Optik 127 (2016) 4736–4740

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

A new algorithm for image recognition and classification based on improved Bag of Features algorithm

Kangshun Li, Fubin Wang*, Lixia Zhang

College of Information, South China Agricultural University, Guangzhou 510642, China

ARTICLE INFO

Article history:
Received 11 September 2014
Accepted 29 August 2015

Keywords:
Bag of Features
Image recognition
Classification
SURF
Spatial pyramid matching

ABSTRACT

This paper presents a method for image recognition and classification based on improved Bag of Features (BOF). In view of the low efficiency and low classification accuracy of the traditional BOF algorithm, a new recognition and classification algorithm combined speeded-up robust features (SURF) and spatial pyramid matching principle is proposed in this paper. SURF algorithm can improve the efficiency and spatial pyramid matching principle can improve the classification accuracy. This paper's method uses SURF algorithm to extract the image feature and generate the codebook. The spatial pyramid matching principle is applied to the image histogram's codebook which can improve the accuracy of the classification. Finally, the method uses the LIBSVM classifier to classify the image histogram's codebook. The experiments are carried out based on Graz, Caltech-256, and Pascal VOC 2012. The results show that our proposed method is better than the traditional method in the efficiency and classification accuracy. In addition, our method is compared with some related research work in classification accuracy, and the results show that our algorithm has obvious advantages.

1. Introduction

With the development of multimedia technology, the image has turned into mainstream visual information. Image recognition and classification has been a popular research subject in the image processing and the application area. With the development of globalization and emergence of various kinds of portals, there is a drastic increase of images. Therefore, it brings great challenges to the image recognition and classification. How the users can accurately get the information in the big data of images is one of the issues to be solved in the computer vision's field. At present, there are many mature image recognition and classification algorithms, and BOF algorithm is a popular one among them. BOF algorithm came from the Bag of Words (BOW) [1,2] and BOW algorithm was originally used for documents matching. In recent years, Csurka et al. [3] transplanted the idea of the algorithm to the field of image processing [4] through the feature extraction algorithms, such that the SIFT convert the images to feature descriptor and get a lot of features to represent keywords. And then using the clustering method (e.g. K-Means) to get the cluster of the words, the cluster center is defined as visual words and K-Means algorithm derived all cluster centers constitute a visual dictionary. Finally, the images are classified by SVM classifier. There are some researchers applying the algorithm to the behavior recognition [5] because it can improve the recognition accuracy rate. And there are some researchers applying the algorithm to the food recognition system [6] which is for diabetic patients. In recent years, BOF algorithm is also been applied in the medical field and has achieved results [7,8].

However, in the BOF algorithm there exists some problem in the application of image recognition and classification such as run is not fast enough and the classification accuracy is not high enough. So it needs to be optimized. The traditional BOF algorithm uses SIFT (scale-invariant feature transform) [9] feature descriptor, K-Means clustering method and SVM (support vector machine) classifier to achieve its purpose. Due to the SIFT algorithm not being robust to extract the feature of the complexity background, in order to improve the robustness of the algorithm [10], we proposed a method based on a partial model, but this method is relatively highly complex. BOF algorithm classification accuracy is not high enough when it just simply applies K-Means clustering and SVM classifiers [11], using probability by Expectation Maximization (EM) algorithm to fit the BOF model; although to some extent it improved the classification accuracy, they remain efficient and classification accuracy of the complexity background image are not high enough.

Aiming at the shortcomings of BOF algorithm, as well as lack of existing research, this paper uses an improved algorithm speeded up robust feature (SURF) [12] to extract image feature descriptor in

* Corresponding author. Tel.: +86 15018733851.
E-mail address: waterbinbin@126.com (F. Wang).

http://dx.doi.org/10.1016/j.ijleo.2015.08.219
0030-4026/© 2016 Published by Elsevier GmbH.

order to reduce the complexity of the BOF algorithm and improve operational efficiency and solve the problem of high storage requirements. Meanwhile, introducing spatial pyramid matching principle to represent histogram based on visual dictionary improves the classification accuracy of the BOF algorithm. Finally, in this paper, a fast and effective classifier (LIBSVM classifier) [13] to improve the efficiency of classification is used.

2. Improved BOF algorithm

In this paper, BOF algorithm has three steps: feature extraction and feature descriptor transformation; build visual dictionary; recognition and classification. Fig. 1 shows the basic flow of BOF algorithm.

According to Fig. 1, the process of BOF can be summarized as follows:

Firstly, use SURF algorithm instead of the traditional SIFT algorithm and use the Hessian matrix to detect the feature points, each with a 64-dimensional feature vector representation, extracted from the image of a representative of the global and local features as an image descriptor. Then, using the K-Means clustering method (this paper does not specify the cluster center), our method selects cluster centers randomly, through a limit number of iterations, and finally we get the cluster center which is the dictionary. The traditional method is represented as a statistical histogram shown by the image in Fig. 2. The article refers to the space pyramid matching scheme to improve the accuracy of classification. Finally, the LIBSVM classifier is used to classify.

