Performance Analysis of Improved Affinity Propagation Algorithm for Image Semantic Annotation

Dong Yang, Ping Guo*

Image Processing and Pattern Recognition Laboratory,
Beijing Normal University, Beijing 100875, China
{d.yang, pguo}@ieee.org

Abstract. In an image semantic annotation system, it often encounters the large-scale and high dimensional feature datasets problem, which leads to a slow learning process and degrading image semantic annotation accuracy. In order to reduce the high time complexity caused by redundancy information of image feature dataset, we adopt an improved affinity propagation (AP) algorithm to improve annotation by extracting and re-grouping the repeated feature points. The time consumption is reduced by square of repetition factor. The experiments results illustrate that the proposed annotation method has excellent time complexity and better annotation precision compared with original AP algorithms.

Keywords: Information retrieval, Image semantic annotation, Clustering, Affinity propagation.

1 Introduction

For semantic annotation of natural images or human-activity images, in order to describe complex and elaborated image semantics, as many as possible images are required in most existing annotation systems. Because of both the image redundancy and image features redundancy, such as overlapping sampling and repeated emergence of similar image regions, it is very common there are a lot of repeated or near same image feature samples before applying vector quantization (VQ) techniques [15] to optimize training data in large datasets.

For unsupervised annotation, each semantic label is considered as a variable. The annotation is treated as a joint probability modeling problem, with clustering representations of image features or words, such as N-cut based method [10] and cross media relevance models [11]. For supervised annotation, each semantic label is considered as a class. The annotation is regarded as a classification problem, which adopts clustering to get sparse representation or posterior probabilistic distribution of image features for each class, and some supervised learning techniques such as the divide and conquer strategy-based scene classifier [12], supervised multi-class

* Corresponding author.

labeling algorithm [13], and multi-class annotation using optimized training data [4] are applied to annotate the images.

In both unsupervised and supervised annotation systems, clustering has become an important process to handle the huge number feature samples. Affinity propagation (AP) clustering algorithm has been validated powerful for image categorization and annotation [1][2][4], because of its excellent performance, such as automatically determining cluster number, and using similarity of data pairs instead of data values. Recently, weighted AP (WAP) [5][7] and AP with repeated points (APRP) [8] are developed to process large dataset with repeated points.

Time consumption problem for large dataset has been studied, including following three aspects: (1) introducing prior knowledge such as sparse similarity matrix [6]; (2) divide-and-conquer strategy such as hierarchical method [7], partition method[9], and sampling techniques [3]; (3) vector quantization techniques [14], which is powerful especially for large and high-dimensioned data.

In this paper, we study the problem to reduce time complexity based on vector quantization and APRP algorithm and analyze the performance of the algorithm for supervised image annotation.

2 Image Annotation using APRP

Image semantic annotation can be regarded as a multi-class classification problem, which maps image features to semantic class labels, through the procedures of image modeling and image-semantic classification [4]. For image modeling, APRP is adopted instead of original AP (OAP) and WAP to process image feature dataset with repeated points. The algorithm flowchart is illustrated as in Fig. 1.

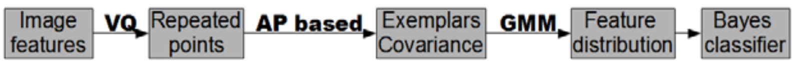

Fig. 1. The algorithm flowchart

2.1 OAP, WAP and APRP

Brief reviews of OAP, WAP and APRP are given according to the literatures [1], [7] and [8], respectively. Repeated point is usually represented as ($\mathbf{x}_i$, n_i), with datum $\mathbf{x}_i$ and its repetition factor n_i [7].

AP algorithm is a graph-based message-passing clustering algorithm. Each data vector is viewed as a point in the graph, and real-value messages are recursively transmitted along edges of the graph until a relatively small number of exemplars and corresponding clusters emerge. The similarity $s(i, k)$ is the negative distance square between datum i and j, and the self similarity $s(k, k)$ is called preference.

