Multimedia Systems
DOI 10.1007/s00530-016-0530-9

SPECIAL ISSUE PAPER

Automatic image annotation using fuzzy association rules and decision tree

Zhixin Li[1,2] · Lingzhi Li[1] · Kaobi Yan[1] · Canlong Zhang[1,2]

© Springer-Verlag Berlin Heidelberg 2016

Abstract The problem of sharp boundary widely exists in image classification algorithms that use traditional association rules. This problem makes classification more difficult and inaccurate. On the other hand, massive image data will produce a lot of redundant association rules, which greatly decrease the accuracy and efficiency of image classification. To relieve the influence of these two problems, this paper proposes a novel approach integrating fuzzy association rules and decision tree to accomplish the task of automatic image annotation. According to the original features with membership functions, the approach first obtains fuzzy feature vectors, which can describe the ambiguity and vagueness of images. Then fuzzy association rules are generated from fuzzy feature vectors with fuzzy support and fuzzy confidence. Fuzzy association rules can capture correlations between low-level visual features and high-level semantic concepts of images. Finally, to tackle the large size of fuzzy association rules base, we adopt decision tree to reduce the unnecessary rules. As a result, the algorithm complexity is decreased to a large extent. We conduct the experiments on two baseline datasets, i.e. Corel5k and IAPR-TC12. The evaluation measures include precision, recall, F-measure and rule number. The experimental results show that our approach performs better than many state-of-the-art automatic image annotation approaches.

Keywords Sharp boundary · Fuzzy classification · Automatic image annotation · Fuzzy association rule · Decision tree

✉ Zhixin Li
lizx@gxnu.edu.cn

Lingzhi Li
lilingzhizwc@sina.com

Kaobi Yan
yankaobi@sina.com

Canlong Zhang
clzhang@gxnu.edu.cn

1 Guangxi Key Lab of Multi-source Information Mining and Security, Guangxi Normal University, Guilin 541004, China

2 Guangxi Experiment Center of Information Science, Guilin 541004, China

1 Introduction

With the development of digital imaging and storage technology, the size of image data increases more and more quickly. At the same time, mobile equipment becomes more and more popular in our daily life. So a large amount of image data emerged in mobile equipment is also needed to be indexed and retrieved effectively and efficiently. To solve the problem, there exist two distinct approaches in the literature. In the original text based image retrieval [2, 21], images were annotated manually and retrieved as documents. Since manually annotation is expensive and subjective, text based image retrieval is difficult to deal with the large image database. Therefore, many works turn their focus on content based image retrieval (CBIR). Under this paradigm, images are retrieved by low level visual features, such as color, shape and texture. However, most CBIR systems were not able to describe images with semantic representation automatically, which lead to the notorious semantic gap. As a result, automatic image annotation (AIA) [32] has attracted more and more attention in recent years.

Generally speaking, the approaches of AIA include two-step. First, image segmentation and feature extraction is executed to get the visual feature representation

Published online: 24 September 2016

of images. Second, image classification or annotation is done based on the correlation of visual features and textual words extracted from training images. In most AIA systems, images are represented by global features, block-based local features, or region-based local features. Duygulu et al. [4] propose translation model (TM) to annotate images. This approach learns the correlation between keywords and image regions by computing the joint probability. Then annotation is automatically accomplished by translating regions into keywords. Jeon et al. [10] propose cross-media relevance model (CMRM), which acquires region-based local features to describe each image through image segmentation, assuming an image is described by a small vocabulary of blobs which were clustered by region features. Lavrenko et al. [14] propose similar continuous-space relevance model(CRM), in which the word probabilities are estimated using multinomial distribution and the blob feature probabilities using a non-parametric kernel density. Afterwards, Feng et al. [5] propose multiple Bernoulli relevance model (MBRM), which uses block-based local features to describe each image. In addition, a multiple Bernoulli distribution is used to generate words instead of the multinomial one as in CRM. Zhang et al. [30, 31] propose an image classification framework by leveraging the non-negative sparse coding,low-rank and sparse matrix decomposition techniques. This approach utilize max pooling along with spatial pyramid matching to get the feature vectors to represent images. A linear SVM(Support Vector Machine) classifier is employed for final classification. Furthermore, many approaches accomplish the task of annotation by combining several feature representation methods. For example, to combine global, regional, and contextual features, Wang et al. [29] present an extended approach of CMRM, which incorporate the three kind of representation by estimate their joint probability. Caneiro et al. [1] propose supervised multiclass labeling (SML), which employs optimal principle of minimum probability of error and treats annotation as a multiclass classification problem where each of the semantic concepts of interest defines an image class. Monay et al. [22] propose an approach by modelling multi-modal co-occurrences. The approach is based on probabilistic latent semantic analysis (PLSA) and constrains the definition of latent space to ensure its consistency in semantic words. Moreover, Li et al. [16] propose a hybrid approach combining continuous PLSA and ensembles of classifier chains. This approach can learn semantic classes of images and consider the correlation between labels at the same time. In addition, many kinds of learning technology are used to annotate images, such as neural network, multiclass SVM and the nearest neighbor classifiers etc. A BLIR approach [25] is proposed based on boosting learning. In the stage of feature extraction, 2D-MHMM (2 dimensional Multi-resolution hidden Markov model) is used to get image representation. In the stage of semantic learning, the boosting algorithm is utilized to combine the keywords and the low-level feature for annotating images. Sumathi and Hemalatha [27] proposed an innovative hybrid hierarchical model to annotate images automatically. The approach utilizes low level image features and a simple combination of basic distances to find the nearest neighbor of a given image. Then semantic concepts are assigned to new images by support vector machine (SVM). Zhang et al. [33] propose a new model to extract foreground and background annotation words. In this model, foreground semantic concepts are obtained from visual saliency analysis and multiple Nystrom-approximating kernel discriminant analysis. At the same time, region semantic analysis was used to get annotation words of background. Makadia et al. [18] propose joint equal contribution (JEC) to annotate images. JEC finds the nearest neighbour of a given image by using global low-level features and a combination of basic distance measures. Then it assigns keywords to corresponding image using a greedy label transfer mechanism.

