

Chapter 40

K-Means Clustering Based on Density for Scene Image Classification

Ke Xie, Jin Wu, Wankou Yang and Changyin Sun

Abstract K-means clustering has been extremely popular in scene image classification. However, due to the random selection of initial cluster centers, the algorithm cannot always provide the most optimal results. In this paper, we develop a density-based k-means clustering. First, we calculate the density and distance for each feature vector. Then choose those features with high density and large distance as initial cluster centers. The remaining steps are the same with k-means. In order to evaluate our proposed algorithm, we have conducted several experiments on two-scene image datasets: Fifteen Scene Categories dataset and UIUC Sports Event dataset. The results show that our proposed method has good repeatability. Compared with the traditional k-means clustering, it can achieve higher classification accuracy when applied in multiclass scene image classification.

Keywords K-means · Density · Scene image classification

40.1 Introduction

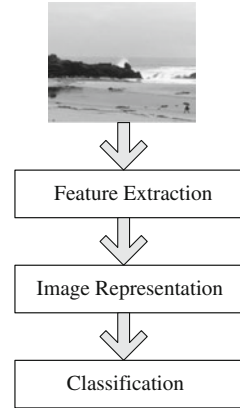
Scene image classification is widely applied in many fields, like content-based image and video retrieval, remote sensing image analysis, biometrics recognition, and intelligent video surveillance [1].

K. Xie · J. Wu · W. Yang · C. Sun (✉)
School of Automation, Southeast University, Nanjing 210096, China
e-mail: cysun@seu.edu.cn

K. Xie · J. Wu · W. Yang · C. Sun
Key Lab of Measurement and Control of Complex Systems of Engineering,
Ministry of Education, Southeast University, Nanjing 210096, China

K. Xie · J. Wu · W. Yang
Key Laboratory of Child Development and Learning Science of Ministry of Education,
Southeast University, Nanjing 210096, China

Fig. 40.1 The flow chart of scene image classification



The existing methods of scene image classification can be divided into three categories. (1) Methods based on the global low-level features of an image. Oliva and Torralba [2] presented a low-dimensional image representation, spatial envelope. (2) Methods based on image segmentation. Vogel [3] used image segmentation to get some meaningful target area. (3) Methods based on the visual vocabulary of image blocks. Fei-fei Li [4] proposed BOW (bag-of-words) model which has been very popular in scene image classification. This model treats image as a document consisting of many “visual words,” then computes a statistical histogram for classification. Figure 40.1. gives the flow chart of BOW model for scene image classification.

The first step is to extract features from local patches. Generally, an image will generate a large amount of features. How to organize these features to represent the image becomes very important. BOW used clustering algorithm to classify the features into groups (namely clusters).

K-means algorithm [5] is widely used in clustering problem. It assumes the closer the distance between two objects, the greater the similarity. However, the k-means clustering cannot always find the optimal configurations due to the randomness of the initial cluster centers and the uncertainty of the number of cluster centers. The density-based clustering method is used to select initial cluster centers according to the density distribution of data points [6–8]. Redmond [9] got the initial cluster centers by building the Kd-tree. Recently, Rodriguez and Laio [10] proposed a new clustering algorithm by fast research and to find density peaks. The clusters were defined as local maxima in the density of data points. There is no need to iterate for many times and the clustering assignment can be finished in only one step. However, the density will be computationally costly when there are a lot of data points.

Here, we propose an improved clustering algorithm which combines the k-means clustering and density-based clustering. It can achieve higher classification accuracy rate when applied in the scene image classification.

The remainder of the paper is organized as follows. Section. 40.2 presents our proposed clustering algorithm in detail. Section. 40.3 gives some experimental results and compares our method with previous algorithms on scene image classification. Section. 40.4 concludes this paper and provides some ideas for the future work.

40.2 Density-Based K-Means Clustering

K-means is a very simple way to solve the clustering problems. However, the algorithm is sensitive to the selection of the initial cluster centers. Moreover, the number of cluster center (namely k) is uncertain. When applied in real-world problem, we must try many times to find the optimal value. In this paper, we propose an improved k-means clustering and we call it density-based k-means clustering.

Compared with the traditional k-means algorithm, the initial cluster center is determined by density and distance. As mentioned in [10], for each point x_i , the density ρ_i is defined as:

$$\rho_i = \sum_j \chi(d_{ij} - d_c)$$

$$\chi(x) = \begin{cases} 1, & x < 0 \\ 0, & x \geq 0 \end{cases} \quad (40.1)$$

d_c is a cutoff distance. It is not a fixed value. ρ_i is equal to the number of points that are closer than d_c to point x_i . δ_i is the minimum distance between x_i and any other point with higher density :

$$\delta_i = \min_{j: \rho_j > \rho_i} (d_{ij}) \quad (40.2)$$

For the point with highest density, take $\delta_i = \max_j (d_{ij})$. Only those points with relative high ρ_i and high δ_i are considered as cluster centers. Therefore, we can define $\gamma_i = \rho_i \delta_i$ and choose those points with large γ_i as cluster centers.

