

Research Article

A Hybrid Clustering Approach for Bag-of-Words Image Categorization

Hui Huang  and Yan Ma 

College of Information and Electrical Engineering, Shanghai Normal University, Shanghai 200234, China

Correspondence should be addressed to Yan Ma; ma-yan@shnu.edu.cn

Received 29 December 2018; Revised 17 January 2019; Accepted 28 January 2019; Published 11 February 2019

Academic Editor: Piotr Jędrzejowicz

Copyright © 2019 Hui Huang and Yan Ma. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The Bag-of-Words (BoW) model is a well-known image categorization technique. However, in conventional BoW, neither the vocabulary size nor the visual words can be determined automatically. To overcome these problems, a hybrid clustering approach that combines improved hierarchical clustering with a K-means algorithm is proposed. We present a cluster validity index for the hierarchical clustering algorithm to adaptively determine when the algorithm should terminate and the optimal number of clusters. Furthermore, we improve the max-min distance method to optimize the initial cluster centers. The optimal number of clusters and initial cluster centers are fed into K-means, and finally the vocabulary size and visual words are obtained. The proposed approach is extensively evaluated on two visual datasets. The experimental results show that the proposed method outperforms the conventional BoW model in terms of categorization and demonstrate the feasibility and effectiveness of our approach.

1. Introduction

Bag-of-Words (BoW), which was originally implemented in the field of text categorization, has been widely used for image categorization. The BoW model for image categorization includes three essential steps: (1) extracting local features from images; (2) constructing a visual vocabulary by clustering the local features to visual words; (3) mapping each local feature to a visual word and representing an image as a vector containing the count of each visual word in that image. Therefore, in BoW, the image is described as a vector, which reduces the time complexity of image categorization [1]. However, a visual vocabulary designed in this manner is not necessarily effective for categorization, and the vocabulary size will influence the level of effectiveness. Therefore, obtaining a suitable vocabulary with an appropriate size and visual words is quite important. In fact, the vocabulary size and visual words correspond to the number of clusters and cluster centers, respectively, so the essence of the aforementioned problem is to determine the optimal number of clusters and cluster centers.

A fundamental problem of cluster analysis is determining the best estimate for the number of clusters, as this has a

significant effect on the clustering results. The most widely used clustering algorithms are K-means and hierarchical clustering. K-means is efficient and simple, but the initial centers are randomly chosen and the number of clusters K is not known beforehand. Hierarchical clustering achieves its final results by iteratively splitting or merging clusters [2]. This method is simple to implement and can solve multi-scale spatial clustering problems, although its termination condition is difficult to determine. Many improved methods that estimate the number of clusters K in BoW have been developed. A novel algorithm combining Gaussian Mixture Models and the Bayesian Information Criterion claims to obtain the true number of visual words in the vocabulary [3], and the number of visual words has been computed by a data mining method based on a cooccurrence matrix [4]. By removing inconsistent edges, the minimum spanning tree based clustering method can identify clusters of arbitrary shape [5]. A clustering quality curve has been introduced to estimate the number of clusters in hierarchical clustering [6], and the original cluster centers have been selected dynamically by constructing a minimum spanning tree [7]. The true number of clusters can also be determined using a clustering validity index [8].

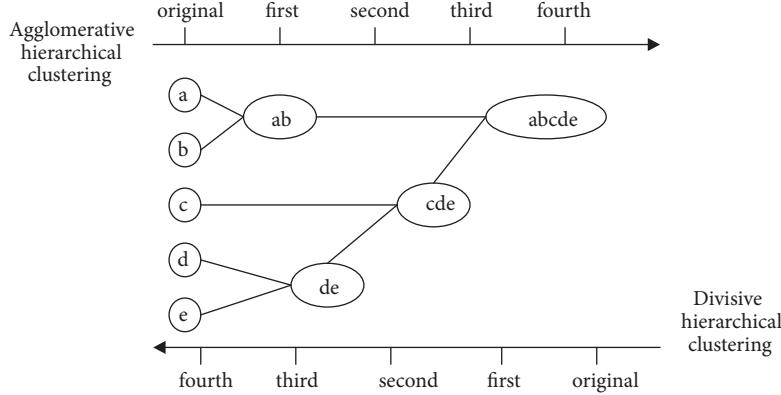


FIGURE 1: Schematic of agglomerative and divisive hierarchical clustering algorithms.

In this paper, we present a novel clustering method that combines hierarchical clustering and K-means. The proposed clustering method is adopted in BoW to complete the process of image categorization. First, in terms of hierarchical clustering, a new cluster validity index is introduced to determine the vocabulary size, i.e., the optimal number of clusters for a set of image features. After the optimal number of clusters has been determined, the optimal initial cluster centers are calculated by an improved max-min method. The best number of clusters and initial centers are then fed into K-means, and the final result for the cluster centers (that is, the vocabulary) is achieved. Finally, the images are categorized according to histogram vectors that represent images with BoW. In short, the vocabulary size for the BoW is adaptively adjusted to reduce errors introduced by image quantization, leading to better image categorization performance.

The rest of this paper is organized as follows: Section 2 reviews hierarchical clustering algorithms and presents the proposed algorithm, which includes a cluster validity index and an improved max-min distance method. Our experimental results and analysis then evaluate the performance of the proposed method in Section 3. Finally, Section 4 draws together some conclusions from this study.

