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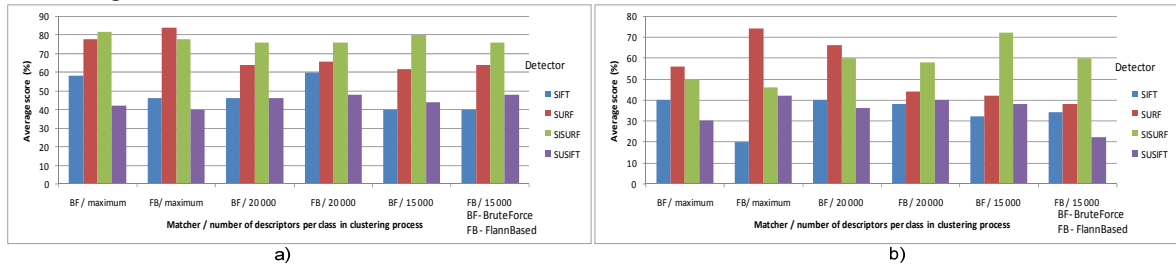

Fig.4. a) Average classification score for SIFT descriptor, b) Average classification score for SURF

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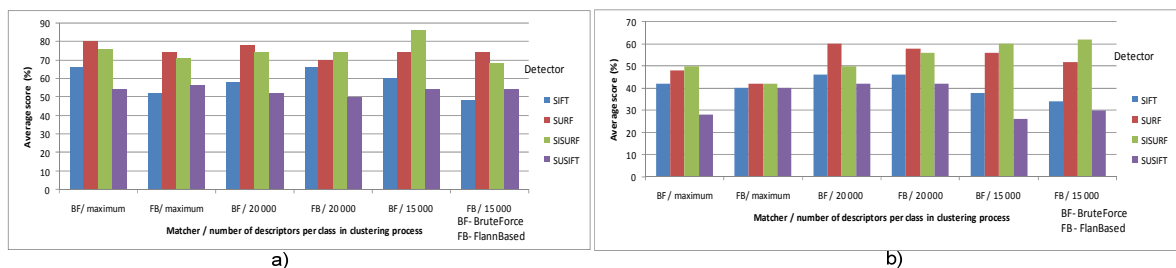

Fig.5. a) Average classification score for OpponentSIFT descriptor, b) Average classification score for OpponentSURF descriptor

6. Conclusion

In this paper, two hybrid key points detectors were presented and tested in comparison to other detectors. The combination of BOW and SVM classifier were experimented. Experiments showed, that highest classification success rate 86% was achieved by algorithm based on combination SISURF detector, OpponentSIFT descriptor, BruteForce matcher and 15.000 descriptors per class in clustering process. Moreover, success rate higher than 80 % was achieved in four other 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 50% only.

Acknowledgements

References

[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 J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73.
[2]Ganesharajah B., Mahesan S, Pinidiyaarachchi U. A. J., 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.
[3] Islam M. K., Jahanm F., Min J, Baek j., Object classification based on visual and extended features for video survivalence application, Control Conference (ASCC), 2011 8th Asian, 15-18 May 2011, Page(s): 1398 – 1401, Print ISBN: 978-1-61284-487-9
[4] Lotz A., Steck A., Schlegel Ch., Analysing solution quality of anytime Bag of Words object classification for a service robot, Technologies for Practical Robot Applications (TePRA), 2012 IEEE International Conference on, 23-24 April 2012, Page(s): 1-6, Print ISBN 978-1-4673-0855-7.
[5] Wilber M.J., Scheirer W.J., Leitner P., Heflin B., Zott J., Reinke D., Delaney D.K., Boult T. E., 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
[6] Ngan K.N., Li H., Video Segmentation and Its Applications,   Springer Science+Business Media 2011, ISBN 978-1441994813.
[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., 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] Czurka G., Dance CH. R., Fan L., Willamowski J., Bray C., Visual Categorization with Bags of Keypoints, Xerox Research Centre Europe, France 2001
[10] 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.