Associative classification is a popular method in data mining field [12]. In addition, it is also successfully applied in image classification [7] and visual concept detection [19, 20]. It is proved that the approach has higher classification accuracy than other famous techniques, such as SVM, Naive Bayes, Neural networks etc. However, these association classification approaches were poor in dealing with the "sharp boundary". Sharp boundary includes fuzzy and uncertain factors commonly exist in image processing, such as edge, boundary and texture, etc. To overcome this problem, Kuok et al. [13] propose a fuzzy technique for mining association rules, which is the origin of fuzzy association rules (FARs). The approach using FARs has been widely applied in data mining. Hence, many approaches based on FARs have been proposed in the field of image classification [3, 28]. However, these approaches have two disadvantages. First, when dealing with the large image database, the size of FARs base becomes very large. Second, since FARs are not optimized adquately, the classification performance is not as good as expectation.

Decision tree based methods are widely used in data mining and decision support applications. The most popular method is based on the algorithm ID3 [24]. In addition, there exist many algorithms including C4.5, CART etc. In this paper, we integrate FARs and decision tree and provide a way to reduce the association rules for annotation. Our approach has several advantages. On the one hand, we reduce fuzzy association rules in the training stage, which determine the degree of correlation between features and concepts of training images. Instead of the traditional discriminative approach which assigns a sample to a class, the FARs assign the sample points to each class with a membership, which aims to handle "sharp boundary" problem.

FREE paper freepaper.me دانلود کننده مقالات علمی

Moreover, in this way, correlation between visual features and semantic concepts can be got intuitively, which could bridge the "semantic gap" to a large degree. On the other hand, the effect on annotating the multi-label images can be demonstrated better than other traditional machine learning methods. In the multi-label image database, we can obtain several rules which assign the same low level feature to multiple semantic concepts. Considering an image with sunset and mountain label, we can get two fuzzy association rules from the image. Finally, to tackle the great fuzzy association rules base, the well-known decision tree algorithm is employed to reduce the number of rules. As a result, the complexity of the algorithm is greatly decreased, while the efficiency of the algorithm is increased. Moreover, decision tree algorithm filters the unnecessary long rules and weak rules from the fuzzy rule base. Consequently, the experimental results show that the performance of image annotation and retrieval is improved obviously, which proves the effectiveness and robustness of our approach.

This paper is organized as follows. Section 2 introduces image segmentation and feature extraction. Section 3 discusses the framework of our proposed approach, which integrates FARs and decision tree to improve the performance of AIA. Section 4 includes experiment design and result analysis. Finally, some concluding remarks are drawn in Sect. 5.

2 Image preprocessing

2.1 Image segmentation

In the field of automatic image annotation, images are generally represented by either global or local features. In our approach, each image is segmented into sub-units so that region-based local features can be obtained from sub-units. Image segmentation is one of the key procedure in image preprocessing. Segmented target areas can be used as units for following feature extraction. The method of segmentation includes edge detection, edge tracing, region growth etc. Our approach use region growth method for image segmentation. This method can segment an image into connected region with similar features and can provide corresponding edge information. The region growth method can be described as following steps. First, finding an appropriate growing point. That is, choosing an element e in matrix M. Second, selecting the growth criterion. We find its neighbors set N that have values within δ appointed by specific algorithm. Then for each element e in N, we find corresponding neighbours that have values within δ. Third, determining the stop condition of the growth. This previous process is repeated until no new neighbors are identified.

$$\begin{bmatrix} 4 & 3 & 7 & 3 & 3 \\ 1 & 7 & (8) & 7 & 5 \\ 0 & 5 & 6 & 1 & 3 \\ 2 & 2 & 6 & 0 & 4 \\ 1 & 2 & 1 & 3 & 1 \end{bmatrix}$$

(a)

$$\begin{bmatrix} 4 & 3 & (7) & 3 & 3 \\ 1 & (7) & (8) & (7) & 5 \\ 0 & 5 & 6 & 1 & 3 \\ 2 & 2 & 6 & 0 & 4 \\ 1 & 2 & 1 & 3 & 1 \end{bmatrix}$$

(b)

$$\begin{bmatrix} 4 & 3 & (7) & 3 & 3 \\ 1 & (7) & (8) & (7) & 5 \\ 0 & 5 & (6) & 1 & 3 \\ 2 & 2 & 6 & 0 & 4 \\ 1 & 2 & 1 & 3 & 1 \end{bmatrix}$$

(c)

$$\begin{bmatrix} 4 & 3 & (7) & 3 & 3 \\ 1 & (7) & (8) & (7) & 5 \\ 0 & (5) & (6) & 1 & 3 \\ 2 & 2 & (6) & 0 & 4 \\ 1 & 2 & 1 & 3 & 1 \end{bmatrix}$$

(d)

Fig. 1 An example of image segmentation using image region growth

Example 1 Figure 1 shows an example of image region growth. The matrix in Fig. 1a represents the original image. The number in the matrix shows the gray value of an image. The pixel whose gray value equals 8 is taken as the initial growing point, named $f(i, j)$. The growth criterion is choosing several points from its 8 neighbors, with the condition that the difference of the gray value of these points and the initial point is not greater than 1. After the first region growth, the matrix in Fig. 1a turns into the matrix in Fig. 1b. Since $f(i-1, j)$, $f(i, j-1)$ and $f(i, j+1)$ all satisfy the growth criterion, they are merged. The matrix in Fig. 1c shows the gray value after the second growing, $f(i+1, j)$ is merged. The matrix in Fig. 1d shows the gray value after the third growing, $f(i+1, j-1)$ and $f(i+2, j)$ are merged. At last, there is no pixels satisfy the growth criterion. Then region growth is stop.

Figure 2 shows several examples of using region growth method to divide an image into homogenous regions. Generally speaking, image segmentation is an advantageous step in image pre-processing. As a result, our approach can acquire semantic concepts in a smaller granularity by image segmentation.