2. Cluster Validity Index and Initial Cluster Centers

2.1. Hierarchical Clustering. Hierarchical clustering creates a hierarchy of clusters that can be represented in a tree structure. There are two types of hierarchical clustering: agglomerative and divisive. Agglomerative hierarchical clustering algorithms work by merging the nearest clusters in a bottom-up fashion. Each individual data point is first assigned to a cluster, and then the two clusters that are closest to each other are merged into a new cluster. This process continues until all clusters have merged into one. The divisive hierarchical clustering algorithm works by splitting clusters into separate clusters in a top-down fashion. A cluster containing all data is created, and this is then divided into two clusters with respect to the amount of separation between data. This process is repeated until the final clusters contain only one data point

[9]. The two types of hierarchical clustering are illustrated in Figure 1.

The agglomerative method has lower computational complexity than the divisive approach, and can obtain better classification performance. In summary, for the computational complexity considered here, we adopt the agglomerative method of hierarchical clustering.

The typical agglomerative hierarchical clustering approach consists of the following steps:

- (1) For a dataset of size N , each sample is regarded as a separate cluster.
- (2) Calculate distances between all pairs of clusters and create a distance matrix.
- (3) Find the minimum distance in the distance matrix, then merge the two corresponding clusters.
- (4) Repeat Steps (2) and (3) until all elements have been merged into one cluster containing N samples.

Hierarchical clustering algorithms are simple to implement and can solve multiscale spatial clustering problems. However, they have some limitations; in particular, the terminal condition is difficult to determine.

2.2. Cluster Validity Index. The quality of clustering is measured by the cluster validity index. Furthermore, the optimal number of clusters k is estimated based on the cluster validity index. It has been verified that there is a knee point in the validity index curves representing the correct number of clusters. Various cluster validity indices have been developed, such as Calinski–Harabasz (CH) [10], Davies–Bouldin (DB) [11], and Krzanowski–Lai (KL) [12]. These indices are implemented to evaluate the clustering results according to the dataset itself and the statistical properties of the clustering results. However, the CH method may present unstable results, with the number of clusters varying with the search range. The DB method is only appropriate for well-separated clusters, and the KL method is designed for datasets whose structure can be easily estimated. Thus, to obtain better clustering results, we propose a novel cluster validity index for an agglomerative hierarchical clustering algorithm. The index is described in more detail below.

Let the dataset be X . Record the distances between all pairs of data samples in a matrix $D_{n \times n}$. As the data samples are clustered by agglomerative hierarchical clustering, a hierarchical clustering tree H containing n levels is formed, that is, $H = \{H_1, H_2, \dots, H_n\}$. Any level H_k of H consists of c clusters with n_i samples. Compactness and separation are implemented to measure the similarity between within-cluster and between-cluster samples based on the Euclidean distance [13].

Definition 1. The within-cluster compactness is the longest edge of the minimum spanning tree formed by all samples in the cluster, that is, the maximum weight. In detail, any level H_k of the hierarchical clustering tree H formed by X consists of c clusters $\{G_1, G_2, \dots, G_c\}$, and each cluster has n_i samples ($i = 1, 2, \dots, c$). The within-cluster compactness $cd(i)$ for the i th cluster is then defined as:

$$cd(i) = \max(w(x_m, x_n)), \quad x_m \in G_i, \quad x_n \in G_i \quad (1)$$

where $w(x_m, x_n)$ denotes the edge weight in the minimum spanning tree formed by all samples in the i th cluster.

Definition 2. The between-cluster separation is the minimum distance from a given data sample to the closest sample from another cluster. In detail, any level H_k of the hierarchical clustering tree H formed by X consists of c clusters $\{G_1, G_2, \dots, G_c\}$, and each cluster has n_i samples ($i = 1, 2, \dots, c$). The between-cluster separation $sd(i)$ for the i th cluster is then defined as:

$$sd(i) = \min(\min(\text{dist}(x_i, x_j))), \quad 1 \leq j \leq c, \quad x_i \in G_i, \quad x_j \in G_j \quad (2)$$

where $\text{dist}(x_i, x_j)$ represents the Euclidean distance between sample x_i in cluster i and sample x_j in cluster j .

Definition 3. Suppose that any level H_k of the hierarchical clustering tree H formed by X consists of c clusters $\{G_1, G_2, \dots, G_c\}$, and each cluster has n_i samples ($i = 1, 2, \dots, c$). The cluster validity index $CS(i)$ for the i th cluster is the ratio of the within-cluster compactness to the between-cluster separation, that is:

$$CS(i) = \frac{cd(i)}{sd(i)} \quad (3)$$

The index CS reflects the cluster validity for the clustered samples in a dataset. Lower CS values denote better clustering results. For a dataset, we can analyze the clustering result by averaging the CS index for all clusters. A lower average indicates better clustering results for the dataset. We write $\text{avg } CS(k)$ for the average of the CS indices for k clustered classes and denote the optimal number of clusters as k_{opt} :

