SVM-KNN Algorithm for Image Classification Based on Enhanced
HOG Feature

Zhaquan Cai1, Pan Yu2, Yihui Liang2*, Bin Lin2, Han Huang2
1Huizhou University, Huizhou, 516007, P.R. China
2 School of Software Engineering, South China University of Technology, Guangzhou, 510006, P.R. China
*corresponding author: Yihuiliang@outlook.com

Abstract

After decades of research, a large number of ways have already been applied in image classification and achieved good results. However, they still suffer from such problems not keep the details of the image well when extraction image feature and, the classifier is not efficient enough and so on. We propose an image classifier algorithm based on the enhanced HOG feature. The approach can use HOG to describe the local features of the object properties and strengthen them, at the same time, it combines the excellent performance of the SVM and KNN algorithm to make classify effects more prominent. Experimental results show a considerable improvement practicality of the image classification algorithm based on Enhanced HOG Feature.

Keywords: image classification, HOG, SVM-KNN

1. Introduction

Image classification plays an important role in many applications, such as monitoring, diagnosis and medical image retrieval. Many computer vision problems can be redefined as an image classification problem. Thus, image classification is a vital step of multi-media content analysis. In the recent decades of years, many advanced multimedia content analysis technology has been widely used. Such as clustering, decision tree method, hidden Markov models, and neural networks.

Support vector machine proposed by Corinna Cortes et al. (3) is most commonly used in image classification. In applying this method, the data set is often a huge number, and very high dimension. If we place the image matrix into classifier directly, the classification cannot run well because it’s time-consuming and involved in complex high operation, therefore principal component analysis (PCA) is usually used to dimension reduction. However it’s linear mapping method, unable to resolve the relationships of image internal structures well.

Zhang et al. proposed a combination of SVM and nearest neighbor, new SVM-KNN method in the literature (4). The main idea of SVM-KNN method is to look for the nearest neighbor of test sample and trained local support vector machines which retained function of distance among the neighbor set. In practice, compared with SVM, accuracy of this method definitely improved and had a good performance, especially when applied in large scale multi-type data sets. However, due to the method of image feature extraction and not put too much energy, so that the performance of the method still leaves some room for improvement.

Dalal et al. proposes a feature description methods HOG, based on statistical properties of the local gradient in the literature (5). Its main idea is that, in an image, appearance and shape of the target can be local gradient or edge direction density distribution well described. O. Ludwig et al.,(6) applied this method to detect pedestrians, showing a fine performance, but did not work on image classification application research furtherly.

He et al. (9) put forward a method of expressing HOG in scale space; Wang et al. (10) proposed the combination of LPP (Local Preserving Projection) and HOG. In spite of the good performance in pedestrian’s detection, they are not developed in image classification.

We inevitably encounter the following two problems no matter what the method is: firstly, how to extract the structural information perfectly, secondly, how to find a higher classification accuracy of the classifier. Therefore, this paper aims to build a feature expression vector to retain image structure as much as possible, and seek an excellent classifier to classify.

Based on HOG (Histograms of Oriented Gradients) (3)
feature extraction methods and to improve it, we propose a new image classification method using SVM-KNN as classifier. Enhanced HOG feature extraction method retains the structural information of image effectively; SVM-KNN optimizes the classification performance of the algorithm within an acceptable time complexity greatly. Thus, a new image classification algorithm is proposed and shows prominent practical effects. Furthermore, we improved HOG feature extraction methods and used it on some common classification for reference and comparison. Simultaneously, providing some meaningful and practical ideas for image classification research.

2. The Proposed Algorithm

The structure information is positive to image classification. Therefore, we design a strategy of enhanced HOG feature to mine the structure information of images. The enhanced HOG feature is obtained by processing the image with CLAHE, subtraction and enhancing strategy.

2.1 Using CLAHE Method to Enhance the Image Feature

The basic idea of CLAHE is to divide the image to several non-overlapping regions. Every block is processed by histogram equalization, and the neighbor blocks are optimized by bilinear interpolation to reduce the blocky effect. We use a modified CLAHE (MCLAHE) as the initialization of the enhanced HOG feature. The procedure of MCLAHE is presented by Algorithm 1.

Algorithm 1: MCLAHE

Input: an tackled image $G_0$

Step 1: Divide the image $G$ into $m \times n$ regions of equal sizes. Each region is denoted as $R[i, j]$.

where $i = 1, 2, ..., n$ and $j = 1, 2, ..., m$.

Step 2: // histogram equalization

for i=1 to n do

for j=1 to m do

Calculate the histogram of $R[i, j]$.

Equalize the histogram of $R[i, j]$.