2.2 Feature extraction

Feature extraction is an important step in image annotation. The low-level feature provides the foundation for distinguishing images. Furthermore, it plays an important role for improving the performance of image annotation. After image segmentation, the low-level visual feature can

FREE paper freepaper.me
دانلود کننده مقالات علمی

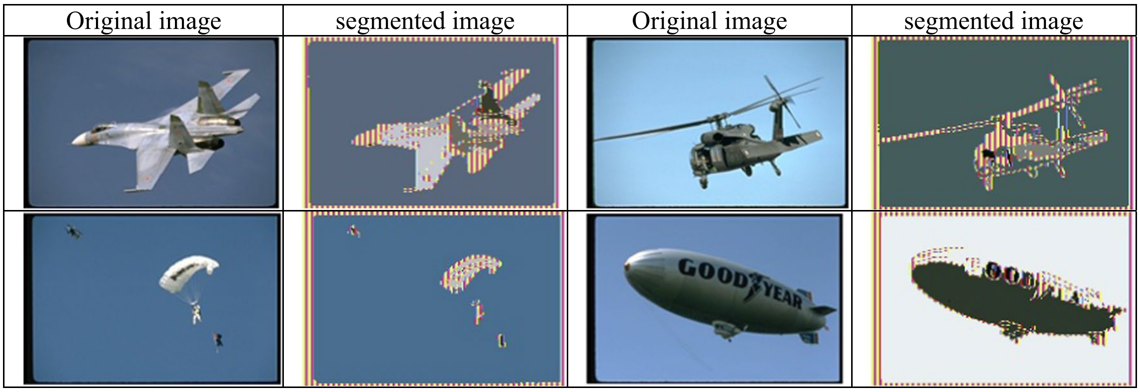


Fig. 2 Original images and corresponding segmented images by region growth method

be extracted from the segmented regions. Feature vectors present different attributes and characteristics of an image. There are many kinds of features for image representation, such as color and texture features defined in the MPEG-7 Standard [15], color coherence vector [23], SIFT feature [17], etc. In this paper, we utilized the color, edge and texture feature for image representation.

Color is one of the most important features of images. There exists many different color spaces, such as RGB, LUV, HSV and HMMD model. Since HSV color model is closely related to human visual perception, we defined color features subject to HSV color space in our approach. After image segmentation, color features are extracted from regions. We encode a histogram to describe the color distribution of image features. It quantizes color space into 9 different bins and counts the frequency of pixels belonging to each color bin.

Textures can capture patterns in the image data roughly and smoothly. It has been well studied in image processing and computer vision. The approach of texture analysis often adopts Gaussian mixture model, Markov random fields or Gabor filter. In our approach, 9-dimensional texture feature vector is acquired through Gabor filter.

Edge features are particularly important for some darker images. Lines in four directions (0, 45, 90 and 135) were extracted using canny edge detection method. Then 4-dimensional feature vector are dedicated for edge features. All the extracted features are combined into a 22-dimensional vector. Figure 3 shows the description of low level image feature.

3 The proposed approach

In general, AIA is a two-step approach. First, image segmentation and feature extraction are executed to get image representation. Second, a learning model is trained to

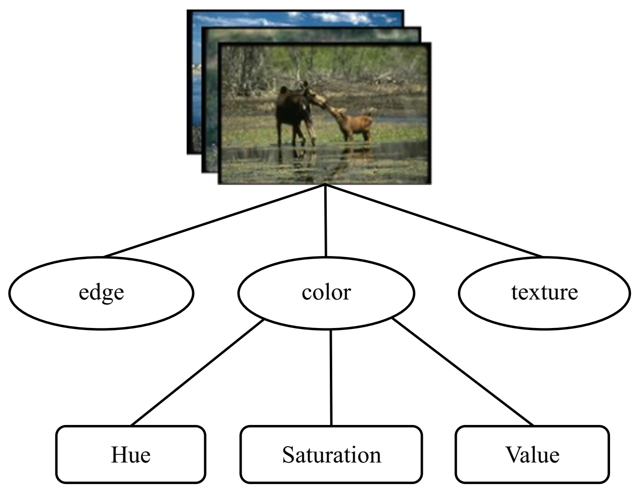


Fig. 3 The low level visual features of images

generate image annotations automatically. After region-based features are extracted in image preprocessing, our approach first divides each attribute of visual features into fuzzy patitions. Then it gets the correlation between visual features and semantic concepts, which forms FARs. Finally, it deals with redundant rules by decision tree, which makes annotation performance much better. Figure 4 shows the framework of our approach, which describes the training and testing stage, illustrating the generative procedure of FARs simultaneously.

3.1 Fuzzy association rules

Our approach includes two stages. In the training stage, low-level features are first extracted. Then FARs base is generated by computing fuzzy support (FS) and fuzzy confidence (FS). Finally, decision tree algorithm is employed to reduce the redundant rules. In the testing stage, annotation is propagated from FARs base. As a result, unlabeled

FREE paper freepaper.me

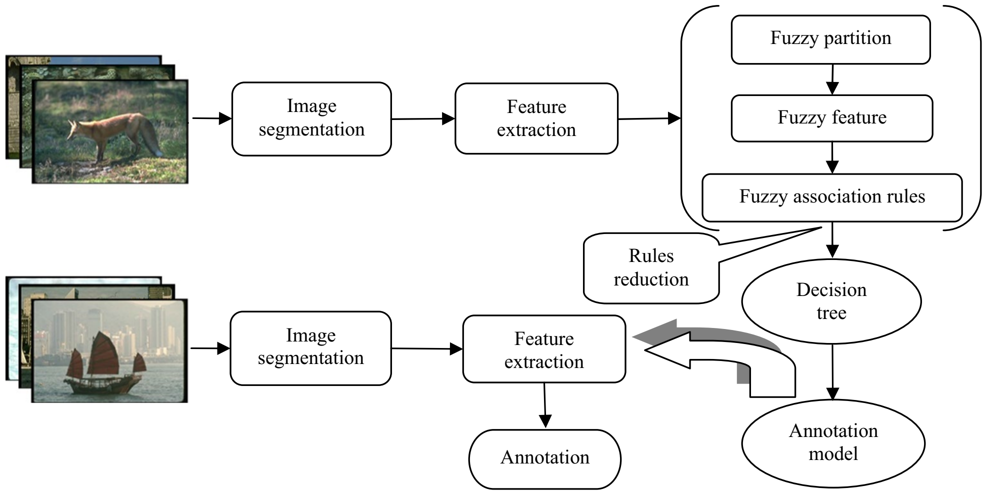


Fig. 4 Framework of automatic image annotation based on fuzzy association rules and decision tree

images could be annotated by several words as accurate as possible.

In this section, we first describe the basic definitions of FARs. Then we discuss how to mine FARs and present how to use FARs for annotation. We consider a two-phase technique to obtain fuzzy rules. The first phase is finding fuzzy feature vector by fuzzy partition according to the original features. The second phase is generating effective FARs from fuzzy feature vector by computing FS and FC.