$$\text{avg } CS(k) = \sum_{i=1}^k \frac{CS(i)}{k} \quad (4)$$

$$k_{opt} = \arg \min(\text{avg } CS(k)), \quad 2 \leq k < n \quad (5)$$

The criterion for judging the cluster validity considers both within-cluster compactness and between-cluster separation. In the case of within-cluster compactness, a smaller distance between a pair of samples in the i th cluster denotes better clustering results. The maximum or minimum distance between a pair of samples is not representative, and samples in a cluster can form the minimum spanning tree. Therefore, it is reasonable that the within-cluster compactness can be measured by the weight of the minimum spanning tree. In the case of between-cluster separation, a greater distance between cluster i and the nearest neighbor cluster j suggests better clustering results. Therefore, it is important to exploit the minimum distance between a pair of clusters to measure the between-cluster separation. The CS index synthesizes these two factors: a smaller value of the CS index implies that the clustering result is better, i.e., a smaller within-cluster compactness and greater between-cluster separation. The best partition (that is, the optimal number of clusters K) is obtained when CS_k reaches a minimum.

In determining the optimal number of clusters using agglomerative hierarchical clustering, we search the effective range $[k_{min}, k_{max}]$. We set k_{min} to 2, and choose \sqrt{n} be the upper bound, as related research shows that $k_{max} \leq \sqrt{n}$ where n is the number of samples in the dataset [14]. The optimal partition result can then be obtained using the proposed CS index. The proposed method is as follows:

- (1) *input*: $k, k \in [k_{min}, k_{max}]$
- (2) *while* $k \leftarrow k_{min}$ to k_{max}
 - (a) Calculate clustering validity index according to Eq. (3);
 - (b) Calculate the average CS according to Eq. (4);
 - (c) Calculate the proper number of clusters according to Eq. (5);
- (3) *end*
- (4) *output*: k, CS

The effectiveness of the CS index is demonstrated through the following experiments. Three datasets were selected from the UCI database [14]. We applied the DB, CH, KL, and CS indexes to the three datasets, and obtained the optimal number of clusters using each index. Table 1 illustrates the results for the above validity indexes with the datasets divided into the same number of clusters. The optimal number of clusters is indicated in bold.

A detailed description of these datasets can be found in the UCI database. The size, number of attributes (dimension of the data), real number of clusters in the datasets, and proper number of clusters obtained with the four aforementioned methods are listed in Table 2.

Tables 1 and 2 indicate that errors occur, and there exist some differences between the proper number of clusters and the true number of clusters when the DB, CH, and KL clustering validity indices are applied to the datasets. In contrast, the proposed index gives a number of clusters that is consistent with the true number. This indicates the effectiveness and feasibility of the proposed index for evaluating clustering results.

TABLE 1: Values of four cluster validity indexes for the Iris/Wine/Breast-Cancer-Wisconsin datasets.

K	DB	CH	KL	CS
2	0.0666/0.4400 /115.4636	0.2149/0.4099/ 28.1092	11.4453/3.1191/1.0830	6.3420/2.4650/ 4.6660
3	0.4640/0.4962/63.9422	1.0816 /0.7709/14.0356	60.6343/7.0889/3.5436	1.9846/0.6920 /28.3850
4	1.1963/0.6330/ 40.3679	0.9327/0.7344/27.4826	1.4063/0.6718/0.4943	7.9528/2.4210/28.4230
5	1.5125/0.6583/61.6671	0.7129/1.1670/20.5832	0.9956/5.9125/1.1797	8.8753/2.9340/28.8990
6	2.6920/1.5296/62.6160	0.5731/1.1521/16.4438	1.9325/0.6848/1.1472	10.9862/2.9940/28.9480
7	7.6907/1.4995/60.0437	0.4929/ 1.3889 /13.6843	0.1244/2.1073/1.1246	12.7995/3.3610/28.9420
8	20.9400/2.9274/50.3963	0.6470/1.0086/11.7129	1.7109/4.9006/1.1081	13.7742/3.6750/29.0550

TABLE 2: Proper number of clusters given by four cluster validity indexes.

Dataset	Number of attributes	True Clusters	DB	CH	KL	CS
Iris	4	3	2	3	3	3
Wine	13	3	2	7	3	3
Breast-Cancer-Wisconsin	11	2	4	2	3	2

2.3. *Initial Cluster Centers.* The basic principle of the K-means algorithm is to use k initial cluster centers and classify the samples to the closest cluster center. The cluster centers are then updated iteratively until a convergence condition is satisfied [15]. A detailed description of K-means is as follows:

- (1) Randomly select k samples from the dataset as initial cluster centers.
- (2) Calculate the distance between each sample and the k cluster centers, and classify the sample into the closest class.
- (3) Recalculate and update the cluster centers according to the above results.
- (4) Repeat Steps (2) and (3) until the clustering does not change the results.

K-means is popular because of its simplicity and efficiency. Despite its advantages, K-means suffers from some limitations, particularly that the number of clusters and the initial cluster centers must be determined a priori. Poor quality initialization can lead to a poor solution. Therefore, it is important to determine an appropriate number of clusters and initial cluster centers. The method proposed in this paper can be exploited to solve the problem of the number of clusters. In the case of initial cluster centers, different centers will lead to different clustering results. The max-min distance method is introduced to optimize the initial cluster centers [16].