Step 3: Use a bilinear interpolation to produce a new image.

Output: a new image $G_1$

Step 2 of Algorithm 1 contains two main operators. First, the histogram of each region is produced. After that, the histogram is equalized by the following model. Let $L$ be a limit threshold, $g(k)$ be the number of pixels whose gray level is $k$ and $N_L$ be the cardinality of the set \( \{ x \mid g(x) > L \} \). For $k = 1, 2, ..., 255$, the operator will make the histograms satisfy Expressions (1)-(2).

\[
N_L = \begin{cases} 
N_L - (L - g(k)) & g(k) \geq L - N_L / L \\
N_L - N_L / L & g(k) < L - N_L / L 
\end{cases} \quad (1)
\]

\[
g(k) = \begin{cases} 
L & g(k) \geq L - N_L / L \\
g(k) + S / L & g(k) < L - N_L / L 
\end{cases} \quad (2)
\]

Step 3 is to use a bilinear interpolation to calculate the gray value of each pixel. A new image is generated after MCLAHE is implemented. Furthermore, the image has more structure information to strengthen the HOG feature than the original one.

2.2 Removing the Background by Subtraction Method

After the structure information is added to the tackled image by MCLAHE, the background will be removed by a hybrid subtraction method. We design the method by hybridizing the frame subtraction method with the low-pass Gaussian filter to obtain a high-quality HOG feature.

The hybrid method includes several steps. The Gaussian filter is used to initialize the image $G_1$. First, $G_1$ is processed by the low-pass Gaussian filter. Second, the gradient of every pixel is calculated to produce a new matrix of image which is denoted by $I$. Finally, a new image $G_2$ is generated by Expression (3).

\[
G_2(x, y) = \begin{cases} 
0 & |I(x, y) - G_1(x, y)| < \delta \\
G_1(x, y) & |I(x, y) - G_1(x, y)| \geq \delta 
\end{cases} \quad (3)
\]

The motivation of the hybrid strategy is to make the local feature and shape of the image more obvious than the
2.3 Enhancing HOG Feature

After the subtraction method, we use an enhancing HOG feature to remain the detailed feature of the image as far as possible. The basic idea of HOG is to show the local feature and shape of the image object by the density distribution of gradient and edge.

![Diagram of image, blocks and cells](image)

Figure 1. The relation of image, blocks and cells

The process of HOG abstracting includes several steps. First, the tackled image is divided into $N \times N$ local regions named blocks, and the blocks are divided into smaller regions named cells. Their relation is shown by Figure 1. We use a center symmetrical operator to calculate the gradient in horizontal and vertical directions. Let $g_x(x, y)$ and $g_y(x, y)$ be the horizontal and vertical gradients of the pixel $(x, y)$. The calculation is implemented by Exp. (4).

$$g(x, y) = \sqrt{g_x(x, y)^2 + g_y(x, y)^2}$$  \hspace{1cm} (4)$$

where $g_x(x, y) = [-1, 0, 1] \cdot G(x, y)$ and $g_y(x, y) = [-1, 0, 1]^T \cdot G(x, y)$. The gradients direction is calculated by the expression that $\theta(x, y) = \arctan\left(\frac{g_y(x, y)}{g_x(x, y)}\right)$. Second, the image is divided into $n \times n$ cells, and the gradient direction is divided into $k$ bins on average. The histogram of every direction is added up to produce a k-dimension vector for all the pixels of every cell. Third, the histogram of each blocks is normalized by L2-norm. The HOG feature of the overlapping blocks is collected to form a high-dimension feature vector by a step of $n$.

2.4 KNN-SVM Hybrid Classification Algorithm for Image

After enhancing HOG feature, we use KNN and SVM to design a hybrid classification algorithm for image. The algorithm is shown by Algorithm 2. The first step produces an enhanced HOG feature denoted by a vector. The vector is used to make $K$ classification labels of the tackled image by the other steps.

### Algorithm 2: KNN-SVM Hybrid Classification Algorithm

**Input:** a training set of images, a tackled image and a parameter $K$

**Step 1:** For every image of the training set and the tackled image do

- Implement Algorithm 1 to produce the HOG-feature vectors.

**End For**

**Step 2:** Calculate the distances between the vector of the tackled image and the vectors of the training set.

**Step 3:** Select the $K$ neighbors.

**Step 4:** If the labels of the $K$ neighbors are the same then

- Let the label of the tackled image be the label of the $K$ neighbors.

Else

- Use the two-vector distances of the $K$ neighbors to calculate a kernel function.
- Apply the SVM algorithm with the kernel function to generate the classification label of the tackled image.