3.1.1 Fuzzy partition

In the real world, fuzzy and uncertain factors exist everywhere. Just as the procedure of image preprocessing, there exists ambiguity and vagueness in some definition, such as edges, boundaries, regions and textures. Ambiguity and vagueness is very useful in interpreting low-level image processing results. Recently, the fuzzy set theory has been used more and more frequently because of its simplicity and similarity to human reasoning. In this paper, we adopt fuzzy set for FARs method.

Example 2 Taking a person's income level for a simple example, if the value of membership 'high' is 0.1 and the value of membership 'standard' is 0.8, then his income level obviously belongs to 'standard' category more strongly.

Fuzzy set represents human levels categories. People can recognize how strong one belongs to a category according to membership functions which ranging from 0 to 1. The first approach on fuzzy partition by a simple fuzzy grid was employed in [9], where a 2-dimensional pattern space was divided into 22-dimensional fuzzy subspace. The original feature values are replaced with membership functions low, mid, high. The fuzzy feature is described by a 66-dimensional vector. Basically, fuzzy mining algorithms employ membership functions to transform each quantitative low-level visual feature value into a fuzzy set in linguistic terms. The key question of fuzzy partition is how to get the appropriate membership function of the fuzzy set based on numerical attributes.

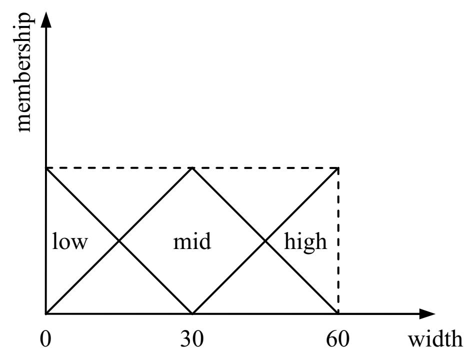


Fig. 5 Triangular membership functions

Figure 5 shows the triangular membership functions of fuzzification. The membership functions stand for three linguistic terms, i.e. low, mid and high. As a result, features belong to classes with membership value in the interval [0, 1], instead of belonging entirely to a certain class.

3.1.2 Mining fuzzy association rules

Given a image database D, low level feature vector $A = \{a_1, a_2, \ldots, a_m\}$, where m is the feature dimension. Fuzzy vector set $A_f = \{A_1^1, A_1^2, \ldots, A_1^{p1}, \ldots, A_2^1, A_2^2, \ldots, A_2^{p2}, \ldots, A_m^1, A_m^2, \ldots, A_m^{pm}\}$, Fuzzy set $F = \{1, 2, \ldots, p_1, 1, 2, \ldots, p_2, \ldots, 1, 2, \ldots, p_m\}$,

FREE paper freepaper.me

Table 1 From original feature vector to fuzzy feature vector

Image ID	Original feature vector		Fuzzy feature vector						Concepts
	a_1	a_2	$Fuzzy(a_1)$			$Fuzzy(a_2)$			
			$u_{low}(a_1)$	$u_{mid}(a_1)$	$u_{high}(a_1)$	$u_{low}(a_2)$	$u_{mid}(a_2)$	$u_{high}(a_2)$	
1	10	33	0.67	0.33	0	0.67	0.33	0	A,C,E,F
2	10	54	0.83	0.17	0	0.17	0.83	0	A,D
3	12	18	0.6	0.4	0	0.4	0.6	0	A,C,E
4	33	54	0	0.9	0.1	0	0.2	0.8	A,B,C
5	51	48	0	0.3	0.7	0	0.4	0.6	B,C,D,F
6	45	46.5	0	0.5	0.5	0	0.45	0.55	A,B,E
7	30	10.5	0	1	0	0.65	0.35	0	C,D
8	57	15	0	0.1	0.9	0.5	0.5	0	A,C
9	45	9	0	0.5	0.5	0	0.45	0.55	A,B,E
10	23.4	39	0.22	0.78	0	0	0.7	0.3	B,D,F
11	27	49.5	0.1	0.9	0	0	0.35	0.65	B,C,D
12	20.1	33	0.33	0.67	0	0	0.9	0.1	A,B,D,F

where the element of fuzzy set p_j is related to a_i. Assuming that $u_{pj}(a_i)$ represents the degree of membership of a_i in the fuzzy set p_j, Table 1 shows the fuzzy process from original feature vector to fuzzy feature vector. Given an original feature vector $A = \{a_1, a_2\}$ consisting of 2-dimensional low level features, the fuzzy feature vectors could be obtained according to the triangular membership function shown in Fig. 5. In this sample, fuzzy set MFs = {low, mid, high}, so both a_1 and a_2 are divided into 3 fuzzy partitions.

The procedure of mining FARs includes two steps.

The first step is computing fuzzy support (FS) and fuzzy confidence (FC). FS and FC are the most common measure of a fuzzy association rule. FS of a rule in the image set D addresses fuzzy percentage of images including both the feature a_i and concept C_j in all images. It measures the significance of the rule. FC presents the accuracy of the rules prediction. It calculates the percentage of images containing both the feature a_i and concept C_j in the total of images containing the feature a_i. A rule in this context is the relationship among transaction items with enough support and confidence.

Example 3 For example, if a rule is "a_i *is low* $\rightarrow C_j$", then

$$FS(a_i \text{ is low} \rightarrow C_j) = \frac{\sum_{a_i \in C_j} u_{low}(a_i)}{|T|}. \tag{1}$$

$$FC(a_i \text{ is low} \rightarrow C_j) = \frac{\sum_{a_i \in C_j} u_{low}(a_i)}{\sum_{a_i \in C} u_{low}(a_i)}. \tag{2}$$

where $|T|$ is total count of transaction in data set D, and C is the set of all concepts. If the FS equals 0.5, it means 50 % images contain the feature a_i and are annotated concept C_j. If the FC equals 0.5, it means the probability of concept C_j in the image set.

Table 2 shows some FARs and their measures extracted from Table 1 based on Eqs. (1) and (2). $A_n^m(m(low, mid, high), n(x_1, x_2))$ is one candidate fuzzy grid, then we can obtain the k fuzzy itemsets according to fuzzy support.