The core idea of the max-min distance method is to select cluster centers according to the maximum distance and to classify samples according to the minimum distance, thus avoiding the initial cluster centers from being too close together. For the max-min distance method, the number of clusters is unknown in advance, and the proportional coefficient θ is selected as the limitation condition. To determine the $(k+1)th$ center, we calculate the distance d_{ij} between sample x_i (which has not been selected as a center) and the k predetermined cluster centers. We then find $d_r = \max(\min(d_{i1}, d_{i2}, \dots, d_{ik}))$ and take sample x_r corresponding

to the condition $d_r > \theta \|c_1 - c_2\|$ as the $(k+1)th$ center c_{k+1} . However, it is difficult to determine the value of θ .

In this paper, the max-min distance method is extended to solve the aforementioned problem and optimize the initial cluster centers under the condition that the number of clusters is known a priori. The proposed method to determine initial cluster centers can therefore be summarized as follows:

- (1) Calculate the average of all samples and select the sample closest to the average as the first cluster center c_1 .
- (2) When the number of clusters is 2, select the unclassified sample farthest from c_1 as the second cluster center c_2 .
- (3) When the number of clusters is 3, calculate the distances between samples that have not been selected as centers and the two initial cluster centers c_1, c_2 . Find the minimum distance d_t , and choose the t th sample as the third center c_3 according to $d_t = \max(d_i)$.
- (4) When the number of clusters is k and $k \leq k_{max}$, calculate the distances d_{ij} between samples that have not been selected as centers and the $k-1$ determined cluster centers, and find the sample satisfying the condition $d_r = \max(\min(d_{i1}, d_{i2}, \dots, d_{i(k-1)}))$. Then, choose the r th sample as the k th center, and output the k initial cluster centers.

In this section, the proposed algorithm is evaluated to illustrate its performance in automatically determining the number of clusters and the cluster centers. The experiments were implemented on the Iris dataset by two methods: the conventional and improved max-min distance algorithm. For the former, we set θ to 0.7, 0.4, 0.1 and obtained corresponding values for K of 2, 3, and 5, respectively. For the improved max-min method, we calculated K = 3 with the cluster validity index CS and obtained appropriate initial cluster centers. Figure 2 illustrates the two corresponding clustering results and the ROC (Receiver Operating Characteristic) curve.

The clustering results for different values of θ are shown in Figures 2(a)-2(c). Figure 2(d) illustrates the clustering