Suppose we have n feature vectors x_1, x_2, \dots, x_n obtained from one training image, and k ($k < n$) is the number of clusters. The procedure of density-based k-means for the scene image classification can be described in the following steps:

- (1) Compute ρ_i , δ_i and γ_i for each feature vector x_i ($i = 1, \dots, n$), and sort γ in decreasing order.
- (2) Choose k feature vectors with large γ as initial cluster centers c_1, c_2, \dots, c_k .
- (3) Assign each remaining feature vector to the nearest cluster center.
- (4) Update the cluster centers by calculating the average of the vectors belonging to the same cluster.
- (5) Repeat step 3 and step 4 until k cluster centers no longer change.

40.3 Experiments

The scene image classification is implemented as follows. First, we use SIFT algorithm [11] to extract features for all the training and testing images. Then we use different clustering methods to organize the features and build the BOW model for each image. Finally, we use the LIBSVM package [12] to implement the SVM classification. Barla A et al. has found, compared with other standard kernel function, the histogram intersection can achieve very promising results on image classification [13]. Therefore, we construct a histogram intersection kernel as the kernel function of SVM. In our experiment, we choose one-against-all classification.

40.3.1 Dataset Used

In order to evaluate our proposed clustering algorithm, we have implemented three different experiments on two well-known scene image datasets. For each class in the dataset, we take 50 images for training and 20 images for testing.

40.3.1.1 Fifteen Scene Categories Dataset

The Fifteen Scene Categories dataset is built by Lazebnik [4]. The dataset contains 15 classes in JPG format, 8 of which are from MIT Scene dataset [2]. They are mostly grayscale images in 300 * 300 pixels. We randomly select eight classes of scene images as our experimental data. They are CALsuburb, coast, forest, highway, inside city, mountain, street, and tall building. Figure 40.2 gives some examples of eight scene classes.

40.3.1.2 UIUC Sports Event Dataset

The UIUC Sports Event dataset is collected by Li and Fei-Fei [14]. It contains eight sports event categories: Rowing (250 images), badminton (200 images), polo (182 images), bocce (137 images), snowboarding (190 images), croquet (236 images),



Fig. 40.2 Examples of eight scene classes



Fig. 40.3 Samples of eight sports categories

sailing (190 images), and rock climbing (194 images). Figure 40.3. gives some samples of the dataset.

The images in this dataset are color. Therefore before computing features, we convert each color image to gray image. And in order to reduce computational complexity, we scaled all images to within 300 * 300 pixels.

40.3.2 K-Means Clustering

Table 40.1 gives the experimental results of k-means clustering for scene image classification. Since the number of cluster center k is uncertain, we choose k from 50 to 450 to get the optimal parameter.

From Table 40.1, we can see that when the number of cluster centers increase, the iteration times increase at the same time. For the Fifteen Scene Categories dataset, although the accuracy has reached the highest at k = 450 the iteration times is very large. In order to achieve a balance between iteration times and accuracy, we take k = 300. The accuracy rate is 92.5 %. For the UIUC Sports Event dataset, the optimal k is also 300. The corresponding accuracy rate is 66.25 %.

Table 40.1 Results of k-means clustering

k (The number of centers)	Fifteen Scene Categories dataset		UIUC Sports Event dataset	
	Iteration times	Accuracy (%)	Iteration times	Accuracy (%)
50	18	89.375	8	57.5
100	13	90	13	58.75
150	16	89.375	10	60.625
200	19	91.25	13	62.5
250	18	91.25	15	65
300	15	92.5	17	66.25
350	18	92.5	18	65.625
400	20	92.5	19	65
450	25	93.75	24	65.625

40.3.3 Density-Based Clustering

We carried out the density-based clustering proposed by Rodriguez and Laio. Tables 40.2 and 40.3 give the results according to different dc.

From Table 40.2, it is easy to find the optimal dc = 0.5 for the Fifteen Scene Categories dataset. The corresponding accuracy is 93.125 % which is higher than that in k-means clustering. For the UIUC Sports Event dataset, the optimal dc = 0.9. The corresponding accuracy is 67.5 % which is also higher than that in k-means clustering.

40.3.4 Density-Based K-Means Clustering

This time, we tried our proposed clustering algorithm for scene image classification. According to Table 40.1, we take k = 300 (the number of centers). The parameter dc is varied. Tables 40.4 and 40.5 give the final results.