The second step is building FARs base. Algorithm 1 traverses all fuzzy feature vectors, obtaining frequent itemsets according to minimum fuzzy support (mFS). Algorithm 2 defines an procedure to obtain effective FARs. It describes if FC(R) is larger than or equal to the minimum fuzzy confidence (mFC), then R is effective and can be reserved. Our approach obtains FARs based on Apriori algorithm.

Algorithm 1 Obtaining the frequent itemsets F

Input: Training set T, mFS.

Output : frequent itemsets F.

Process:

1. $C_1 \leftarrow$ fuzzy partion (T)
2. $F_1 \leftarrow$ fuzzy frequent 1 itemset
3. ***for*** $k = 2; F_{k-1} \neq \Phi; k{+}{+}$ ***do***
4. $C_k = Candidate(F_{k-1})$;
5. ***for*** each transaction $t \in T$ ***do***
6. ***for*** each candidate $c \in C_k$ ***do***
7. $F_k = Checking(C_k, mFS)$;
8. ***return*** $F = \bigcup_k F_k$

FREE freepaper.me paper دانلود کننده مقالات علمی

Table 2 Fuzzy support and fuzzy confidence of rules

ID	Fuzzy rules	Fuzzy support	Fuzzy confidence
1	a_1 *is low* $\rightarrow A$	0.23	0.88
2	a_1 *is low* $\rightarrow B$	0.05	0.24
3	a_1 *is low* $\rightarrow C$	0.11	0.50
4	a_1 *is low* $\rightarrow D$	0.12	0.54
5	a_2 *is low* $\rightarrow A$	0.20	0.79
6	a_2 *is low* $\rightarrow B$	0	0
7	a_2 *is low* $\rightarrow C$	0.24	0.95
8	a_2 *is low* $\rightarrow D$	0.07	0.27
9	a_1 *is mid* $\rightarrow A$	0.30	0.55
10	a_1 *is mid* $\rightarrow B$	0.34	0.62
11	a_1 *is mid* $\rightarrow C$	0.37	0.68
12	a_1 *is mid* $\rightarrow D$	0.32	0.58
13	a_2 *is mid* $\rightarrow A$	0.34	0.70
14	a_2 *is mid* $\rightarrow B$	0.25	0.51
15	a_2 *is mid* $\rightarrow C$	0.25	0.51
16	a_2 *is mid* $\rightarrow D$	0.29	0.60
17	a_1 *is high* $\rightarrow A$	0.17	0.74
18	a_1 *is high* $\rightarrow B$	0.11	0.48
19	a_1 *is high* $\rightarrow C$	0.18	0.81
20	a_1 *is high* $\rightarrow D$	0.06	0.26
21	a_2 *is high* $\rightarrow A$	0.12	0.48
22	a_2 *is high* $\rightarrow B$	0.16	1
23	a_2 *is high* $\rightarrow C$	0.17	0.68
24	a_2 *is high* $\rightarrow D$	0.14	0.55

Algorithm 2 Obtaining FARs

Input: Frequent itemsets F, mFC.

Output : FARs.

Process:

1. ***for*** every feature attribute $c \in C$ ***do***
2. ***while*** $F_k \neq \Phi$ ***do***
3. ***for*** every frequent itemset $f \in F_k$ ***do***
4. ***if***$(fuzconf(f \rightarrow c) \geq mFC)$ ***then***
5. Output the rule $f \rightarrow c$ with $conf = fuzconf(f \rightarrow c)$;

3.2 Decision tree

The FARs base obtained from the approach is large and complex because the size of the image database is usually large. Many redundant rules exist in FARs base. As a result, the test phase is time costly and low efficiency. Therefore, our main goal is to select rules from the base to construct a compact rule set which can minimize the number of rules. Genetic Algorithms have been proposed to deal with this problem [6, 8]. However, this algorithm is not efficient enough. In our approach, we use the well-known algorithm of decision tree to reduce association rules. To filter the unnecessary long rules and weak rules from the FAR base, the basic idea is to determine how to split the FARs [11].

After the FARs are obtained, the former condition of rule is used to construct the new attribute for the decision tree.

Example 4 For example, a rule is described as $R1 : (a_1\ is\ low,\ a_2\ is\ mid) \rightarrow A$, then $a_1 \bigcap a_2$ is candidate attribute of decision tree.

However, it is doubtful whether the new candidate attribute will be valuable for annotation. Therefore, a further evaluation is proposed in the approach. First, candidate attribute AN_j is selected as root of a decision tree. The information content is described as Eq. (3).

$$G_{root} = -\sum_{j=1}^{k} \frac{|C_j|}{|T|} log_2 \frac{|C_j|}{|T|}. \tag{3}$$

where $|T|$ stands for total count of transaction in data sets, k stands for the count of class, $|C_j|$ is the count of jth class.

For one candidate attribute (that is, when $AN_j = 1$ is satisfied), the number of data records is described as Eq. (4). In these dataset records, the number of data records belonging to C_j is described as Eq. (5).

$$N_1 = |T| \frac{sup(R_i)}{conf(R_i)}. \tag{4}$$

$$N_2 = |T| sup(R_i). \tag{5}$$

The information content of data records which satisfy $AN_j = 1$ is described as Eq. (6). When $AN_j \neq 1$, the number of data records is described as Eq. (7).

$$G_1 = \frac{N_1}{|T|}[-\frac{N_2}{N_1} log_2 \frac{N_2}{N_1}]. \tag{6}$$

$$N_3 = |T| - N_1 = |T| \frac{conf(R_i) - sup(R_i)}{conf(R_i)}. \tag{7}$$

In the data records which satisfy $AN_j \neq 1$, we use N_{3k} to represent the number of data records whose class label is marked $C_k (k = 1, 2, , n)$. Then the information content of this kind of data records is described as Eq. (8).

$$G_2 = \frac{N_3}{|T|}[-\sum_{k=1}^{n} \frac{N_{3k}}{N_3} log_2 \frac{N_{3k}}{N_3}]. \tag{8}$$

To sum up, the information gain of new candidate attribute can be computed by $Gain(AN_j) = G_{root} - G_1 - G_2$. If

دانلود کننده مقالات علمی freepaper.me FREE paper

$Gain(AN_j) > 0$, then attribute AN_j will be still kept in use, otherwise it will be abandoned. The procedure is described in detail as algorithm 3.