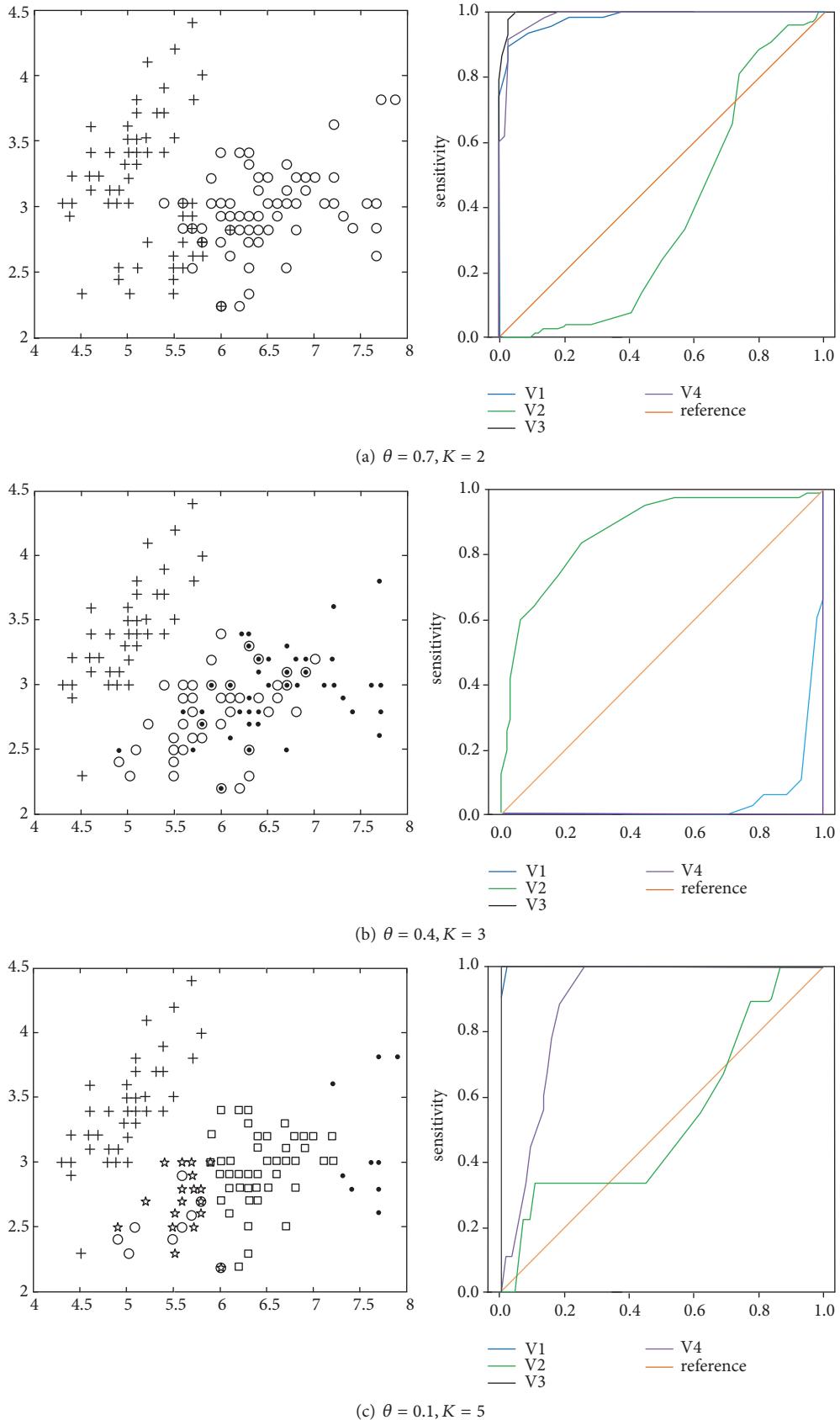


FIGURE 2: Continued.

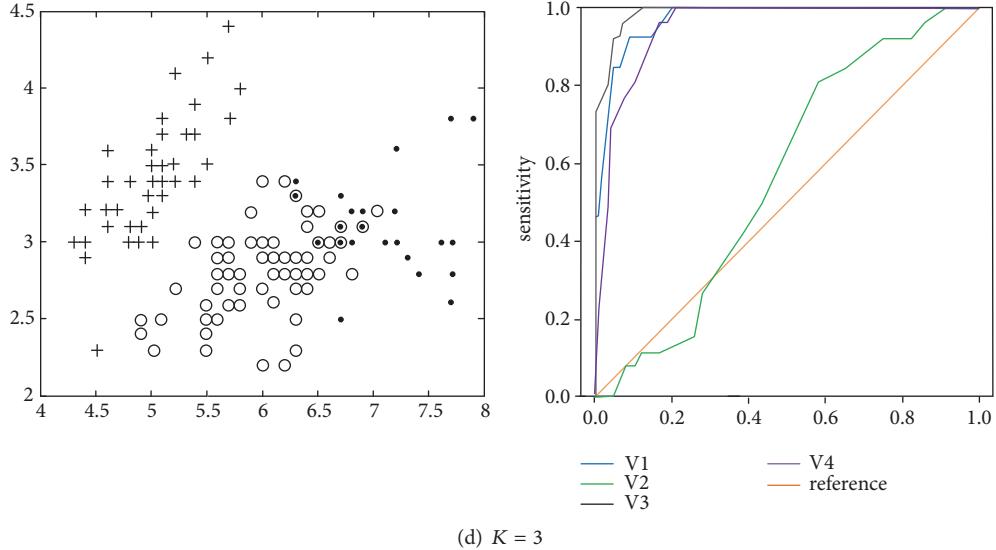


FIGURE 2: Clustering results and ROC curves given by the conventional and improved max-min method.

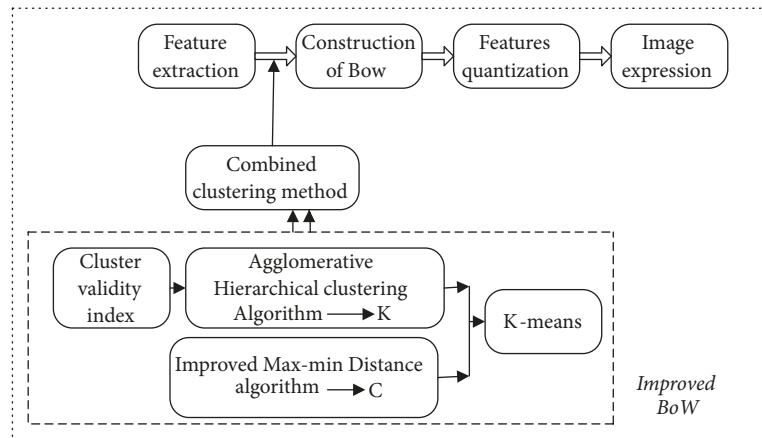


FIGURE 3: Overview of the proposed approach.

results given using the improved max-min distance method. As shown in Figure 2(a), the data samples are obviously partitioned into two clusters, which does not conform to the actual situation. The number of clusters is correct in Figure 2(b), but the clustering results are not ideal. Neither the number of clusters nor the clustering results are ideal in Figure 2(c). Compared with these results, Figure 2(d) shows that the improved max-min distance method can enhance the clustering results with the correct number of clusters and a more concentrated data distribution. The right-hand side of Figure 2 illustrates the ROC curves, which reflect the performance of the classification method. A greater area under the ROC curve (AUC) signifies that the performance of the classification method is better. The AUC in Figures 2(a)–2(c) is lower than in Figure 2(d), which indicates that the improved max-min approach generates a better clustering result. This means that the improved max-min distance method gives a better classification of the dataset in the clustering process.

2.4. Proposed Method. The above sections have described the cluster validity index for the optimal number of clusters and the improved max-min distance method for initial cluster centers to achieve the optimal clustering result. Based on the aforementioned methods, it is possible to improve the conventional BoW: we combine the agglomerative hierarchical clustering algorithm with K-means to automatically determine the optimal number of clusters and the initial cluster centers, i.e., the vocabulary size and the visual words. Figure 3 gives an overview of the proposed approach.