**End If**

**Output:** the classification label of the tackled image

3. Numerical Results

In this section, we present three experimental results to show how the proposed enhanced HOG feature is helpful for image classification. The first experiment is to show whether the enhanced HOG feature can improve the face-image classification by KNN methods. The second one is to confirm whether the enhanced HOG feature is generally efficient to other algorithms for image classification. The last one is to investigate how the parameters of the enhanced...
HOG feature impact on the performance of the proposed algorithm.

3.1 Experimental data set

In the experiment, we chose two images database for test data.

The first is the Asian face database PF01 (8), it was created by the South Korean Pohang University of Science and Technology, contains 103 individual photos (53 males, 50 females), each have 17 different images (one usually face, four faces under different light, eight faces at different positions, four faces with different expressions). The people in database are all Asian, mostly Korean. The size of each image is 256 * 256 pixels. In the test problems, we take everyone's previous nine images as the training sample, late eight images as a test sample, as shown in Figure 6. We will carry out the test of improving HOG performance using them.

The second is Amsterdam Library of Object Images (ALOI), ALOI database(15) is a collection of colorful images of 1000 objects, each object contains more than 100 images, and the entire database contains 110,250 images. Different images of the same object are obtained by changing the viewing angle of the observer, the color and angle of the light. Sample image shown in Figure 7. We will test the performance differences between new algorithm and other algorithms with this database. In the experiment, we randomly picked 200 objects, each object randomly picked 50 images, in which 20 training samples, 30 as a test sample.

In the method involving HOG, we adopt default parameter settings(3); this paper uses some default parameter settings: preprocess like Gamma correction is not used; gradient calculation using the simplest symmetry center operator[−1 0 1]; the size of sampling window (i.e., block) is 8 × 8, the size of cell is 4 × 4 pixels; using standardized methods of L2-norm.

Test program is written with Matlab R2012, the running platform is 64-bit Windows 7, the machine processor is Intel Core Intel Core i5 CPU M460 @ 2.53GHz, memory is 4GB (3.68GB available).

3.2 Comparison of experimental results

The test results are reflected by the accurate rate of recognition. During the test process, the program will predict each category an image belongs to, and compare it with its real category, if they are the same, the recognition is considered correct. The accurate rate of recognition is the correct image recognition divided by the total number of test samples.

Table 1 illustrates when K is assigned with different values, how KNN, HOG-KNN, IHOG-KNN recognition results vary. It turns to be that, as K increases, an overall recognition rate tends to decrease gradually, combining feature extraction methods base on HOG makes recognition accuracy significantly improved, and the application of enhanced HOG makes the accuracy rate of image classification further improved. The results illustrate as for this data set, the accuracy rate of image classification is higher when K value is small, it will increase when K reduces, and the improvement of experiment for HOG is apparently effective.
Table 1 Accuracy rate of recognition using PF01 face image database

<table>
<thead>
<tr>
<th></th>
<th>KNN</th>
<th>HOG-KNN</th>
<th>IHOG-KNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50.10%</td>
<td>82.80%</td>
<td>87.40%</td>
</tr>
<tr>
<td>2</td>
<td>41.20%</td>
<td>76.40%</td>
<td>87.40%</td>
</tr>
<tr>
<td>3</td>
<td>39.50%</td>
<td>76.20%</td>
<td>86.20%</td>
</tr>
<tr>
<td>4</td>
<td>38.40%</td>
<td>73.10%</td>
<td>81.50%</td>
</tr>
<tr>
<td>5</td>
<td>39.10%</td>
<td>71.90%</td>
<td>79.00%</td>
</tr>
<tr>
<td>6</td>
<td>37.90%</td>
<td>68.80%</td>
<td>76.40%</td>
</tr>
<tr>
<td>7</td>
<td>36.20%</td>
<td>66.60%</td>
<td>73.80%</td>
</tr>
</tbody>
</table>

Table 2 shows accurate identification rate of the image classification when ALOI database uses different methods. The results show that Several methods of face recognition algorithm based on enhanced HOG features extraction and mainstream classifier have achieved good recognition results, the methods based on enhanced HOG feature extraction has higher performance than the original method.

3.3 Experimental results under different parameter settings

We used the "mean filter", "circular area mean filter", "Gaussian low pass filter", "Motion Blur operator", four kinds of filter operator to test, image classification methods for test are the best-behaved when using ALOI database. Test results are obtained through the application of IHO-G-SVM-KNN (K = 10).

Data in Table 3 shows that the Gaussian low-pass filtering algorithm enables the operator to achieve the highest recognition accuracy rate.