Algorithm 3 Reducing FARs according to new rules

Input: FARs.

Output: a decision tree.

Process:

1. Generate new candidate attribute in the FARs;
2. **for** every candidate attribute AN_j **do**
3. **if** $Gain(AN_j) > 0$ **then**
4. Keep AN_j in use;
5. **else** abandon AN_j;
6. Delete the FARs which including the AN_j;

4 Experimental results and analysis

4.1 Datasets

We adopt Corel5k [4] and IAPR TC12 [18] as two baseline image datasets in our experiments. These two datasets are widely used as basic comparative data for recent research in image annotation.

Corel5k image database includes relatively abundant images which covers multiple topics. The database is composed of 50 CDs which are containing 50 semantic topics. Each CD contains 100 images with the same size. The images could be converted into a variety of formats. Each CD represents a semantic topic, such as building, people, bus, elephant, beach, etc. Corel 5k dataset usually divided into three parts. 4000 images constitute a training set. 500 images constitute a validation set which is used to estimate model parameters. The rest 500 images are used to test algorithm performance. After finding optimal model parameters with validation set, the training set of 4000 images and the validation set of 500 images are combined into a new training set consist of 4500 images. Each image in the image database is annotated by 1–5 keywords. Training set containing 374 keywords in total. In addition, 260 keywords are used in test set.

The IAPR TC12 dataset has an extensive application in automatic image annotation and multimedia information retrieval. The dataset includes 20,000 segmented images, including pictures of different sports and actions, photographs of people, animals, cities and many other categories of images. In the feature extraction phase, 99535 feature vectors are extracted from segmented regions. Each feature is composed of color spaces LAB, texture and space location. Each segmented region is assigned a label from a carefully defined vocabulary. The annotation vocabulary is organized according to a conceptual hierarchy. The annotator goes through from top to bottom to look for the best label of each object.

4.2 Parameters setting

In our approach, mFS, mFC and partition number K are the parameters that can affect the annotation performance. To obtain the optimal parameters, we have done experiments with different values of the parameters. Figure 6 shows the annotation precision with different mFS and mFC. Notice that when the mFS and mFC of the rules is 0.05 and 0.85, respectively, the method has the best annotation precision 92.5 %. From the results, we can see that the annotation precision is more sensitive to larger mFS, which indicates that a smaller mFS could be a better choice. On the other hand, the results show that the higher the mFS is, the more robust the annotation performance is. Hence, from the figure we can see the best annotation precision 92.5 % is simultaneously obtained with (mFS, mFC) = (0.05, 0.85). Likewise, we have used these two best parameters in subsequent experiments. We define K as the maximal number of fuzzy partitions in each quantitative attribute.

Figure 7 shows the precision with K varying from 3 to 8. We obtained the best performance when K is 6. That is to say, our experiments divide each quantitative attribute into

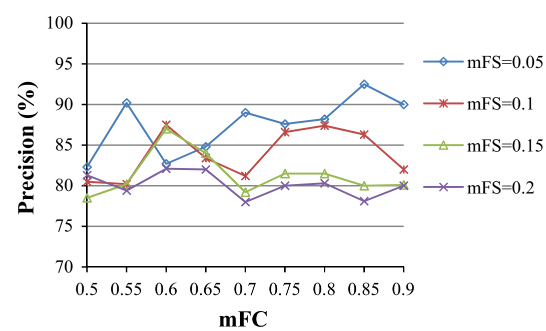


Fig. 6 Precisions of image annotation with different mFS and mFC

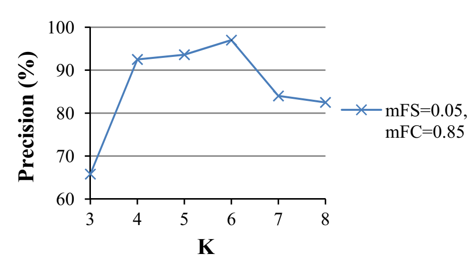


Fig. 7 Precisions of image annotation with different K

FREE paper
freepaper.me

6 fuzzy partitions. Then the fuzzy feature is described by a 22 × 6-dimensional vector.

4.3 Evaluation measures

The algorithms are coded with Visual C++ 2012 and run on Intel 2.66 GHz Pentium4 CPU. The Operating system is Microsoft Windows 7 pro. In our approach, we use 80 % pictures in training part and the remain images are used as testing part. In Corel5k dataset, the segmentation algorithm is employed to segment 4000 images, and totally obtains 42,379 segmented regions. There are 6 largest regions are selected at most. Each region is represented as a 22-dimensional visual feature. Each image is annotated with 3–5 top-ranked keywords. The standard of performance evaluation includes precision, recall and F-measure, which are defined as

$$\text{Precision} = \frac{a}{b}. \tag{9}$$

$$\text{Recall} = \frac{a}{c}. \tag{10}$$

$$\text{F-Measure} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}. \tag{11}$$

In above equations, a is the number of images annotated one certain keyword correctly, and b is the number of images which annotated the same word in the retrieval, and c is the number of all the images with the keyword as its original annotation. Tables 3 and 4 reveal precision, recall and F-measure of our approach on different database. Experimental results show that our approach could annotate testing images effectively. Especially, for the category building on Corel 5 k and the category landscape on IAPR TC12, we obtained an outstanding performance.

4.4 Comparative experiments

Our approach automatically annotates images by combining FARs and decision tree. The approach can mine the correlation between the visual features and semantic keywords of images. Furthermore, it uses decision tree to reduce unnecessary rules and assure the efficency of the algorithms. To show the advantage of our approach, we compare our approach with other state-of-the-art approaches, including TM [4], CMRM [10], CRM [14], MBRM [5], SML [1], JEC [18], PLSA-WORDS [22], HGDM [16] and CFAR [26]. TM, CMRM, CRM, MBRM, SML and PLSA-WORDS are traditional and famous annotation models. HGDM [***8] is a hybrid model. JEC is a relatively new model which acquires high performance. Moreover, CFAR is a semantic annotation model by dealing with fuzzy association rules.