We can summarize the proposed clustering algorithm as follows:

- (1) The dataset is clustered by agglomerative hierarchical clustering, and a hierarchical clustering tree is generated.
- (2) Define a new cluster validity index to evaluate the clustering results in each level, and determine the optimal number of clusters K .

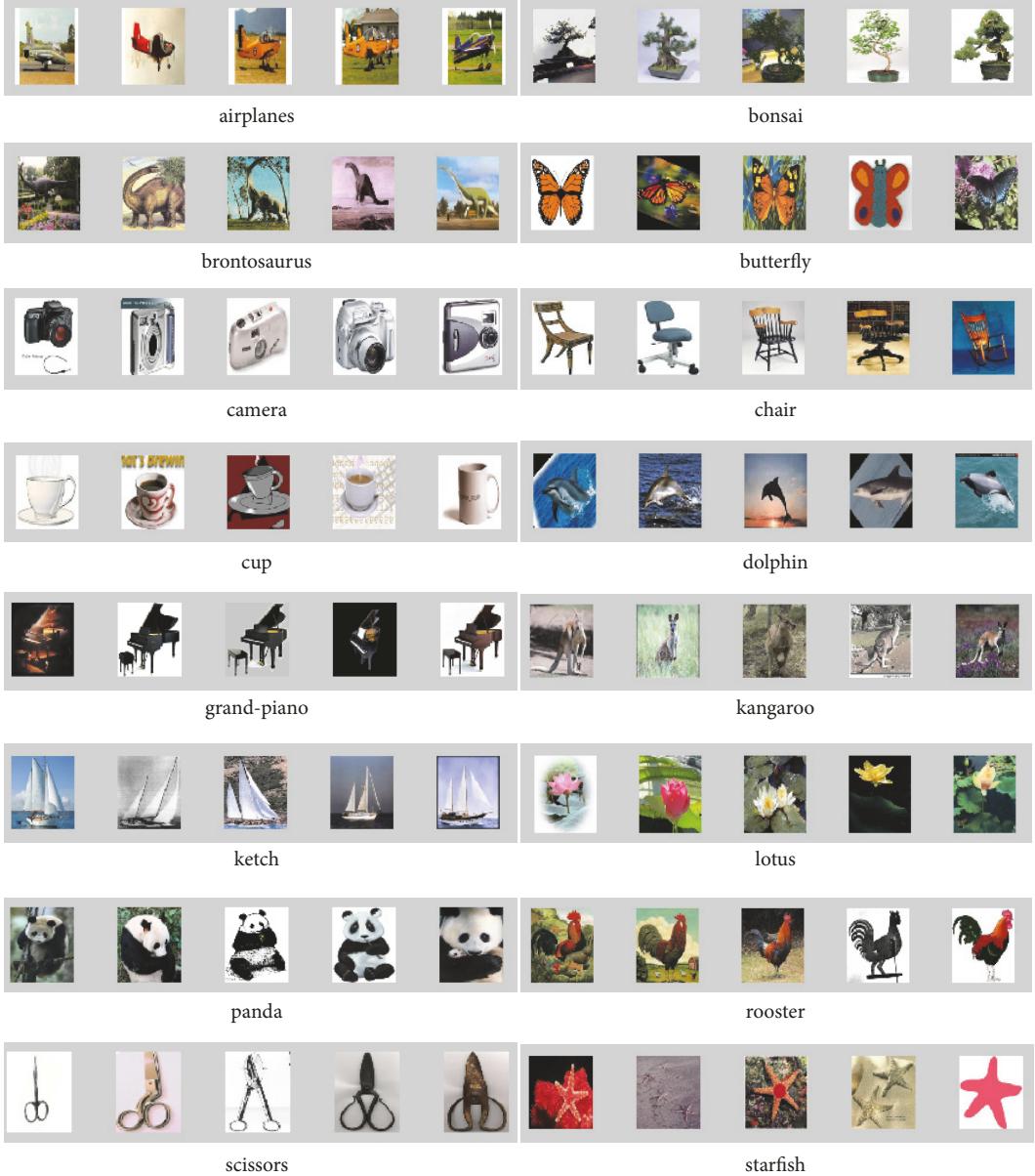


FIGURE 4: Some samples in sixteen categories from the Caltech 101 dataset.

- (3) Apply the improved max-min distance algorithm to determine the initial cluster centers C according to K .
- (4) Input K and C into the K-means algorithm, and cluster the dataset again to improve the results from the agglomerative hierarchical clustering algorithm. Finally, we achieve the optimal clustering results, which correspond to the partitioned dataset.

3. Experimental Results and Analysis

To evaluate the proposed method, we used the Caltech 101[17] and 15 Scenes [18] datasets. Caltech 101 includes 101 object categories, whereas 15 Scenes includes the 15 scene categories of Store, Office, Tall building, Street, Open country,

Mountain, Inside city, Highway, Forest, Coast, Living room, Kitchen, Industrial, Suburb, and Bedroom.