During the process of HOG improvement, the image enhancement (using adaptive histogram equalization given in this paper), in order to explore the effect of different image enhancement techniques, we used the "gray-scale transformation," "histogram equalization", "adaptive histogram equalization" to do the experiment, image classification methods for test are the best-behaved when using ALOI database. Test results are obtained through the application of IHO-G-SVM-KNN (K = 10).

Data in table 4 shows that the use of "adaptive histogram equalization" can achieve the best image enhancement and achieve the most accurate recognition. We can conclude through experiment 1-3: HOG can keep internal structure of images, HOG for the feature extraction is already a good choice. However, enhanced HOG will further enhance the performance of feature extraction method. In the process of improvement for HOG, the use of adaptive histogram equalization to enhance the image achieves the most satisfactory results, and the choice of filter in the use of “difference image method” is not fixed, it needs to be adjusted according to data circumstances.

3.4 Experimental results analysis

The reason of results appear in Table 1 is that HOG operation on the local cell of the image, better able to save the outline feature of the image, the geometric and optical distortion of image can keep good invariance. Thus, HOG as a feature description method than the traditional way of dimensionality reduction, shows more excellent effect. The enhanced HOG is improved to enhance the image so that the features of image turn to be more obvious, in addition to the removal of unrelated background factors interference, so its higher performance than HOG is expected.

The reason of results appear in Table 2 is that KNN, SVM itself is relatively good classifier, it has been fully reflected during decades of research in the field of image classification. However, KNN uses only individual samples of training sample space, and does not use distance information between the test sample and the training samples, it seems not comprehensive enough, but SVM-KNN takes full use of these information.

The reason of results appear in Table 3-4 is that adaptive histogram equalization method considers the local features of the image better than other methods so that achieves the best reinforcing effect of image. Practice has proved that the combination of enhanced HOG feature extraction and SVM-KNN has a high practical value. Its high recognition rate and a good processing speed makes it can occupy a place among image classification activities in the future. Therefore, the proposed method can be extended in everyday applications. Meanwhile, the idea of combining enhanced HOG with popular classifier has a certain guiding significance for new classification algorithm in the following days.

4. Conclusion

This paper propose an original method, which combines SVM-KNN with HOG and improves HOG, effectively raised the accuracy of the image classification, and achieved excellent experimental results. Meanwhile, the paper also combined enhanced HOG with some common classifiers, and got satisfactory results, providing a very good value for references. What’s more, this paper put forward a new way of thinking for the study of image classification algorithm:
Now that the study of classifier tend to be more and more sophisticated, starting from feature models from to improve classification method is feasible and worth considering. Precisely because of this, this paper does not focus too much on improving the classification performance, so in future studies on the effect and speed of classifier, there are still some room for improvement. In addition a more reliable image enhancement techniques and background removal method will also improve feature extraction method performance, this can also be used as future research directions.

Acknowledgement

This study was funded by National Natural Science Foundation of China (61370102, 61170193, 61370185), Guangdong Natural Science Foundation (2014A030306050, S2012010009865, s2013010013432, S2013010015940), the Fundamental Research Funds for the Central Universities, SCUT(2015PT022), Science and Technology Planning Project of Huizhou City(2011P002, 2011g012, 2011P005,2011P003, 2011g011) and Science and Technology Planning Project of Guangdong Province (2011B090400041, 2012B010100039,2012B040305011, 2012B010100040), Education and Science Programs of Guangdong Province (11JXZ012, 14JXN065), Discipline Construction Programs of Guangdong Province (2013LYM00874), Key Technology Research and Development Programs of Huizhou(2013-13, 2013B020015008, 2014B050013016), Science and Technology Plan Project of Huizhou University(2012QN09).

References


(19) Zixi Xiang; Xueqiang Lv; Kai Zhang: An Image Classification Method Based on Multi-feature Fusion and Multi-kernel SVM. IEEE Conference on Computational Intelligence and Design (ISCID), 2014

(20) Li Weixing; Su Haijun; Pan Feng; Gao Qi; Quan Bin: A fast pedestrian detection via modified HOG feature, Control Conference (CCC) IEEE, 2015

(21) Diqing Sun; Junzo Watada: Detecting pedestrians and vehicles in traffic scene based on boosted HOG features and SVM, IEEE International Symposium on Intelligent Signal Processing (WISP), 2015

(22) Fiza Murtaza; Muhammad Haroon Yousaf; Sergio A. Velazini: Multi-view Human Action Recognition Using Histograms of Oriented Gradients (HOG) Description of Motion History Images (MHIs), International Conference on Frontiers of Information Technology (FIT), IEEE, 2015