The responsibility $r(i, k)$, which is sent from datum i to potential exemplar k, reflects the accumulated evidence for how appropriate datum k is the exemplar of datum i, considering other potential exemplars of datum i. The availability $a(i, k)$, which is sent from datum i to potential exemplar k, reflects the accumulated evidence for how appropriate it would be for datum i to choose datum k as its exemplar, considering the support from other data that datum k should be an exemplar.

The exemplar is determined by combing the availability and responsibility.

$$e = \max_{k}\{a(i,k) + s(i,k)\}. \tag{1}$$

The main differences of OAP, WAP and APRP are at the message-passing, similarity and preference setting. These differences lead to that WAP and APRP treats repeated points as one point thus to reduce the time complexity, while OAP does not consider the repeated point problem.

Frey and Dueck [1] set the initial $a(i, k)$ as 0, then the $r(i, k)$ and a(i, k) are iteratively updated:

$$r(i,k) \leftarrow s(i,k) - \max_{k' \text{ s.t. } k' \neq k}\{a(i,k') + s(i,k')\}. \tag{2}$$

$$a(i,k) \leftarrow \min\left\{0, r(k,k) + \sum_{i' \text{ s.t. } i' \notin \{i,k\}} \max\{0, r(i',k)\}\right\}. \tag{3}$$

$$a(k,k) \leftarrow \sum_{i' \text{ s.t. } i' \neq k} \max\{0, r(i',k)\}. \tag{4}$$

OAP does not refer to the repetition factor n_i of the datum i, therefore there should be n_i copies of $\mathbf{x}_i$ in the dataset for OAP.

Zhang *et al*. [7] changed similarity and preference as:

$$s'(i,j) = n_i s(i,j), \quad i \neq j. \tag{5}$$

$$s'(i,i) = s(i,i) + (n_i - 1)\varepsilon_i, \quad \varepsilon_i \geq 0. \tag{6}$$

The formula (5) expresses the repetition factor n_i of the datum i by changing the similarities from datum $\mathbf{x}_i$ to all other data. However, if n_i grows large, i tends to be dissimilar with all other data, which makes $\mathbf{x}_i$ tends to be an isolated point.

Yang and Guo [8] set the preference similar to formula (6) and change the objective function as:

$$S(c) = \sum_{i=1}^{N} n_i s(i,c_i) + \sum_{k=1}^{N} \delta_k(c). \tag{7}$$

Where c is the clustering result that maps datum i to its exemplar, $\delta(c)$ is a function that will produce a large value if an exemplar does not choose itself as its exemplar, and $S(c)$ is the summation of the similarity of all data and their exemplars, adding the penalty factor $\delta(c)$.

The update rule of availability is changed as:

$$a(i,k) \leftarrow \min\left\{0, r(k,k) + \sum_{i' \text{ s.t. } i' \notin \{i,k\}} \max\{0, n_{i'} r(i',k)\}\right\}. \tag{8}$$

$$a(k,k) \leftarrow \sum_{i' \text{ s.t. } i' \notin \{i,k\}} \max\{0, n_{i'} r(i',k)\}. \tag{9}$$

The formula (8) and (9) considers the repetition factor n_i because of the summation operation. APRP does not change the similarity because the similarity is a distance-like measure, which is not influenced by repetition factor.

2.2 Applying APRP to large dataset

Each image is divided into sub-blocks of 8×8 pixels with overlapping margin. For each sub-block, discrete cosine transform is applied on each color channel, the coefficient values are quantized and zigzag scanned from left top to right down. The first 16 values are selected as the feature vector of this color channel. Then the three vectors from three color channels are concatenated as a 48-dimensional feature vector of the sub-block. Vector quantization algorithm [15] is adopted on the vector group, while the codebooks are trained once in the whole dataset and are fixed in each quantization process.