We used sixteen object categories from Caltech 101 as shown in Figure 4 and six object categories from 15 Scenes (bedroom, CALsuburb, industrial, MITcoast, MITforest, and MITinsidecity) for the experiments. We chose 20 images per category for training and 10 per category for testing. For convenience, the size of all images was set to 100×100 pixels. Each training image was represented as a BoW. First, scale-invariant feature transform (SIFT) descriptors extracted from the training images are chosen as visual features. Second, the dataset constructed by the visual features is clustered by the proposed clustering approach. The size of the dataset is dependent on the total number of the SIFT descriptors. In our experiments, the dataset includes 34328 feature vectors.

class	Categorization results									
airplanes										
bonsai										
brontosaurus										
butterfly										
camera										
chair										
cup										
dolphin										
grand-piano										
kangaroo										
ketch										
lotus										
panda										
rooster										
scissors										
starfish										

FIGURE 5: Image categorization results on Caltech 101.

Finally, the BoW representation was fed into a support vector machine (SVM) [19] using LIBSVM [20] to improve the classification performance. In this paper, we analyze the BoW model with both known and unknown vocabulary sizes.

(1) *Known Vocabulary Size.* We fixed the vocabulary size to 150. While clustering the image data, the value of K was set to 150 and the SVM was applied to classify the image. Figures 5 and 6 show the categorization results (the first eight images are selected), where red blocks denote the misclassified images. Tables 3 and 4 list the precision and recall [21] corresponding to the two databases, respectively.

Figure 5 shows that a large number of the Caltech 101 image categories were misclassified. There exist two misclassified images for the bonsai, brontosaurus, and camera classes, and three misclassified images for the butterfly class. There was generally one misclassified images in each of the 15 Scenes classes. That is, different levels of misclassification

exist. From Tables 3 and 4, it can be seen that the precision and recall corresponding to both databases was not high.

(2) *Unknown Vocabulary Size.* In the case of an unknown BoW vocabulary size, we used the proposed method to automatically determine the optimal number of clusters and the initial cluster centers. We then performed a clustering analysis to categorize the images and expressed the images according to the distribution of visual words. By calculating the number of clusters, it was found that the vocabulary sizes for Caltech 101 and 15 Scenes were 475 and 264. The SVM classifier was then used for the classification task. Figures 7 and 8 show the categorization results (the first eight images are selected), where red blocks denotes the misclassified images. Tables 5 and 6 list the precision and recall corresponding to the two databases, respectively.

Figures 7 and 8 show that the image categorization results for both databases have been improved, and the number

class	Categorization results							
bedroom								
CALsuburb								
industrial								
MITcoast								
MITforest								
MITinsidecity								

FIGURE 6: Image categorization results on 15 Scenes.

TABLE 3: Precision and recall on Caltech 101.

	airplanes	bonsai	brontosaurus	butterfly	camera	chair	cup	dolphin
Precision (%)	100.00	87.50	100.00	87.50	87.50	100.00	100.00	87.50
Recall (%)	80.00	60.00	60.00	50.00	60.00	70.00	60.00	50.00
	grand-piano	kangaroo	ketch	lotus	panda	rooster	scissors	starfish
Precision (%)	87.50	87.50	100.00	75.00	100.00	100.00	100.00	100.00
Recall (%)	60.00	65.00	80.00	50.00	80.00	85.00	85.00	80.00

TABLE 4: Precision and recall on 15 Scenes.

	bedroom	CALsuburb	industrial	MITcoast	MITforest	MITinsidecity
Precision (%)	87.50	87.50	75.00	100.00	87.50	100.00
Recall (%)	70.00	70.00	60.00	100.00	70.00	100.00

TABLE 5: Precision and recall on Caltech 101.

	airplanes	bonsai	brontosaurus	butterfly	camera	chair	cup	dolphin
Precision (%)	100.00	100.00	75.00	87.50	87.50	100.00	100.00	87.50
Recall (%)	80.00	80.00	60.00	70.00	70.00	80.00	80.00	60.00
	grand-piano	kangaroo	ketch	lotus	panda	rooster	scissors	starfish
Precision (%)	87.50	95.00	100.00	85.00	100.00	90.00	100.00	85.00
Recall (%)	60.00	80.00	70.00	75.00	80.00	70.00	80.00	70.00

of misclassified images has decreased to at most one. From Tables 5 and 6, it can be seen that the precision and recall have improved compared to the case of a preset vocabulary size, which illustrates that the improved BoW model can express images more accurately and automatically determine the vocabulary size. Hence, the proposed method obtains improved categorization performance.

4. Conclusions

In this paper, we have presented an improved method that automatically determines the vocabulary size of the BoW model. We performed a clustering analysis on image datasets to judge the optimal number of clusters and initial cluster centers according to a cluster validity index, and improved max-min distance method by combining K-means

with agglomerative hierarchical clustering. The experimental results on three datasets validated the effectiveness and feasibility of the proposed approach. To evaluate the categorization performance of the proposed approach, experiments with known and unknown vocabulary sizes were conducted on the Caltech 101 and 15 Scenes datasets. The results demonstrated that the proposed BoW model based on the hybrid of hierarchical clustering and K-means algorithms classifies images more accurately and achieves better categorization performance.