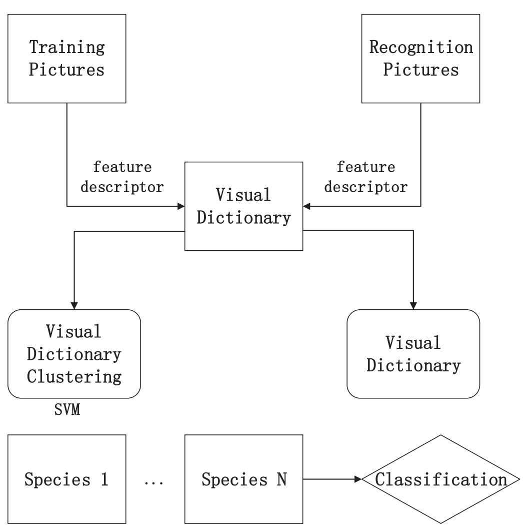

Fig. 1. The basic flow of BOF algorithm.

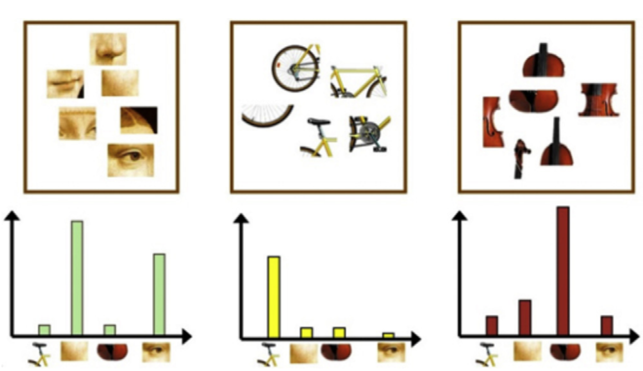

Fig. 2. The histogram based on visual words.

2.1. SURF algorithm

To extract the feature at BOF algorithm, there are commonly used algorithms such as SIFT, Gradient Location-Orientation Histogram (GLOH), and PCA-SIFT [14]. SIFT algorithm is an invariant feature detection technique based on scale space, image scaling, rotation, affine transformation even illumination changes remain invariant image local characterization methods. This method must calculate the amount of data and use high time complexity. Aiming at the shortcomings of BOF algorithm, this paper uses SURF to extract image feature descriptor in order to reduce the complexity of the BOF algorithm and improve operational efficiency.

The basic flow of SURF algorithm includes feature detection (build Hessian matrix, build scale space, precise positioning of feature points), which determines the main direction and generates feature point descriptor.

2.1.1. Feature point detection

Given a point $X = (x, y)$, the Hessian matrix at scale σ is defined as follows:

$$(1) H(X,\sigma) = \begin{bmatrix} L_{xx}(X,\sigma) & L_{xy}(X,\sigma) \\ L_{xy}(X,\sigma) & L_{yy}(X,\sigma) \end{bmatrix}$$

In formula (1), L_{xx} is the convolution of the Gaussian second order derivative $\partial^2/(\partial x^2) g(\sigma)$ with the image $I = (x, y)$, and $g(\sigma) = (1/2\pi\sigma^2)e^{-(x^2+y^2)/2\sigma^2}$, and similarly for L_{xy} and L_{yy}.

In this paper, box filters were used instead of the second order Gaussian filter, and used the calculation which is unrelated with the template of the integral image [15] in order to increase computing speed and accelerate convolution. In Fig. 3, the 9×9 box filters are approximations of a Gaussian with $s = \sigma = 1.2$ and represent the lowest scale. The value of D_{xx}, D_{yy}, D_{xy} represent the value of box filters and image convolved, instead of the value of L_{xx}, L_{yy}, L_{xy}.

In this paper it uses the experience value of 0.9 in the weight of the matrix, and then obtains the Hessian matrix det expression:

$$(2) \det(H_{\text{approx}}) = D_{xx}D_{yy} - (0.9D_{xy})^2$$

The first column is the Gaussian second order partial derivative in y-direction to approximation for second order Gaussian partial derivative in y-direction; the second column is the Gaussian second order partial derivative in xy-direction to approximation for the order Gaussian partial derivative in xy-direction.

To achieve scale invariance of feature points, this paper uses an image pyramid to achieve scale space by Gaussian convolution with

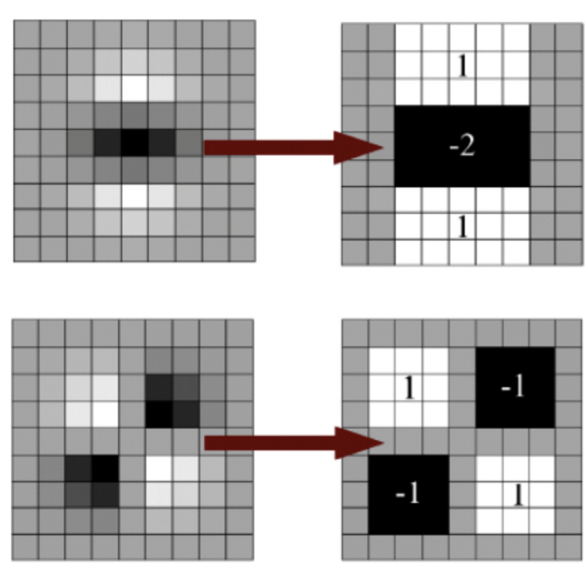

Fig. 3. Box filter approximations for second order partial derivative.