The training set of each class label is composed with feature points from training images that has this label. APRP algorithm is adopted to get exemplars with automatically determined cluster number for each training set. The image feature distribution of each class label is estimated in the form of Gaussian mixture model (GMM) [4].

$$\{\mathbf{e}_k\} = \mathrm{aprp}(\{\mathbf{x}\}), \quad P(\mathbf{x} \mid \boldsymbol{\mu}, \Sigma, \pi) = \sum_{k=1}^{K} \pi_k N(\mathbf{x} \mid \boldsymbol{\mu}_k, \Sigma_k), \tag{10}$$

where $\{\mathbf{x}\}$ is the set of 48-dimensional feature vectors. Each feature vector is assigned an exemplar by APRP algorithm, which means all the feature vectors are clustered into K clusters. Each cluster $\{\mathbf{x}_k\}$ has one exemplar $\boldsymbol{\mu}_k=\mathbf{e}_k$ as the mean vector. The covariance matrix is $\boldsymbol{\Sigma}_k=\mathrm{cov}(\{\mathbf{x}_k\})$ and cluster weight is $\pi_k=\mathrm{num}(\{\mathbf{x}_k\})$.

2.3 Image semantic annotation

Bayesian classifier is adopted for image semantic annotation. The classifier is trained using training images which are manually pre-annotated with single class label.

For each class label c, the re-grouping process collects repeated points from every image of this label c and assign them into groups. Then APRP algorithm is applied on these groups, and the class distribution $P(\mathrm{x}|c,\boldsymbol{\mu}_c,\boldsymbol{\Sigma}_c,\pi_c)$ is computed hierarchically using the GMM image modeling method.

For annotating a test image **I**, the Bayesian decision rule is adopted. For a given class c, the probability that the test image belongs to this class is the product of the probabilities that the image feature samples **x** belong to this class.

$$P(\mathbf{I} \mid c, \boldsymbol{\mu}_c, \Sigma_c, \pi_c) = \prod_{\mathrm{x} \in \mathrm{I}} P(\mathbf{x} \mid c, \boldsymbol{\mu}_c, \Sigma_c, \pi_c) \cdot \tag{11}$$

By computing all class-conditional distributions $P(\mathbf{I}|c)$, the semantic annotation results for this image **I** can be obtained with the labels whose posterior probabilities $P(c|\mathbf{I})$ are the top several large values [4].

2.4 Performance analysis of APRP

Image annotation system usually processes many similar image regions, which increase the probability of forming repeated points after vector quantization of a large number of feature points. On consider this fact, it is expected that APRP-based annotation algorithm works well with the dataset with repeated points.

Assuming the number of feature points is M and the number of repeated points is N, we count the repeated points as one point. N is less than the codebook length of vector quantization. The relationship between M and N is:

$$M = \sum_{i=1}^{N} n_i \cdot \tag{12}$$

Where n_i is the repetition factor of point i, M/N is the repetition factor of the whole group of feature points.

In figure 2, the influence of repetition factor on clustering is illustrated. The distance criterion is the sum of similarities from data to their exemplars, and the sum of similarities from exemplars to their center, which is used to prevents that the cluster number becoming too large.

$$f_{\mathrm{distance}} = \sum_{i=1}^{N} |\mathbf{x}_i - \mathbf{m}_i| + \sum_{i=1}^{S} |\mathbf{m}_i - \overline{\mathbf{m}}| \cdot \tag{13}$$

When the M/N increases from 1.67 times to 3.00 times, the increase of similarity sum of WAP increase almost one time. The APRP performs the best among three algorithms, and when repetition factor is more than 2.00, the performance of APRP is stable compared with WAP and OAP.

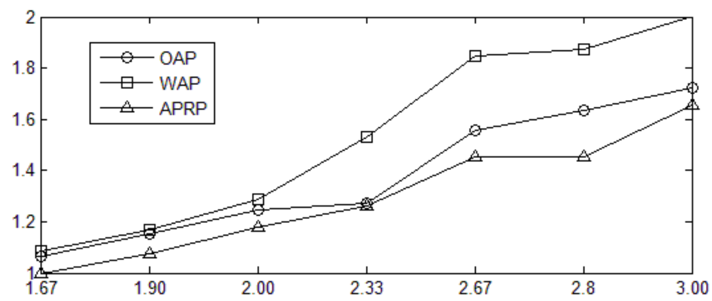

Fig. 2. The repetition factor vs. corresponding distance criterion of clustering

APRP algorithm greatly reduces the time consumption. AP algorithm repeatedly processes every copy of the repeated points. AP algorithm has the loop of comparing each point with other points, and its time consumption is M^2/N^2 times of APRP's time consumption.