Data Availability

The data used to support the findings of this study are included within the article.

class	Categorization results							
airplanes								
bonsai								
brontosaurus								
butterfly								
camera								
chair								
cup								
dolphin								
grand-piano								
kangaroo								
ketch								
lotus								
panda								
rooster								
scissors								
starfish								

FIGURE 7: Image categorization results on Caltech 101.

TABLE 6: Precision and recall on 15 Scenes.

	bedroom	CALsuburb	industrial	MITcoast	MITforest	MITinsidecity
Precision (%)	100.00	87.50	87.50	100.00	100.00	100.00
Recall (%)	80.00	70.00	70.00	80.00	80.00	80.00

class	Categorization results							
bedroom								
CALsuburb								
industrial								
MITcoast								
MITforest								
MITinsidecity								

FIGURE 8: Image categorization results on 15 Scenes.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (nos. 61501297 and 61373004).

References

- [1] Y.-X. Wang, H. Guo, C.-Q. He, Z. Feng, and Q. Jia, “Bag of spatial visual words model for scene classification,” *Computer Science*, vol. 38, no. 8, p. 265, 2011.
- [2] R. C. de Amorim, “Feature relevance in Ward’s hierarchical clustering using the L_p norm,” *Journal of Classification*, vol. 32, no. 1, pp. 46–62, 2015.
- [3] M. Xu, J. Han, and L. Guo, “Determine word number of Visual Bag-of-Words model by model selection method–fbfb,” *Computer Engineering and Applications*, vol. 31, no. 47, pp. 148–150, 2011.
- [4] K. Koonsanit, C. Jaruskulchai, and A. Eiumnoh, “Determination of the initialization number of clusters in K-means clustering application using co-occurrence statistics techniques for multispectral satellite imagery,” *International Journal of Information and Electronics Engineering*, vol. 2, no. 5, pp. 785–789, 2012.
- [5] X. Lv, Y. Ma, X. He, H. Huang, and J. Yang, “CcIMST: A clustering algorithm based on minimum spanning tree and cluster centers,” *Mathematical Problems in Engineering*, vol. 2018, Article ID 8451796, 14 pages, 2018.
- [6] N. Chen, Z.-S. Xu, and M.-S. Xia, “Hierarchical hesitant fuzzy K-means clustering algorithm,” *Applied Mathematics-A Journal of Chinese Universities Series B*, vol. 29, no. 1, pp. 1–17, 2014.
- [7] X.-J. Tong, F.-R. Meng, and Z.-X. Wang, “Optimization to k-means initial cluster centers,” *Computer Engineering and Design*, vol. 32, no. 8, pp. 2721–2723, 2011.
- [8] J. Yang, Y. Ma, X. Zhang, S. Li, and Y. Zhang, “An initialization method based on hybrid distance for k-means algorithm,” *Neural Computation*, vol. 29, no. 11, pp. 3094–3117, 2017.
- [9] J. Wu, H. Xiong, and J. Chen, “Towards understanding hierarchical clustering: A data distribution perspective,” *Neurocomputing*, vol. 72, no. 10–12, pp. 2319–2330, 2009.
- [10] T. Calinski and J. Harabasz, “A dendrite method for cluster analysis,” *Communications in Statistics-theory and Methods*, vol. 3, no. 1, pp. 1–27, 1974.
- [11] D. L. Davies and D. W. Bouldin, “A cluster separation measure,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 2, pp. 224–227, 1979.
- [12] S. Dudoit and J. Fridlyand, “A prediction-based resampling method for estimating the number of clusters in a dataset,” *Genome Biology*, vol. 3, no. 7, Article ID research0036.1, 2002.
- [13] H. He and Y. Tan, “A two-stage genetic algorithm for automatic clustering,” *Neurocomputing*, vol. 81, pp. 49–59, 2012.
- [14] “UCI Machine Learning Repository: Data Sets,” 2018, <http://archive.ics.uci.edu/ml/datasets.html>.
- [15] Z. Jing and D. Fu, “Improved k-means algorithm with meliorated initial centers,” *Computer Engineering and Design*, vol. 34, no. 5, pp. 1691–1694, 2013.
- [16] D. Li, J. Shen, and H. Chen, “A fast K-means clustering algorithm based on grid data reduction,” in *Proceedings of the 2008 IEEE Aerospace Conference*, pp. 1–6, Big Sky, MT, USA, 2008.
- [17] Available: http://www.vision.caltech.edu/Image_Datasets/Caltech%20101/.
- [18] Available: <http://archive.ics.uci.edu/ml/index.php>.
- [19] X.-Y. Wang, Y.-W. Li, H.-Y. Yang, and J.-W. Chen, “An image retrieval scheme with relevance feedback using feature reconstruction and SVM reclassification,” *Neurocomputing*, vol. 127, pp. 214–230, 2014.
- [20] C.-C. Chang and C.-J. Lin, “LIBSVM – A Library for Support Vector Machines,” 2018, <https://www.csie.ntu.edu.tw/~cjlin/libsvm/>.
- [21] J. Zhang, L. Sui, L. Zhuo, Z. Li, and Y. Yang, “An approach of bag-of-words based on visual attention model for pornographic images recognition in compressed domain,” *Neurocomputing*, vol. 110, pp. 145–152, 2013.