Table 3 Evaluation measures of annotation on Corel5k

Concepts	Precision (%)	Recall (%)	F-Measure
Africa	81.2	76.9	0.85
Beach	87.8	80.2	0.90
Building	93.1	89.2	0.93
Mountain	87.1	86.5	0.89
Elephant	92.6	88.6	0.95
Mean	88.4	84.3	0.90

Table 4 Evaluation measures of annotation on IAPR TC12

Concepts	Precision (%)	Recall (%)	F-Measure
Sports	79.5	75.3	0.82
People	87.6	88.6	0.94
landscape	91.4	90.2	0.88
Animal	88.0	80.5	0.87
Mean	87.0	82.6	0.88

Table 5 Performance comparison of different annotation models on Corel5k dataset

Models	#Words with recall > 0	Results on 49 best words		Results on all 260 words	
		MP	MR	MP	MR
TM	49	0.20	0.34	0.06	0.04
CMRM	66	0.40	0.48	0.10	0.09
CRM	107	0.59	0.70	0.16	0.19
MBRM	122	0.74	0.78	0.24	0.25
SML	137	–	–	0.23	0.29
JEC	139	–	–	0.27	0.32
PLSA-Words	105	0.56	0.71	0.14	0.20
HGDM	137	0.78	0.83	0.28	0.32
Our approach	138	0.81	0.85	0.30	0.35

Table 6 Performance comparison of different annotation models on IAPR TC12 dataset

Models	#Words with recall > 0	MP	MR
MBRM	186	0.21	0.14
JEC	196	0.25	0.16
PLSA-Words	177	0.18	0.12
HGDM	194	0.29	0.18
Our approach	199	0.32	0.21

FREE paper freepaper.me دانلود کننده مقالات علمی

For comparison, experimental results on Corel5k are shown in Table 5. We compare the mean precision (MP) and mean recall (MR) in 49 best keywords and all 260 keywords in testing set. From the data on the 49 best keywords set, we can see that the MP of our approach is higher than all other approaches. Also, our approach has best annotation performance on total 260 keywords. From the experimental results, we can conclude that the performance of our approach is superior to these state-of-the-art approaches.

Table 6 shows the experimental results on IAPR TC12. The evaluation measures includes the number of keywords with recall >0, MP and MR. From the table we can see that our approach acquire better performance than all other approaches, which proves that our approach has good robustness when the size of image database become large.

Table 7 adds the comparison with SVM and CFAR. From the table we can conclude the precision of annotation approach based on fuzzy association rules are higher than the non-fuzzy association rules approach. Furthermore, our approach has better effectiveness than CFAR in most case. The small rule number indicates the decision tree performs excellent in reducing redundant rules.

In summary, the experimental results on the datasets of Corel5k and IAPRTC12 indicate that our approach is fairly stable with respect to its parameters. Furthermore, since our approach utilizes fuzzy association rules to relieve the influence of sharp boundary problem and employs decision tree to handle redundant association rules, it acquires higher precision and better effectiveness than many state-of-the-art approaches.

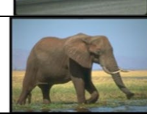
4.5 Examples of automatic image annotation

We select 3–5 top-ranked keywords to annotate the images from Corel 5k dataset. Figure 8 gives some annotation examples, including manual annotation and the annotation generated by our approach. It can be seen from the figure that the semantic concepts recognized by our approach are quite precise. Moreover, the keywords generated by our

Table 7 Mean precision and rule number of SVM, CFAR and our approach

Concepts	Dataset	Image number	SVM precision	CFAR		Our approach	
				Precision	Rule number	Precision	Rule number
Sky	Corel5k	100	0.85	0.83	12	0.83	9
Sea	Corel5k	56	0.76	0.82	7	0.83	5
People	Corel5k	78	0.80	0.81	9	0.82	7
Tree	Corel5k	90	0.69	0.77	6	0.80	6
Flower	Corel5k	30	0.69	0.75	4	0.77	5
Sports	IAPR TC12	86	0.71	0.80	10	0.82	8
People	IAPR TC12	55	0.70	0.74	7	0.80	6
Landscape	IAPR TC12	90	0.71	0.80	9	0.79	6
Animal	IAPR TC12	60	0.72	0.82	8	0.80	7

Fig. 8 Examples of automatic image annotation on Corel5k

Image	Our approach	Manual annotation	Image	Our approach	Manual annotation
	boat, sky, buildings, water	boat, city, buildings, water		cars, road, buildings	cars, grass, road, buildings
	tiger, snow, stone	tiger, snow		elephant, grass, water, sky	elephant, water, grass, sky
	sun, water, seabeach	sun, water, seabeach, sky		sky, water clouds, sidebeach	sky, water, clouds, sidebeach, waves
	birds, tree, branchleaf	birds, branchleaf		snow, mountain, stone, sky	snow, mountain, stone, sky

دانلود کننده مقالات علمی FREE freepaper.me paper

approach are mostly plausible even it is not exists in manual annotated keywords.

5 Conclusions

In this paper, we have proposed an approach which integrates fuzzy association rules with decision tree algorithm to accomplish the task of automatic image annotation. In our approach, fuzzy association rules have been extracted to get the correlation between low level features and semantic concepts. However, since image database is very large in the real world, we adopt decision tree algorithm to process the large fuzzy association rules base, which can improve the performance of image annotation greatly. The experimental results demonstrate that our approach can derive the accurate rules under large database. On one hand, the difference between our approach and other traditional association rules is that we add a process to handle association rules. On the other hand, the difference between our approach and other machine learning method of annotation (SVM, boosting, neutral network, etc.) is that we get the association between visual features and semantic concepts intuitively. Our approach is independent of specific hardware and operating system. After optimization, it can be applied in not only personal computer but also mobile equipment. In our future work, we will focus on selecting a more effective membership function for feature partition and a more proper criterion to calculate the fuzzy association rules.

Acknowledgments This work is supported by the National Natural Science Foundation of China (Nos. 61663004, 61262005, 61363035, 61365009, 61165009), the Guangxi Natural Science Foundation (2013GXNSFAA019345, 2014GXNSFAA118368), the National Basic Research Program of China (No. 2012CB326403) and the Guangxi “Bagui Scholar” Teams for Innovation and Research Project, Guangxi Collaborative Innovation Center of Multi-source Information Integration and Intelligent Processing.