3 Experiments

The annotation of natural images has become a difficult problem but a valuable benchmark to validate the image annotation algorithm. The performance of the proposed algorithm is evaluated in both the image modeling and annotation stages. The different image annotation method is compared with adopting the following clustering algorithms separately: OAP[1], WAP[7], APRP[8].

3.1 Dataset and criteria

The images are selected from databases [18] and [19]. We select a subset of 2000 images which contain seven labels: building, car, crossroad, grass, plant, road and sky. There are 628 single-label images (we name it DBS), and 1372 multiple-label images (we name it DBM). One multiple-label image usually contains two or three labels. Using these images, we set up two training and testing scenarios, as in table 1.

Table 1. Two training and testing scenarios.

	Training	Testing
Scenario1	2/3 DBS	1/3 DBS
Scenario2	DBS	DBM

Each image is divided into some sub-blocks with size of 8×8 pixels, and the adjacent blocks overlap 2 pixels. We resize the image with the resolution 300~500×300~500. A fixed code-vector number of 1000 is set up for vector quantization of each image.

The image modeling criteria are sum of point-to-centroid distance [17], and the logarithm of likelihood [16], which are shown in formula (13) and (14), respectively.

$$f_{\text{likelihood}} = \log P(\mathbf{I}|\boldsymbol{\mu},\Sigma,\pi) = \sum_{\mathbf{x}\in\mathbf{I}}^{N} \log P(\mathbf{x}|\boldsymbol{\mu},\Sigma,\pi). \qquad (\mathbf{14})$$

The image annotation criteria are average recall and precise. For a given semantic class, we assume that there are w_h human annotated images w_{auto} computer annotated images in the test set, of which w_c are correct, the recall and precision are defined as following:

$$recall = \frac{w_c}{w_h}, \quad precise = \frac{w_c}{w_{auto}}. \qquad (\mathbf{15})$$

3.2 Experiment result analysis

In table 2, we compare the time consumption of annotating one image, OAP-based method is several times of that of WAP or APRP. The total annotation time in table 2 contains clustering time. The time consumption is reduced by M^2/N^2 times. For example, if the repetition factor M/N is 3, the time consumption of OAP is nearly 9 times of that of WAP or APRP.

Table 2. Time consumption of annotating one image.

Time (s)	OAP	WAP	APRP
DBS	331.0	54.3	56.3
DBM	487.9	77.2	78.6

In table 3 and table 4, we compare the image modeling results of the three annotation systems, where the APRP-based method performs similarly even a little bit better than OAP, and much better than WAP.

Table 3. Sum of point-to-centroid distance of image modelling in DBS and DBM.

f_{distance}	OAP	WAP	APRP
DBS	24872	33515	24362
DBM	42238	48993	42626

Table 4. Logorithm of likelihood of image modelling in DBS and DBM.

$f_{\text{likelihood}}$	OAP	WAP	APRP
DBS	-48.4	-56.6	-47.7
DBM	-52.1	-73.6	-56.1

In table 5, we compare the image annotation result of the three methods. In scenario 1, we use single-label testing images to simulate image categorization tasks;

in scenario 2, we use multiple-labels testing images to simulate image annotation tasks.

Table 5. Image semantic annotation result.

Annotation	Scenario 1			Scenario 2		
	OAP	WAP	APRP	OAP	WAP	APRP
Recall	85.7	75.3	88.5	74.5	67.8	75.2
Precise	73.5	69.4	75.0	61.2	59.1	60.9

The proposed APRP algorithm improves the accuracy of image annotation than WAP does, and it performs close to even higher than OAP. Considering the time consumption of OAP is several times of that of APRP, the proposed algorithm performs the best among the three algorithms for image annotation on large dataset.

4 Conclusion

On considering there exists redundancy information in image feature dataset, the performance of improved AP algorithm for image semantic annotation is analyzed in this paper. The redundancy information usually appears to be repeated points in the large feature datasets after vector quantization, we propose to adopt improved AP algorithm to solve this problem. The image modeling accuracy is improved and time consumption in annotation is greatly reduced with APRP algorithm. The annotation precision approaches or even outperforms than the original AP algorithm, and is much better than that of WAP algorithm. For the case of repetition factor increases because of vector quantization, the performance of APRP algorithm is more stable compared with OAP and WAP. The proposed algorithm is promising on the effectiveness and response speed of the image semantic annotation.