Apparently, compared with k-means and density-based clustering, the density-based k-means we proposed in this paper achieves higher classification accuracy. For the Fifteen Scene Categories dataset, the most optimal dc = 0.6. The accuracy

Table 40.2 Results of density-based clustering on 15 scene categories dataset

dc	Accuracy (%)
0.1	86.875
0.2	88.75
0.3	90.625
0.4	90.625
0.5	93.125
0.6	90.625
0.7	89.375
0.8	91.875

Table 40.3 Results of density-based clustering on UIUC sports event dataset

dc	Accuracy (%)
0.5	58.125
0.6	61.25
0.7	62.5
0.8	63.75
0.9	67.5
1.0	63.75
1.1	58.75
1.2	58.125

Table 40.4 Results of density-based k-means clustering on 15 scene categories dataset

dc	Iteration times	Accuracy (%)
0.1	17	92.5
0.2	20	91.875
0.3	18	91.875
0.4	20	92.5
0.5	18	92.5
0.6	14	94.375
0.7	13	90.625
0.8	14	90.625

Table 40.5 Results of density-based k-means clustering on UIUC sports event dataset

dc	Iteration times	Accuracy (%)
0.4	15	64.375
0.5	21	69.375
0.6	22	68.75
0.7	19	68.125
0.8	21	69.375
0.9	29	68.75
1.0	22	73.75
1.1	21	69.375
1.2	18	68.125

has reached 94.375 %. As can be seen from Table 40.5, for the UIUC Sports Event Dataset, the most optimal $dc = 1.0$, the corresponding accuracy has reached 73.75 %.

40.3.5 Discussion

The above experiments showed that our proposed density-based k-means clustering can always achieve better performance. But for different database, the optimal parameters are not the same. In the Fifteen Scene Categories dataset, the optimal cutoff distance dc is 0.6. In the UIUC Sports Event dataset, the best $dc = 1.0$.

40.4 Conclusion

In this paper, we proposed a density-based k-means clustering for scene image classification. Compared with the traditional k-means algorithm, the initial cluster centers are not randomly selected. For each feature vector, we calculate two quantities: The density and the distance. Then we can find those points with high

density and high distance, which we call cluster centers. After we have determined the cluster centers, the remaining steps are the same with k-means clustering.

We have done several experiments to compare different clustering algorithm for scene image classification on two well-known scene image datasets. Experimental results show that our proposed method achieves better performance. When applied in multiclass scene image classification, it has good repeatability and high classification accuracy rate.

However, the selection of the cutoff distance d_c is still a problem. How to find the most optimal parameter for the clustering problem of scene images is our next work.

Acknowledgments This project is partly supported by NSF of China (61375001), partly supported by the open project program of Key Laboratory of Child Development and Learning Science of Ministry of Education, Southeast University (No.CDLS-2014-04), partly supported by China Postdoctoral Science Foundation (2013M540404) and partly supported by the Ph.D. Programs Foundation of Ministry of Education of China (No. 20120092110024).

References

1. Zhou L (2012) Research on key technologies of scene classification and object recognition. Graduate School of National University of Defense Technology (in Chinese)
2. Oliva A, Torralba A (2001) Modeling the shape of the scene: a holistic representation of the spatial envelope. *Int J Comput Vision* 42(3):145–175
3. Vogel J, Schiele B (2004) Natural scene retrieval based on a semantic modeling step. In: *ACM international conference image video retrieval*, New York, pp 207–215
4. Lazebnik S, Schmid C, Ponce J (2006) Beyond bags of features: Spatial pyramid classification for recognizing natural scene categories. In: *IEEE computer society conference on computer vision and pattern recognition*. New York, pp 2169–2178
5. Hartigan J, Wang M (1979) A k-means clustering algorithm. *Appl Stat* 28:100–108
6. Ester M, Kriegel HP, Sander J et al (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. In: *Kdd vol 96*, pp 226–231
7. Ankerst M, Breunig MM, Kriegel HP et al (1999) Optics: ordering points to identify the clustering structure *ACM sigmod record*. *ACM* 28(2):49–60
8. Hinneburg A, Keim DA (1998) An efficient approach to clustering in large multimedia databases with noise. In: *KDD vol.98*, pp 58–65
9. Redmond SJ, Heneghan C (2007) A method for initialising the k-means clustering algorithm using Kd-trees. *Pattern Recogn Lett* 28(8):965–973
10. Rodriguez A, Laio A (2014) Clustering by fast search and find of density peaks. *Science* 344(6191):1492–1496
11. Lowe DG (2004) Distinctive image features from scale-invariant keypoints. *Inter J Comput Vis* 2(60):91–110
12. Chang CC, LIN CJ (2001) LIBSVM: a library for support vector machines. <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
13. Barla A, Odone F, Verri A (2003) Histogram intersection kernel for image classification. In: *International Conference on Image Processing vol.3(2):p III-513–16*
14. Li LJ, Fei-Fei L (2007) What, where and who? Classifying events by scene and object recognition. In: *IEEE 11th international conference on IEEE. computer vision ICCV*, pp 1–8