References

1. Carneiro, G., Chan, A.B., Moreno, P.J., Vasconcelos, N.: Supervised learning of semantic classes for image annotation and retrieval. IEEE Trans. Patt. Anal. Mach. Intell. **29**(3), 394–410 (2007)
2. Chang, S.K., Hsu, A.: Image information systems: where do we go from here? IEEE Trans. Knowl. Data Eng. **4**(5), 431–442 (1992)
3. Dong, J., Sheng, G.: Remote sensing image classification based on fuzzy association classification. J. Comp. Res. Dev. **49**(7), 1500–1506 (2012)
4. Duygulu, P., Barnard, K., de Freitas, N., Forsyth, D.: Object recognition as machine translation: learning a lexicon for a fixed image vocabulary. In: Proceedings of the 7th European Conference on Computer Vision, pp. 97–112 (2002)
5. Feng, S.L., Manmatha, R., Lavrenko, V.: Multiple bernoulli relevance models for image and video annotation. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1002–1009 (2004)
6. F.P. Pach, J.A.: Selecting fuzzy if-then rules for classification problems using genetic algorithms. International Scholarly and Scientific Research and Innovation **2**(1), 546–551 (2008)
7. Hu, Y., Chen, R., Tzeng, G.: Mining fuzzy association rules for classification problems. Comp. Ind. Eng. **43**(4), 735–750 (2002)
8. Ishibuchi, H., Nozaki, K.: Selecting fuzzy if-then rules for classification problems using genetic algorithms. IEEE Trans. Fuzzy Syst. **3**(3), 260–270 (1995)
9. Ishibuchi, H., Nozaki, K., Tanaka, H.: Distributed representation of fuzzy rules and its application to pattern classification. Fuzzy Sets Syst. **52**(1), 21–32 (1992)
10. Jeon, J., Lavrenko, V., Manmatha, R.: Automatic image annotation and retrieval using cross-media relevance models. In: Proceedings of the 26th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 119–126 (2003)
11. Jing, G., Zhao, B.: New method about how to construct decision tree based on association rule. In: Proceedings of International Workshop on Open-Source for Scientific Computation, pp. 131–135 (2011)
12. Jukic, N., Nestorov, S.: Comprehensive data warehouse exploration with qualified association rule mining. Decision Supp. Syst. **42**(2), 859–878 (2006)
13. Kuok, C.M., Fu, A., Wong, M.H.: Mining fuzzy association rules in databases. ACM SIGMOD Record **27**(1), 41–46 (1998)
14. Lavrenko, V., Manmatha, R., Jeon, J.: A model for learning the semantics of pictures. Adv. Neural Inform. Process. Syst. **16**, 553–560 (2003)
15. Li, F.F., Fergus, R., Perona, P.: One-shot learning of object categories. IEEE Trans. Patt. Anal. Mach. Intell. **28**(4), 594–611 (2006)
16. Li, Z., Shi, Z., Zhao, W., Li, Z., Tang, Z.: Learning semantic concepts from image database with hybrid generative/discriminative approach. Eng. Appl. Art. Intel. **26**(9), 2143–2152 (2013)
17. Lowe, D.: Distinctive image features form scale-invariant keypoints. Int. J. Comp. Vision **60**(2), 91–110 (2004)
18. Makadia, A., Pavlovic, V., Kumar, S.: Baselines for image annotation. Int. J. Comp. Vision **90**(1), 88–105 (2010)
19. Mangalampalli, A., Pudi, V.: Fuzzy association rule mining algorithm for fast and efficient performance on very large datasets. In: Proceedings of IEEE International Conference on Fuzzy Systems, pp. 1163–1168 (2009)
20. Mangalampalli, A., Pudi, V.: A fuzzy associative classification approach to visual concept detection. Int. J. Uncertain. Fuzz. Knowl.-Based Syst. **22**(3), 429–452 (2014)
21. Markkula, M., Sormunen, E.: End-user searching challenges indexing practices in the digital newspaper photo archive. Inform. Retriev. **1**(4), 259–285 (2000)
22. Monay, F., Gatica-Perez, D.: Modeling semantic aspects for cross-media image indexing. IEEE Trans. Patt. Anal. Mach. Intel. **29**(10), 1802–1817 (2007)
23. Pass, G., Zabith, R.: Histogram refinement for content-based image retrieval. In: Proceedings of the 3rd IEEE Workshop on Applications of Computer Vision, pp. 96–102 (1996)
24. Quanlan, J.: C4.5: programs for machine learning. Morgan Kaufmann, San Francisco, CA, USA (1993)
25. Ru, L., Ma, S., Lu, J.: Boosting-based automatic linguistic indexing of pictures. J. Image Graph. **11**(4), 486–491 (2006)
26. Silaa, C.N., Freitas, A.A.: A survey of hierarchical classification across different application domains. Data Mining Knowl Discov **22**(1–2), 31–72 (2011)
27. Sumathi, T., Hemalatha, M.: An innovative hybrid hierarchical model for automatic image annotation. In: P.V. Krishna, M.R.

FREE paper freepaper.me دانلود کننده مقالات علمی

Babu, E. Ariwa (eds.) Global Trends in Information Systems and Software Applications, vol. 270. Springer (2012)

28. Tazaree, A., Eftekhari-Moghadam, A., Sajadi-Ghaem-Maghami, S.: A semantic image classifier based on hierarchical fuzzy association rule mining. Multi. Tools Appl. **69**(3), 921–949 (2014)
29. Wang, Y., Mei, T., Gonga, S., Huang, X.S.: Combining global, regional and contextual features for automatic image annotation. Patt. Recogn. **42**(2), 259–266 (2009)
30. Zhang, C., Cheng, J., Liu, J., Pang, J., Liang, C., Huang, Q., Tian, Q.: Object categorization in sub-semantic space. Neuro-Computing **142**, 248–255 (2014)
31. Zhang, C., Liu, J., Tian, Q., Xu, C., Lu, H., Ma, S.: Image classification by non-negative sparse coding, low-rank and sparse decomposition. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1673–1680 (2011)
32. Zhang, D., Islam, M.M., Lu, G.: A review on automatic image annotation techniques. Patt. Recogn. **45**(1), 346–362 (2012)
33. Zhang, J., Hu, W.: Effective multi-modal multi-label learning for automatic image annotation. In: Proceedings of the 9th International Conference on Fuzzy Systems and Knowledge Discovery, pp. 1216–1220 (2012)

FREE paper
freepaper.me
دانلود کننده مقالات علمی