Acknowledgments. The research work described in this paper was fully supported by the grants from the National Natural Science Foundation of China, (Project No. 90820010, 60911130513)

References

1. Frey, B.J., Dueck, D.: Clustering by Passing Messages between Data Points. Science 315, 972--976 (2007)
2. Dueck, D., Frey, B.J.: Non-metric Affinity Propagation for Unsupervised Image Categorization. In: IEEE International Conf. on Computer Vision, pp. 1--8. IEEE Press, New York (2007)
3. Sun, C.Y., Wang, C.H., Song, S., Wang, Y.F.: A Local Approach of Adaptive Affinity Propagation Clustering for Large Scale Data. In: IEEE International Joint Conf. on Neural Networks, pp. 161--165. IEEE Press, New York (2009)
4. Yang, D., Guo, P.: Image Modeling with Combined Optimization Techniques for Image Semantic Annotation. Neural Comput. Appl. (2011) (In press)

5. Furtlehner, C., Sebag, M., Zhang, X.L.: Scaling Analysis of Affinity Propagation. Phys. Rev. E 81(6), 006102 (2010)
6. Xiao, J.X., Wang, J.D., Tan, P., Quan, L.: Joint Affinity Propagation for Multiple View Segmentation. In: IEEE International Conf. on Computer Vision, pp. 1--7. IEEE Press, New York (2007)
7. Zhang, X.L., Furtlehner, C., Sebag, M.: Data Streaming with Affinity Propagation. In: Machine Learning and Knowledge Discovery in Databases, pp. 628--643. Springer, Heidelberg (2008)
8. Yang D., Guo P.: Improvement of Affinity Propagation Algorithm for Large Dataset, Workshop of the Cognitive Computing of Human Visual and Auditory Information (2010)(in Chinese) .
9. Zhang, X.Q., Wu, F., Zhuang, Y.T.: Clustering by Evidence Accumulation on Affinity Propagation. In: IEEE International Conf. on Pattern Recognition, pp. 1--4. IEEE Press, New York (2008)
10. Barnard, K., Duygulu, P., Forsyth, D., de Freitas, N., Blei, D.M., Jordan, M.I.: Matching Words and Pictures. J. Mach. Learn. Res. 3, 1107--1135(2003)
11. Jeon, J., Lavrenko, V., Manmatha, R.: Automatic Image Annotation and Retrieval using Cross-Media Relevance Models. In: ACM International Conf. on Research and Development in Information Retrieval, pp. 119--126. ACM Press, New York (2003)
12. Luo, J., Savakis, A.: Indoor VS Outdoor Classification of Consumer Photographs using Low-Level and Semantic Features. In: IEEE International Conf. on Image Processing, pp. 745--748. IEEE Press, New York (2001)
13. Carneiro, G., Chan, A.B., Moreno, P.J., Vasconcelos, N.: Supervised Learning of Semantic Classes for Image Annotation and Retrieval. IEEE Trans. on Pattern Anal Mach Intell. 29, 394--410(2007)
14. Lin, S., Yao, Y., Guo, P.: Speed Up Image Annotation Based on LVQ Technique with Affinity Propagation Algorithm. In: International Conf. on Neural Information Processing, . Springer, Heidelberg , 533-540 (2010)
15. Linde, Y., Buzo, A., Gray, R.M.: An Algorithm for Vector Quantizer Design. IEEE Trans. on Commun. 28(1), 84--95(1980)
16. Bishop, C.M.: Pattern Recognition and Machine Learning. Springer, Heidelberg, Chp.9, Sec.3(2006)
17. Theodoridis, S., Koutroumbas, K.: Pattern Recognition, 3rd ed. Academic Press, Salt Lake City USA, Chp.11, Sec.2 (2006)
18. Visual Object Classes, http://pascallin.ecs.soton.ac.uk/challenges/VOC/voc2010/
19. IGPR Images, http://igpr.bnu.edu.cn/~dyang/imageset/